Trợ lý Lexia

Tài liệu bảo mật cho bộ phận IT

Dữ liệu được lưu ở đâu, gửi đi những gì, ai được xem, và doanh nghiệp kiểm soát thế nào. Tài liệu mô tả phiên bản dịch vụ hiện hành; cam kết ràng buộc nằm trong hợp đồng và Phụ lục xử lý dữ liệu ký với từng doanh nghiệp.

Phiên bản 11/2026 · Công ty Cổ phần Công nghệ Lexia

1. Tóm tắt

- Dịch vụ phần mềm (SaaS) chạy trên trình duyệt; doanh nghiệp không cài đặt gì, chỉ cần kết nối HTTPS (cổng 443).
- Máy chủ ứng dụng, cơ sở dữ liệu, kho tệp, bản sao lưu và nhật ký đặt tại Việt Nam.
- Văn bản được đọc chữ (OCR) tại Việt Nam trước mọi bước khác. Tệp gốc không rời Việt Nam.
- Chỉ phần chữ đã che thông tin nhạy cảm được gửi cho mô hình AI ở nước ngoài; văn bản được đánh dấu Mật không bao giờ được gửi.
- Dữ liệu các doanh nghiệp tách biệt ở tầng cơ sở dữ liệu; phân quyền theo vai trò và phòng ban; xác thực hai lớp.
- Nhân viên của bên cung cấp không xem nội dung của doanh nghiệp, trừ khi Quản trị viên của doanh nghiệp mở lượt hỗ trợ có thời hạn.

2. Kiến trúc và nơi đặt dữ liệu

Thành phần	Vai trò	Nơi đặt
Máy chủ web và ứng dụng	Giao diện, API, kiểm tra quyền	Việt Nam
Cơ sở dữ liệu PostgreSQL	Dữ liệu nghiệp vụ, nhật ký	Việt Nam
Kho tệp	Tệp gốc (không ghi đè), ảnh trang, tệp xuất	Việt Nam
Đọc chữ (OCR), che dữ liệu	Đọc bản scan, ảnh chụp; phát hiện và che thông tin nhạy cảm	Việt Nam
Quét mã độc	Kiểm tra mọi tệp tải lên	Việt Nam
Bản sao lưu	Sao lưu hằng ngày, giữ 30 ngày	Việt Nam
Mô hình ngôn ngữ (AI)	Phân loại, bóc thông tin, tóm tắt, Hỏi AI trên phần chữ đã che	Anthropic, Hoa Kỳ

Luồng xử lý một văn bản: tải lên qua HTTPS → quét mã độc → lưu tệp gốc → đọc chữ tại Việt Nam → phát hiện và che thông tin nhạy cảm → gửi phần chữ đã che cho AI → kiểm chứng kết quả bằng code → người dùng kiểm tra và xác nhận.

3. Dữ liệu gửi cho AI

- **Được che trước khi gửi**: số CCCD, CMND, hộ chiếu, mã số thuế cá nhân; số tài khoản, số thẻ; địa chỉ; số điện thoại, email; tên người (kể cả khi bản scan mất dấu); số tiền; thông tin tài sản; số bản án, số hợp đồng trong thân văn bản; các từ khóa do doanh nghiệp khai báo. Bảng đối chiếu nằm trên máy chủ tại Việt Nam; kết quả trả về được ghép lại tại Việt Nam.
- **Không bao giờ gửi**: tệp gốc, ảnh, chữ ký, con dấu; văn bản được đánh dấu Mật; dữ liệu tài khoản người dùng; nhật ký; mật khẩu, khóa bí mật.
- **Tùy chọn không gửi gì ra nước ngoài**: theo Phụ lục xử lý dữ liệu, doanh nghiệp có thể chọn chỉ xử lý tại Việt Nam; khi đó văn bản vẫn được đọc chữ, lưu, tìm kiếm, và người dùng nhập thông tin.
- Nhà cung cấp AI cam kết không dùng dữ liệu gửi qua API để huấn luyện mô hình. Mỗi lần gọi AI được ghi nhật ký (người, thời điểm, văn bản, nơi xử lý, số token), không lưu nội dung.
- Kết quả AI là đề xuất: chỉ trở thành dữ liệu chính thức khi người dùng xác nhận; thời hạn phải được xác nhận riêng. Câu trả lời của Hỏi AI phải có trích dẫn khớp nguyên văn tài liệu, nếu không hệ thống trả lời là không tìm thấy.

4. Tách biệt dữ liệu và phân quyền

- Mỗi bảng dữ liệu nghiệp vụ gắn mã tổ chức và được tách bằng Row-Level Security của PostgreSQL, cộng với kiểm tra ở tầng ứng dụng. Ứng dụng kết nối cơ sở dữ liệu bằng tài khoản không có quyền chủ sở hữu. Bộ kiểm thử tự động về cách ly giữa các tổ chức là điều kiện để phát hành.
- Bốn vai trò: Quản trị viên, Trưởng bộ phận (phạm vi phòng ban), Chuyên viên (hồ sơ được giao), Chỉ xem. Danh sách, số đếm, tìm kiếm, Hỏi AI và file Excel đều lọc theo quyền ngay trong truy vấn.
- Văn bản Mật: chỉ Lãnh đạo, người được Quản trị viên cho phép xem văn bản mật và người tham gia xử lý xem được; Quản trị viên không tự động xem được.

5. Đăng nhập, phiên làm việc, mã hóa

- Mật khẩu lưu dạng mã băm argon2id; khóa tài khoản 15 phút sau 5 lần sai; giới hạn số lần thử theo email và địa chỉ IP; mật khẩu tạm phải đổi ở lần đăng nhập đầu.
- Xác thực hai lớp (TOTP, ứng dụng Google Authenticator hoặc tương đương), doanh nghiệp bắt buộc được cho mọi tài khoản; 10 mã khôi phục dùng một lần.
- Phiên lưu trên máy chủ; cookie HttpOnly, Secure, SameSite; chống giả mạo yêu cầu (CSRF); hết hạn sau 30 phút không dùng hoặc 12 giờ.
- Đường truyền: HTTPS, chỉ TLS 1.2 và 1.3, HSTS. Lưu trữ: dữ liệu trên hạ tầng tại Việt Nam có mã hóa ổ lưu trữ; khóa xác thực hai lớp được mã hóa thêm ở tầng ứng dụng.

6. Tệp tải lên và tải xuống

- Nhận dạng định dạng theo nội dung tệp (PDF, JPG, PNG, TIFF, WEBP; Word, Excel cho mẫu); từ chối PDF có mật khẩu, Excel có macro; quét mã độc mọi tệp.
- Tệp gốc lưu bất biến: không ghi đè; sửa văn bản tạo phiên bản mới. Tải xuống qua liên kết có thời hạn ngắn, gắn với người đăng nhập.
- Giá trị ghi ra Excel luôn ở dạng chữ hoặc ngày, không bao giờ là công thức.

7. Nhật ký, giám sát, sao lưu

- Nhật ký ghi: đăng nhập, xem, tải, sửa, xóa, xuất Excel, gửi AI, trình, duyệt, trả lại, từ chối. Nhật ký chỉ được thêm, không sửa, không xóa; không chứa nội dung văn bản.
- Giám sát máy chủ 5 phút một lần và từ bên ngoài; cảnh báo cho người trực khi dịch vụ lỗi, ổ đĩa đầy, chứng chỉ sắp hết hạn.
- Sao lưu cơ sở dữ liệu và kho tệp hằng ngày, kiểm tra tính toàn vẹn, giữ 30 ngày, không giữ bản sao lưu dài hạn; diễn tập khôi phục định kỳ.

8. Quyền truy cập của bên cung cấp

- Nhân viên vận hành dùng bảng điều khiển riêng: đăng nhập bằng tài khoản riêng có xác thực hai lớp, chỉ truy cập từ mạng nội bộ của bên cung cấp, chỉ thấy số liệu vận hành (số văn bản, dung lượng, lượt AI), không đọc được văn bản, hồ sơ, công việc.
- Khi cần hỗ trợ, Quản trị viên của doanh nghiệp mở một lượt hỗ trợ tối đa 72 giờ, thu hồi được bất cứ lúc nào. Trong lượt đó người hỗ trợ chỉ xem, không sửa được dữ liệu, không xem được văn bản Mật; mọi thao tác ghi vào nhật ký của doanh nghiệp.
- Xóa dữ liệu của một tổ chức cần hai nhân viên vận hành khác nhau duyệt, mỗi lần nhập mã xác thực hai lớp mới.

9. Vòng đời dữ liệu khi ngừng dùng

Mốc	Việc
Ngày chấm dứt	Chuyển chế độ chỉ xem; gói xuất dữ liệu (tệp gốc, sổ văn bản Excel, dữ liệu có cấu trúc, nhật ký) sẵn sàng tải về
Sau 30 ngày	Xóa dữ liệu chính, hai người duyệt; gửi biên bản xác nhận đã xóa. Gia hạn thêm 30 ngày một lần khi doanh nghiệp đề nghị
Sau 60 ngày	Dữ liệu đã hết khỏi mọi bản sao lưu

Nhật ký bảo mật giữ 24 tháng sau khi xóa; hợp đồng, chứng từ và biên bản xóa giữ 10 năm, không chứa nội dung văn bản.

10. Phát triển và vận hành an toàn

- Mỗi thay đổi mã nguồn chạy bộ kiểm thử tự động (cách ly tổ chức, phân quyền, tính thời hạn, che dữ liệu, không gửi dữ liệu ra ngoài khi không được phép).
- Quét lỗ hổng thư viện và ảnh container ở mỗi lần build; cập nhật bản vá hệ điều hành tự động.
- Mọi lời gọi AI đi qua một cổng duy nhất (AI Gateway) áp dụng quy tắc che dữ liệu và ghi nhật ký; kết quả AI được kiểm tra theo lược đồ trước khi ghi.

11. Doanh nghiệp cần chuẩn bị

- Trình duyệt Chrome, Edge, Firefox hoặc Safari bản mới; cho phép truy cập **lexiaassist.com** và tên miền ứng dụng qua HTTPS.
- Chỉ định Quản trị viên; bật bắt buộc xác thực hai lớp; khóa tài khoản người nghỉ việc; rà phân quyền định kỳ.
- Đánh dấu Mật cho văn bản không được rời Việt Nam. Không tải lên tài liệu mang dấu bí mật nhà nước.

12. Chưa có ở phiên bản hiện tại

Đăng nhập một lần (SSO với Active Directory, Azure AD), triển khai trên máy chủ riêng hoặc tại chỗ, chữ ký số, nhận văn bản qua email, tích hợp hệ thống quản lý văn bản khác. Các mục này nằm trong lộ trình; doanh nghiệp có nhu cầu xin liên hệ để trao đổi.

Liên hệ: Công ty Cổ phần Công nghệ Lexia · 6/143 Đông Các, phường Đống Đa, Hà Nội · lienhe@lexiaassist.com · lexiaassist.com
Bộ phận IT cần thêm thông tin (bảng câu hỏi bảo mật, mẫu Phụ lục xử lý dữ liệu) xin gửi email; chúng tôi trả lời bằng văn bản.