

0. Đối tượng sử dụng

Cả khách hàng và nhân viên, phụ thuộc vào đang áp dụng cho luồng nghiệp vụ nào

1. Hệ thống cần bắt buộc hỗ trợ đầy đủ những tính năng nào - nhập, chỉnh sửa, xóa, text-to-voice, điều hướng hoặc tìm kiếm bằng lệnh thoại? BTC có thể cung cấp danh sách ưu tiên (must-have, should-have, nice-to-have) không?

Tất cả tính năng Thêm, sửa, xóa, tìm kiếm, chỉnh sửa, điều hướng tìm kiếm trên Form nhập liệu. Must-have.

2. BTC có cung cấp dữ liệu mẫu hoặc giao diện form mẫu, giao diện mẫu, hay API có sẵn không?

Để ứng viên chủ động xây form, khuyến nghị Form phức tạp với nhiều ô nhập liệu khác nhau, nhiều tính năng khác nhau.

3. Kỳ vọng độ trễ phản hồi người dùng ở mức nào, chấp nhận lâu hơn để đạt độ chính xác cao? BTC ưu tiên tốc độ hay độ chính xác hơn?

0.1 giây sau câu nói hệ thống cần hiểu hành động và thực thi. Ưu tiên cả 2.

4. Khi người dùng nói cho hệ thống, hệ thống nên thực thi ngay hay chờ đến khi người dùng dừng nói?

Hệ thống cần đảm bảo hiểu được khi nào người dùng kết thúc câu lệnh.

5. Có quy định về việc sử dụng nền tảng không (AWS, Azure, GCP, on-premise)? Có khuyến khích dùng LLM open-source hoặc tự train/fine-tune không?

AWS, LLM open-source

6. BTC có thể chia sẻ các tiêu chí và trọng số chấm điểm không?

Giao diện: Đơn giản, thân thiện với người dùng và dễ tương tác.

Tính năng: Hỗ trợ các chức năng thiết yếu như tương tác bằng giọng nói với AI, nhận dạng giọng nói chính xác, thực thi đúng các lệnh nhập liệu, kích hoạt các chức năng trên giao diện, hiểu được giọng vùng miền và tự động sửa các lỗi chính tả phổ biến.

7. BTC có chuẩn bị sẵn bộ test cases (các câu lệnh mẫu, tình huống đa ý định, mơ hồ hoặc chỉnh sửa lệnh) dùng để chấm không?

Để khách quan, BTC sẽ hỏi theo testcase của BTC khi ứng viên demo.

8. BTC mong đợi hệ thống xử lý tốt các tình huống đặc biệt nào, như tiếng ồn lớn, nhiều người nói cùng lúc, audio kém chất lượng, hoặc lệnh phức tạp/mơ hồ/đối nghịch?

Hệ thống cần lọc được tiếng ồn, bắt được chính xác voice của người nói, cả trong tình huống người nói nhỏ, tiếng địa phương, từ nóng,

Về Nghiệp vụ:

1. BTC có thể làm rõ các loại hình giao dịch mà hệ thống cần hỗ trợ - như mở tài khoản, đăng ký vay, chuyển tiền, mở thẻ, xử lý khiếu nại, KYC/compliance, v.v.? Trong phạm vi hackathon, nên tập trung vào 1-2 workflow chính hay cần xây dựng hệ thống linh hoạt cho nhiều loại nghiệp vụ? Workflow nào được ưu tiên?

Ứng viên có thể demo Luồng vay, mở tài khoản hoặc mở thẻ

2. Một quy trình nghiệp vụ điển hình có bao nhiêu bước và bao nhiêu trường dữ liệu cần nhập ở mỗi bước? Các bước có phụ thuộc tuần tự hay có thể xử lý song song? Có cho phép quay lại bước trước hoặc có logic rẽ nhánh (branching logic) tùy điều kiện khách hàng không?

Bank có rất nhiều luồng nghiệp vụ sản phẩm khác nhau. Có thể tham khảo luồng mở tài khoản hoặc luồng mở thẻ tín dụng của bank.

Bác bước có thể tuần tự hoặc không tùy vào Luồng.

Có thể quay lại bước trước để update thông tin.

3. Mỗi form trung bình có bao nhiêu trường thông tin, và kiểu dữ liệu phổ biến gồm những gì (text, số, ngày, dropdown, checkbox, file upload...)? Có các quy tắc validation phức tạp nào

như định dạng, giới hạn giá trị, cross-field logic hay business rules cần tuân thủ không? BTC có quy chuẩn cụ thể về độ dài ký tự, format ngày, hoặc cách hiển thị tiền tệ không?

BTC khuyến nghị form càng nhiều thông tin, càng nhiều ô nhập liệu với validation cho từng ô tuân thủ theo format tối thiểu, ví dụ ô nhập số, ô ngày tháng, ô thông tin tiền tệ.

4. BTC có thể chia sẻ tỷ lệ lỗi nhập liệu hiện tại (baseline) trong các quy trình thực tế và các loại lỗi phổ biến nhất (chính tả, sai định dạng, thiếu trường bắt buộc...)? Mục tiêu mà BTC mong muốn hệ thống đạt được là giảm tỷ lệ lỗi xuống mức bao nhiêu phần trăm, hoặc giới hạn từng loại lỗi cụ thể ra sao?

Dữ liệu cần phải đúng và đủ nên độ chính xác cần 99%, 1% người dùng sẽ verify trước khi submit.

5. Một form hoặc workflow hiện tại trung bình mất bao lâu để hoàn tất? Thời gian đó phân bổ thế nào giữa thao tác nhập, đọc hiểu, di chuyển giữa các ô, sửa lỗi? BTC kỳ vọng giảm bao nhiêu phần trăm thời gian nhập liệu hoặc tổng thời gian xử lý end-to-end?

Chưa có thời gian đo lường cụ thể. Khi hệ thống hiểu được câu lệnh, fill đúng đủ thông tin là được.

6. Hệ thống có cần tối ưu riêng cho môi trường nhiều tiếng ồn không?

Có, vì nhiều phòng ban nhân viên ngồi chung nên sẽ ồn.

7. BTC có thể chia sẻ vùng miền chính của người dùng để xác định giọng nói cần hỗ trợ (Bắc, Trung, Nam, giọng Huế)? Hệ thống có cần hiểu song ngữ Việt-Anh dùng đan xen và các biến thể ngôn ngữ thường dùng trong ngành (ví dụ: “save/lưu”, “delete/xóa”) không?

Nhân viên của VPB khắp các miền nên cần hỗ trợ đủ (Bắc, Trung, Nam, Huế...). Có vì nhiều form là tiếng Anh.

8. Mức độ chủ động của hệ thống khi tự sửa lỗi chính tả là gì - tự sửa, gợi ý, hay hỏi xác nhận?

Prefer việc tự sửa, trong tình huống không thì có thể gợi ý / hoặc hỏi xác nhận

9. Hành động nào là bắt buộc trong nhập - sửa - xóa? Hệ thống có cần hỗ trợ thêm các thao tác như điều hướng giữa field/section, đọc lại dữ liệu, tìm kiếm, xác minh, lưu/hủy/reset form hay lưu nháp không?

Đầy đủ : thêm, sửa, xóa tìm kiếm, điều hướng, xác minh ...

10. Các loại nút nào cần điều khiển bằng giọng nói (Save, Submit, Cancel, Next, Reset, Print...)? Hệ thống có cần phân biệt nút khả dụng hay bị disable không, và có tự động xử lý

hộp thoại xác nhận (popup “Bạn có chắc chắn muốn xóa?” → “Có/Xác nhận”) không? Ngoài ra, có cần hỗ trợ các phần tử phức tạp như dropdown, date picker, file upload không?

Tất cả các nút các tính năng trên FORM

11. Hệ thống có cho phép người dùng kết hợp giọng nói với bàn phím, chuột (ví dụ: click chọn field rồi nói để nhập) không? Khi người dùng nói sai và muốn sửa bằng tay, có được phép không?

Có, cho linh hoạt

12. Hệ thống có cần hiểu ngữ cảnh trong cùng phiên làm việc (ví dụ: sau “Nhập tên Nguyễn Văn A” thì hiểu “Tuổi là 30” là bước kế tiếp) và nhận biết các lệnh sửa (“Không, là 0987654321”)? Ngoài ra, có cần hỗ trợ hiểu đại từ (“anh ấy”, “nó”)?

Có