

DISCLAIMER

ULTIME AVVERTENZE PRIMA DELLA RIVOLUZIONE

DIMENSIONI E SFIDE
DELLA CYBER SOCIAL SECURITY

GRUPPO N°1

1. Davide Pio Posa
2. Francesca Mantini
3. Nicola Balzano
4. Alessio Carbonara
5. Giuseppe Avitto
6. Teresa Catalano

KNOWLEDGE PARTNER

Problemi

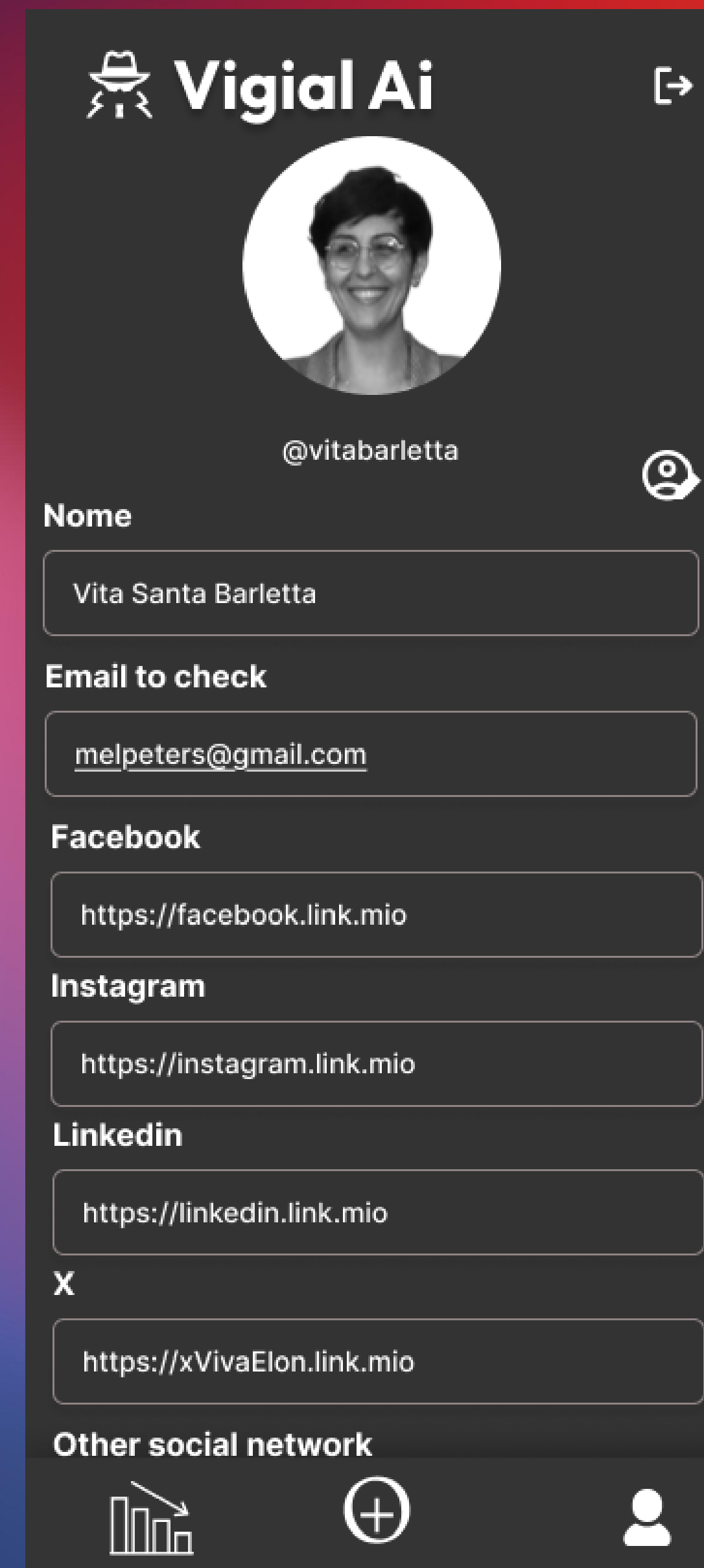
Un pericolo invisibile, ogni giorno più vicino

- L'inganno digitale da parte di chi ti fidi → phishing
- Il peso delle parole nelle chat → violenza e manipolazione digitale
- Il peso delle parole nell'era digitale → cyberbullismo

Vigil Ai

Prototipazione di una Mobile Application:

- Email detector
- Social network Monitoring
- Chat Manager



The image shows a mobile application prototype for 'Vigil Ai'. At the top, there is a header with the 'Vigil Ai' logo (a stylized figure with a lightning bolt) and a share icon. Below the header is a circular profile picture of a woman, with the username '@vitabarletta' underneath it. To the right of the username is a chat icon. The main content area consists of several input fields with labels: 'Nome' (containing 'Vita Santa Barletta'), 'Email to check' (containing 'melpeters@gmail.com'), 'Facebook' (containing 'https://facebook.link.mio'), 'Instagram' (containing 'https://instagram.link.mio'), 'Linkedin' (containing 'https://linkedin.link.mio'), and 'X' (containing 'https://xVivaElon.link.mio'). At the bottom, there is a section labeled 'Other social network' with three icons: a bar chart, a plus sign in a circle, and a person icon.

Vigil Ai

@vitabarletta

Nome

Vita Santa Barletta

Email to check

melpeters@gmail.com

Facebook

https://facebook.link.mio

Instagram

https://instagram.link.mio

Linkedin

https://linkedin.link.mio

X

https://xVivaElon.link.mio

Other social network

Bar chart icon, Plus icon, Person icon

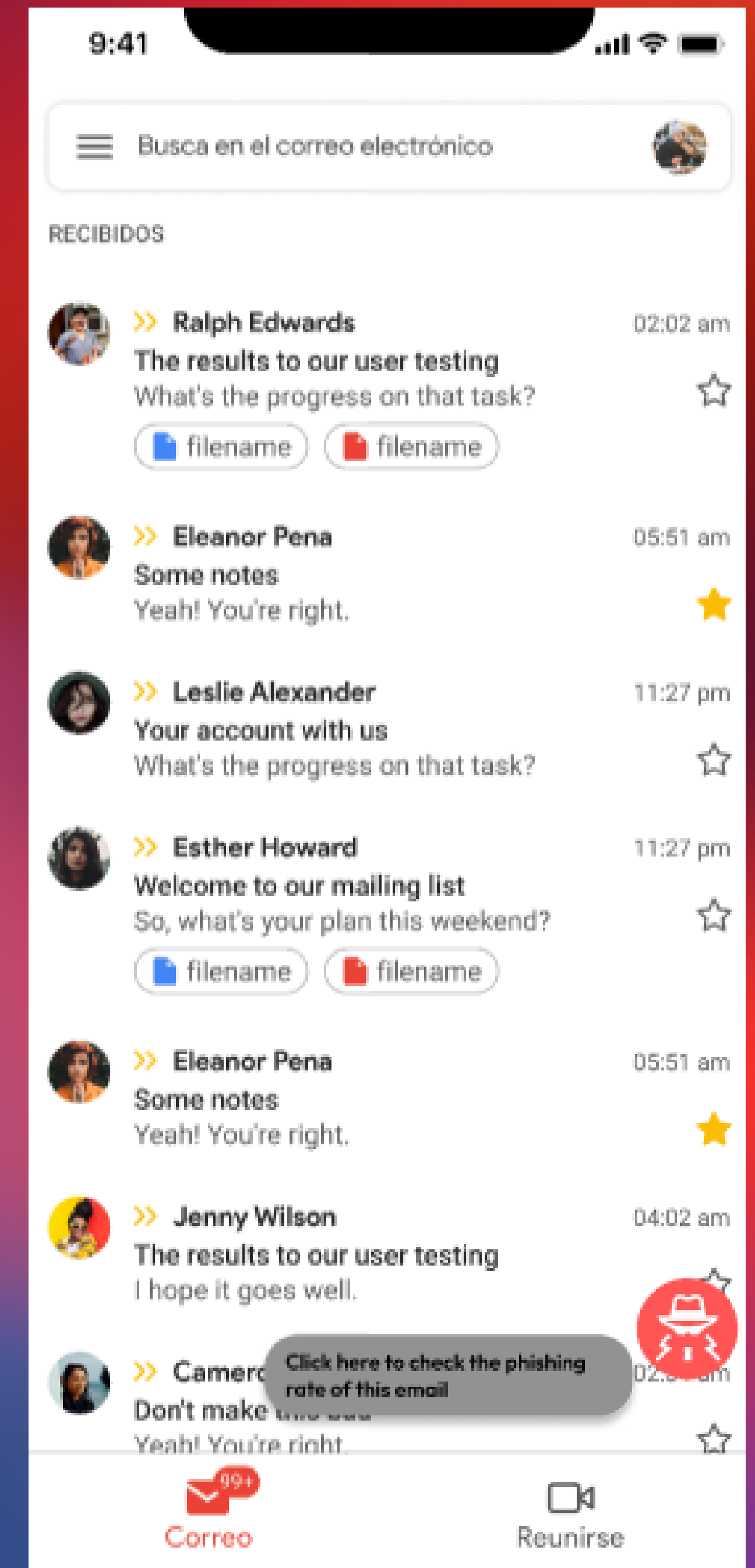
Vigil Ai - Phishing

Scansione multilivello:

- Analisi contenuto artificialmente generato
- Controllo del dominio e relativa affidabilità

Alerting:

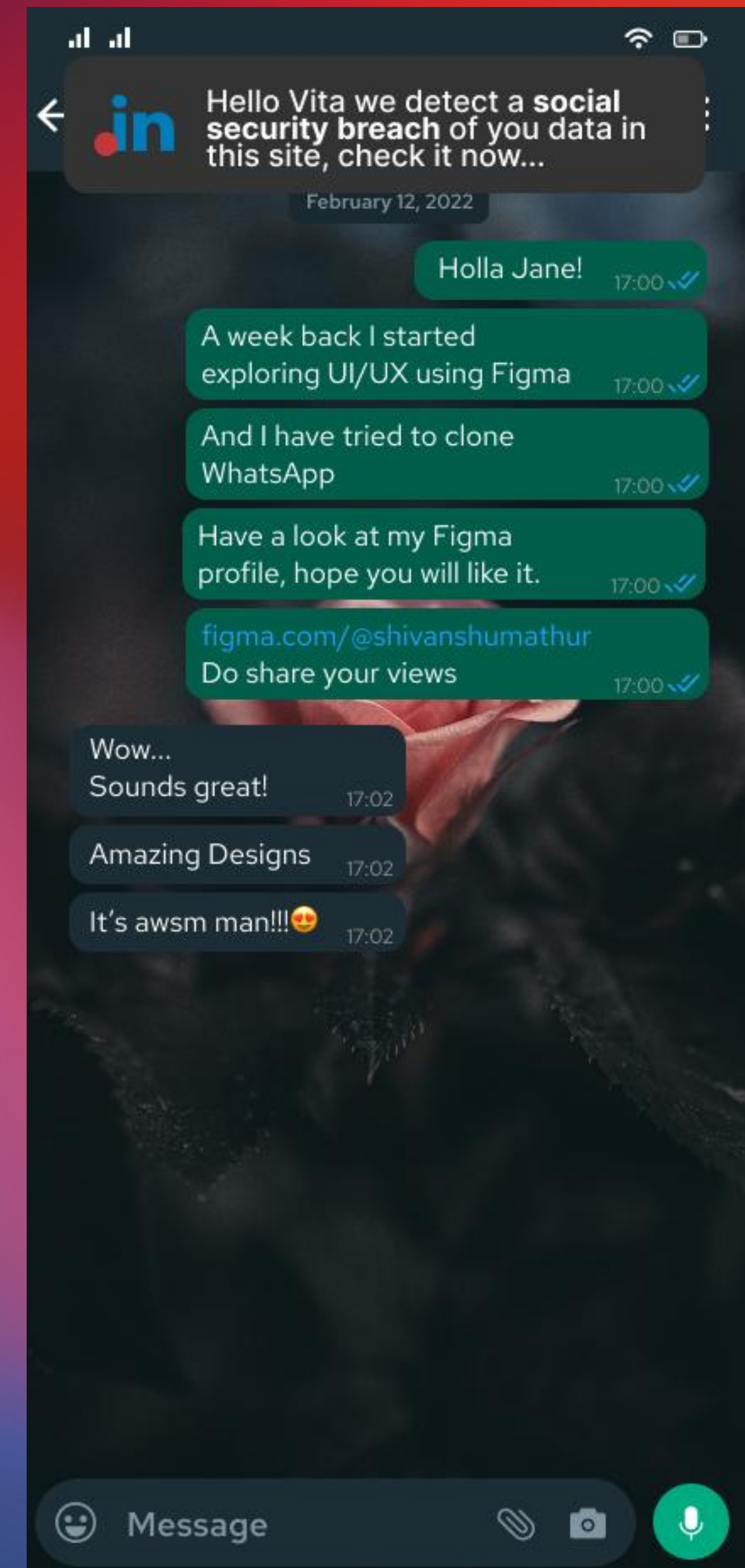
-  Sospetto lieve: AI generated
-  Sospetto medio: Domain sospetto
-  Phishing certo: Combinazione dei precedenti



Vigil Ai - Chat

Modello NLP che effettua:

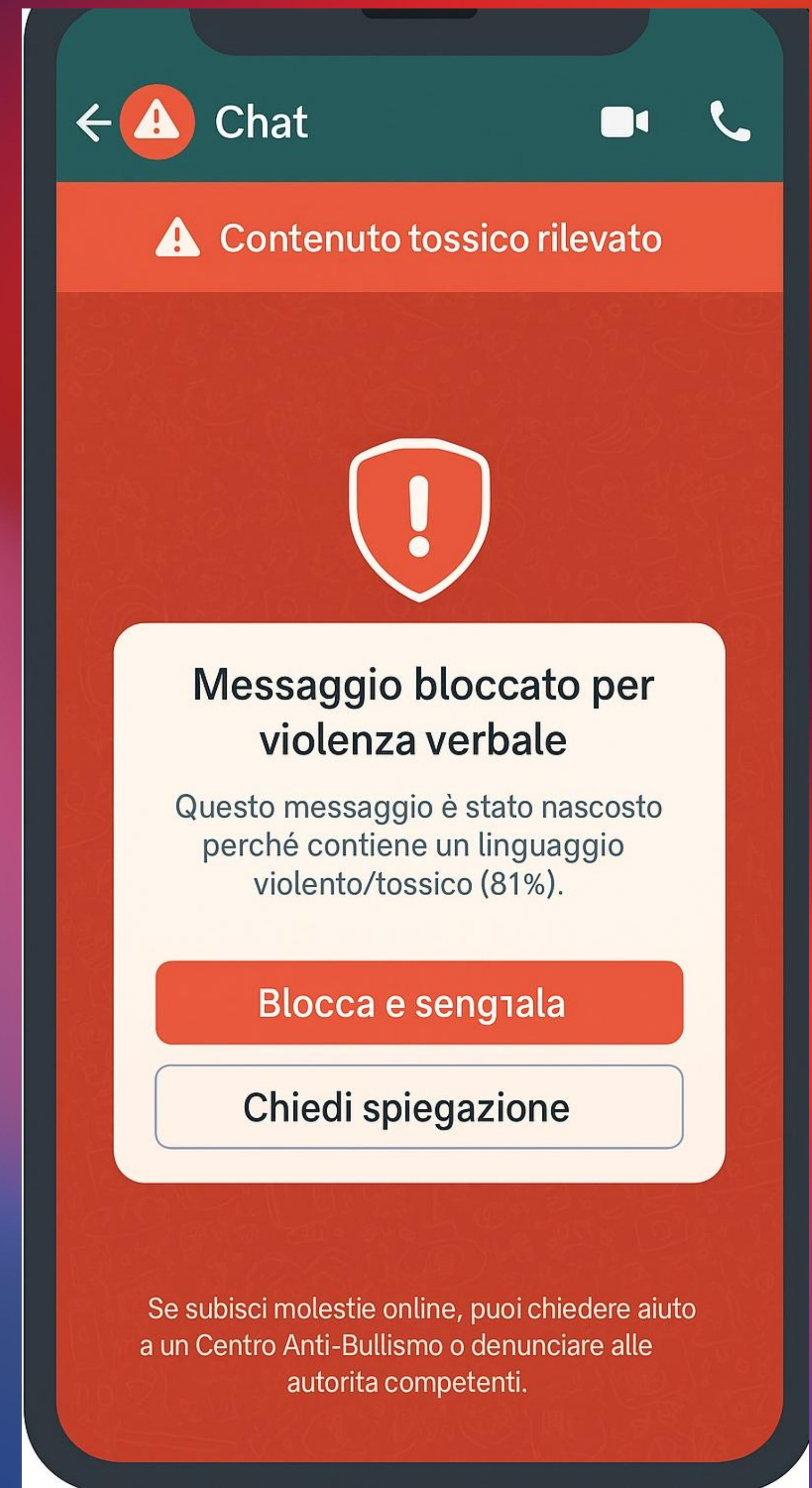
- **Analisi semantica dei messaggi**
- **Rilevamento di contest tossico e violenza verbale + LLaMAntino: spiegazione del rilevamento**
- **Alert bloccante se livello di tossicità >70%**
- **Log dei messaggi più tossici**
- **Sostegno psicologico con contatti sensibili**



Vigil Ai - Support

Proposte di supporto alla vittima:

- **Link su piattaforme polizia postale**
- **Contatti per Centri anti violenza**
- **Contatti per Centri bullismo**
- **Assistenza psicologica**

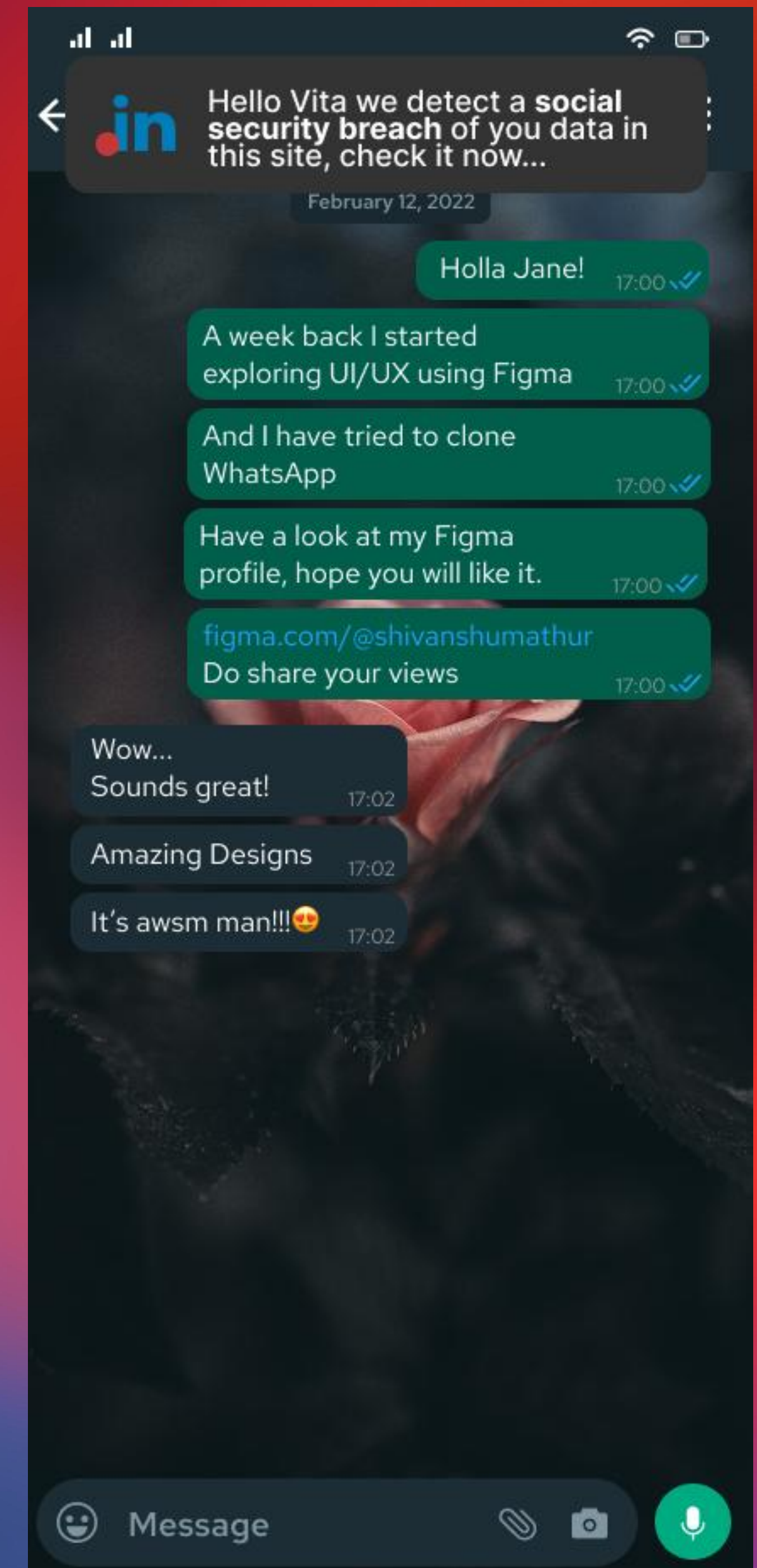


Vigil Ai – Social Media

- **Crawling Profili social**
- **Controllo post, commenti, menzioni**
- **Controllo immagini contenenti l'utente**

Alerting:

-  **Basso** → menzione con insulti isolati
-  **Medio** → ripetizione sistematica di citazioni negative
-  **Alto** → immagini sensibili + link diretto + possibile cyberbullismo/abuso



GRAZIE PER L'ATTENZIONE