



ĐẠI HỌC QUỐC GIA TP. HỒ  
CHÍ MINH  
TRƯỜNG ĐẠI HỌC CÔNG  
NGHỆ THÔNG TIN  
KHOA KHOA HỌC VÀ KỸ  
THUẬT THÔNG TIN  
—o0o—

# HỆ THỐNG GIỚI THIỆU KẾT BẠN DỰA TRÊN TÍNH CÁCH, XỬ LÝ TRỰC TIẾP TRÊN THIẾT BỊ VÀ BẢO VỆ QUYỀN RIÊNG TƯ

Sinh viên thực hiện: **Hoàng Minh Nhật - 24550031**

Giảng viên hướng dẫn: **ThS Ngô Khánh Khoa**

Cử nhân ngành: **Công nghệ thông tin**

On-device personality-based friend recommendation for privacy-preserving social networking

# 1. Mục lục

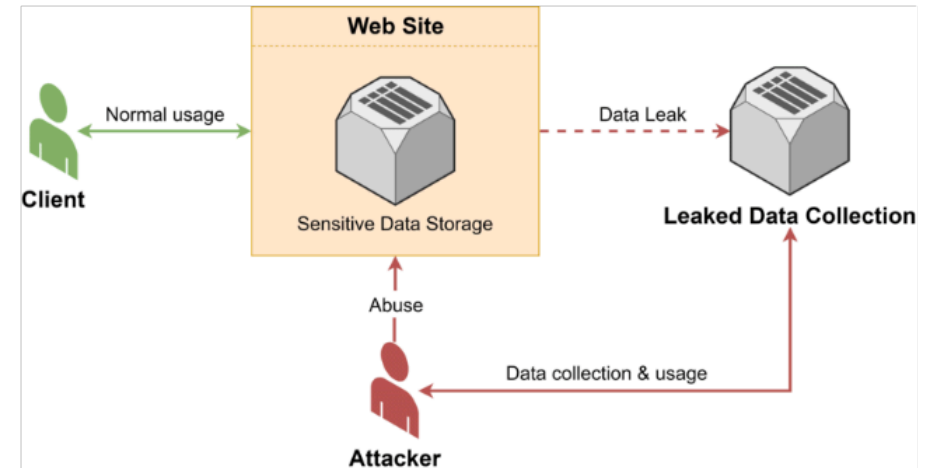
## 1. Mục lục

- Bối cảnh & động lực
- Mục tiêu & đóng góp
- Kiến trúc hệ thống tổng thể
- Dữ liệu tính cách & PCA on-device
- Bảo mật & lưu trữ dữ liệu nhạy cảm
- Mô hình gợi ý & xếp hạng
- Quy trình nghiệp vụ (End-to-End)
- Hiện thực & công nghệ
- Đánh giá & kết quả
- Demo

## 2. Bối cảnh & động lực

## 2. Bối cảnh & động lực

- Ứng dụng kết bạn cần gợi ý phù hợp nhưng phải tôn trọng quyền riêng tư.
- Dữ liệu tính cách (Big Five) và sở thích là dữ liệu nhạy cảm.
- Giải pháp: xử lý trực tiếp trên thiết bị + mã hóa dữ liệu trước khi lưu trữ.



## 3. Mục tiêu & đóng góp

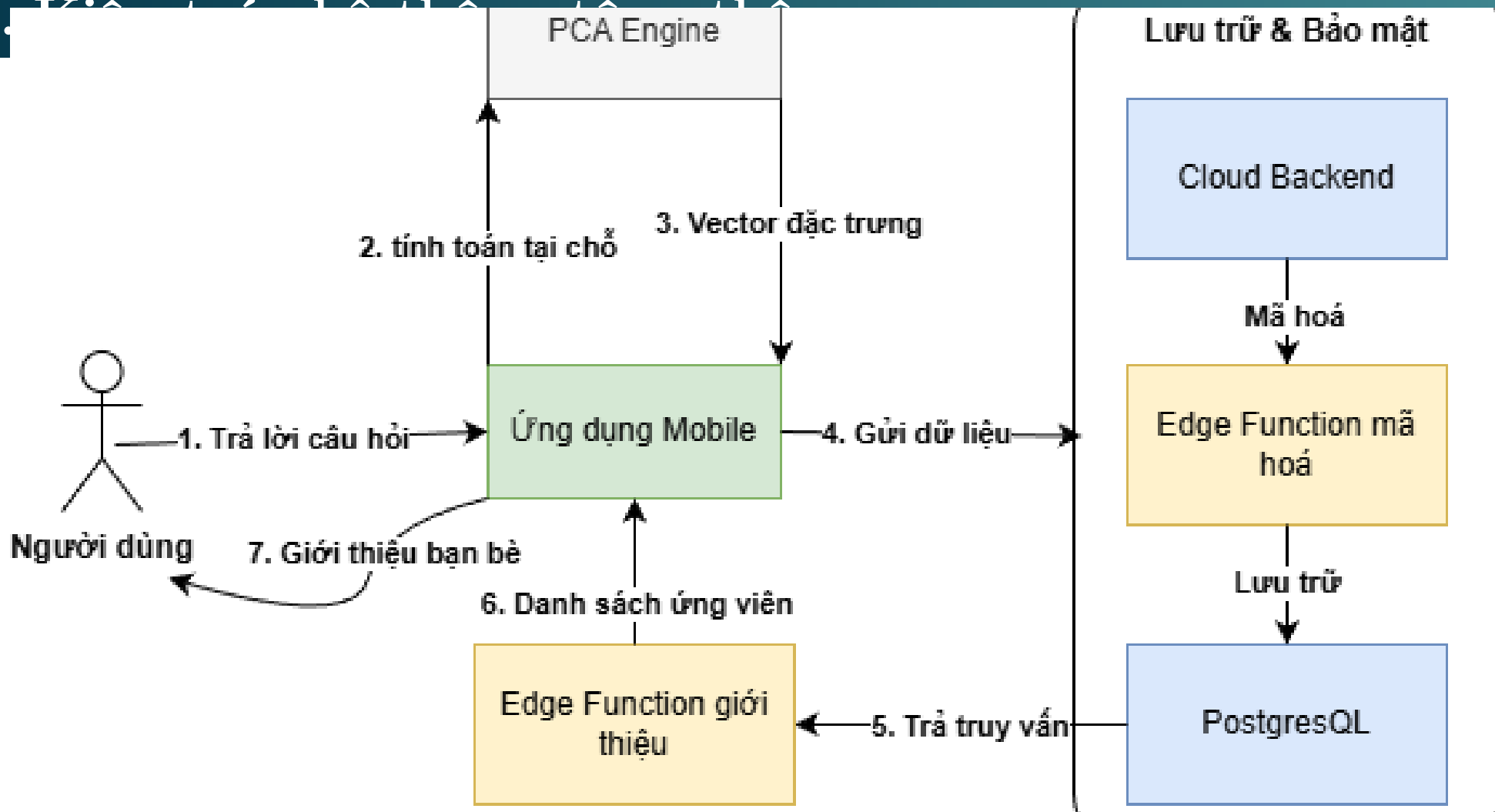
### 3. Mục tiêu & đóng góp

- Xây dựng hệ thống gợi ý kết bạn dựa trên tính cách + sở thích.
- Thiết kế pipeline xử lý tính cách trên thiết bị (PCA 5D  $\rightarrow$  4D).
- Mã hóa dữ liệu nhạy cảm bằng AES-256-GCM qua Edge Functions.
- Kết hợp điểm tương đồng PCA, ELO và sở thích để xếp hạng gợi ý.
- Triển khai ứng dụng di động + backend Supabase.

## 4. Kiến trúc hệ thống tổng thể

## 4. Kiến trúc hệ thống tổng thể

## 4. Trào lưu & phân tích hệ thống



## 4. Kiến trúc hệ thống tổng thể

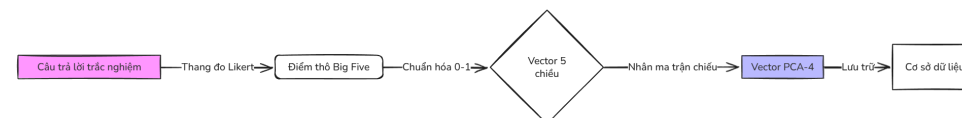
- Client: thu thập Big Five & sở thích, xử lý PCA on-device.
- Backend: lưu trữ mã hóa, tính điểm, gợi ý qua Edge Functions.



## 5. Dữ liệu tính cách & PCA on-device

### 5. Dữ liệu tính cách & PCA on-device

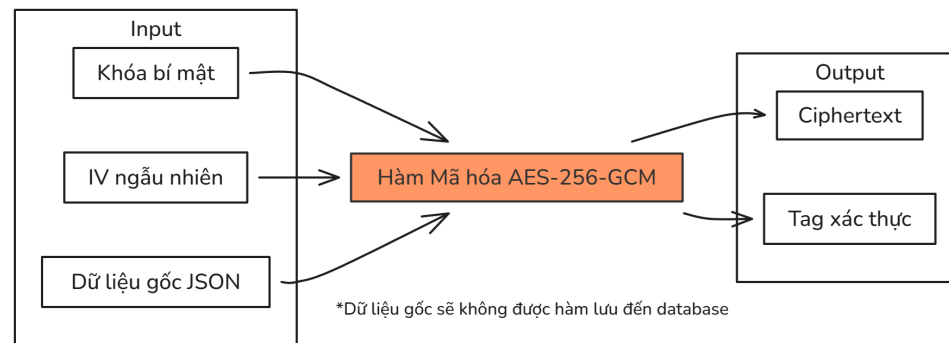
- Thang đo Big Five (IPIP-50).
- Chuẩn hóa theo mean thống kê.
- PCA giảm chiều 5  $\rightarrow$  4 để tăng tốc và ổn định.
- Tính toán ngay trên thiết bị, không cần ML server.



## 6. Bảo mật & lưu trữ dữ liệu nhạy cảm

### 6. Bảo mật & lưu trữ dữ liệu nhạy cảm

- Không lưu plaintext Big Five hoặc sở thích.
- Mã hóa AES-256-GCM qua Edge Function score-crypto.
- Lưu `b5_cipher`, `hobbies_cipher` + `iv`.
- Phân quyền và RLS trên Supabase.

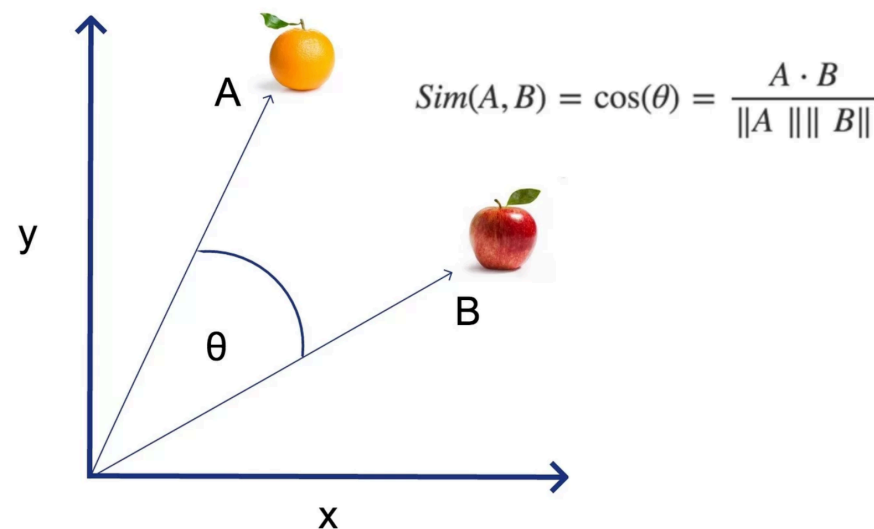


## 7. Mô hình gợi ý & xếp hạng

### 7. Mô hình gợi ý & xếp hạng

- Tương đồng tính cách: cosine trên PCA 4D.
- Độ gần ELO: ưu tiên mức tương đương.
- Sở thích: vector embedding + similarity.
- Điểm tổng: kết hợp có trọng số theo cấu hình.

#### Cosine Similarity



## 8. Quy trình nghiệp vụ End-to-End

## 8. Quy trình nghiệp vụ End-to-End

## 8. Demo 1: Tính cách

The screenshot displays a personality assessment application interface. On the left, a dark-themed card shows the statement "I am quiet around strangers." with five response options: a red 'X', an orange downward arrow, a grey circle, a cyan upward arrow, and a blue checkmark. On the right, a radar chart visualizes the results across five dimensions: Intellect/Imagination, Agreeableness, Conscientiousness, Emotional Stability, and a fifth unlabeled dimension. The chart shows high scores in Emotional Stability and Agreeableness. Below the chart, a box titled "Here's your final result!, nha!" lists the top strengths: Emotional Stability (70%), Agreeableness (50%), and Conscientiousness (50%).

I am quiet around strangers.

× ↓ ● ↑ ✓

Here's your final result!,  
nha!

Top strengths:

- Emotional Stability: 70%
- Agreeableness: 50%
- Conscientiousness: 50%

## 8. Quy trình nghiệp vụ End-to-End

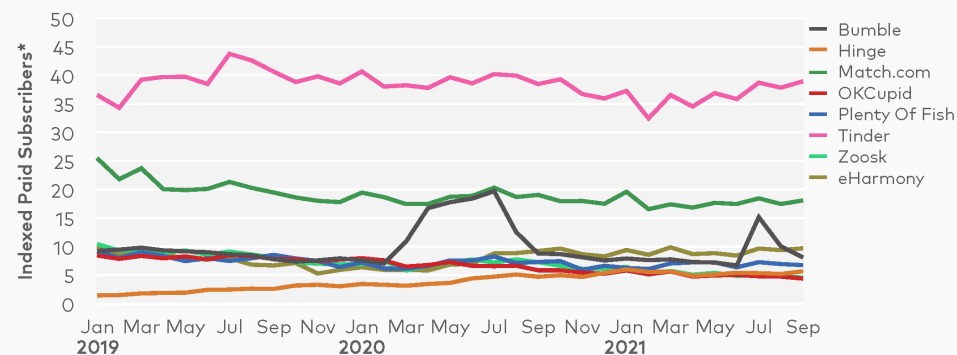
- Người dùng trả lời Big Five + chọn sở thích.
- PCA on-device → mã hóa → lưu DB.
- Edge Function gợi ý & trả về danh sách.

## 9. Hiện thực & công nghệ

### 9. Hiện thực & công nghệ

- Frontend: React Native (Expo), Zustand, i18n.
- Backend: Supabase (Auth, DB, Edge Functions).
- On-device ML: PCA weights nhúng trong app.
- Realtime: Supabase Realtime + RealtimeManager.

Monthly Indexed Paid Customers for Select Online Dating Services



\* Indexed to total paid subscribers in January 2019

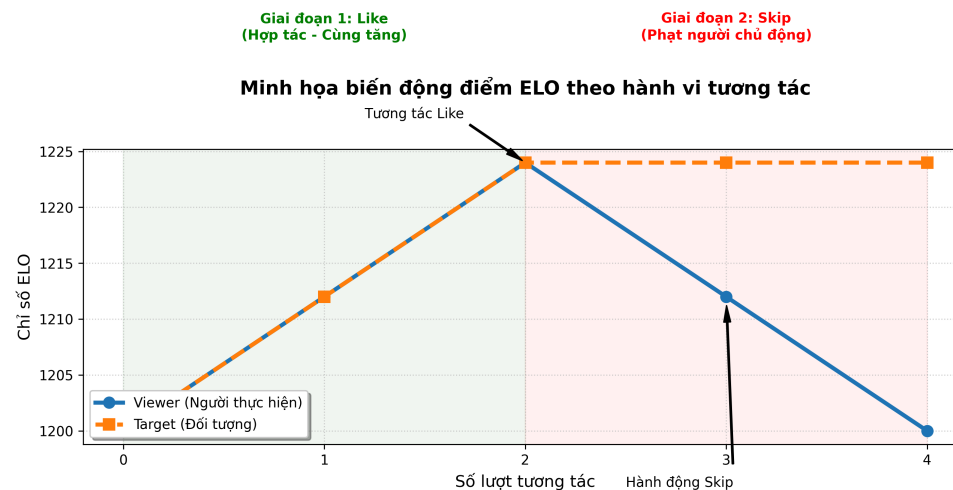
\* Data does not include free plans or Apple App Store transactions

 Bloomberg  
Second Measure

# 10. Đánh giá & kết quả

## 10. Đánh giá & kết quả

- Độ tương đồng PCA ổn định sau giảm chiều.
- Pipeline gợi ý chạy nhanh, đáp ứng realtime.
- Mã hóa đảm bảo dữ liệu nhạy cảm không lộ.





## 11. Demo (15 phút)

### 11. Demo (15 phút)

- Đăng nhập tài khoản demo.
- Trả lời một phần Big Five, cập nhật sở thích.
- Xem danh sách gợi ý và thao tác like/match.
- Kiểm tra hồ sơ chi tiết và lịch sử match.

## 12. Kết thúc

## 12. Kết thúc

Xin cảm ơn hội đồng và thầy/cô đã lắng nghe

Q&A