Phân tích · 16 tháng 5, 2026

← Trước·Tiếp theo →

Mã hóa không có nghĩa là quyền riêng tư: siêu dữ liệu nói gì về bạn

Nội dung được mã hóa và siêu dữ liệu hiển thị là hai việc khác nhau. Khi một dịch vụ nói về "mã hóa đầu-cuối", họ chỉ đang kể một nửa câu chuyện.

Ổ khóa không bảo vệ được tất cả mọi thứ

Phần lớn các dịch vụ nhắn tin hiện nay đều quảng cáo mã hóa đầu-cuối. Và đó là sự thật: nội dung tin nhắn di chuyển dưới dạng mã hóa, sao cho không ai trên đường đi – kể cả nhà cung cấp dịch vụ – có thể đọc được văn bản khi nó đang được truyền đi. Cho đến nay, khẳng định này là chính xác.

Vấn đề là nội dung chỉ là một phần của câu chuyện. Mặc dù không ai có thể đọc được những gì bạn nói, nhưng dịch vụ biết những điều khác với độ chính xác rất cao: bạn nói chuyện với ai, vào lúc nào, tần suất ra sao, từ vị trí xấp xỉ nào, trên thiết bị nào, bạn gửi bao nhiêu tin nhắn và nhận được bao nhiêu, bạn chia sẻ bao nhiêu tệp. Tất cả những thứ này được gọi là siêu dữ liệu (metadata). Và siêu dữ liệu, trong nhiều trường hợp, nói lên gần như nhiều điều bằng chính tin nhắn đó.

Siêu dữ liệu tiết lộ điều gì

Người ta không cần đọc tin nhắn để biết nhiều điều. Nếu một người gọi điện hoặc nhắn tin cho một bác sĩ ung thư vào mỗi sáng thứ Ba lúc chín giờ trong suốt sáu tháng, không cần nghe cuộc trò chuyện cũng có thể đoán được chuyện gì đang xảy ra. Nếu hai người trao đổi hàng trăm tin nhắn mỗi ngày và đột nhiên dừng lại, bạn không cần đọc bất kỳ tin nhắn nào để hiểu chuyện gì đã xảy ra. Nếu một chuyên gia tư vấn thuế nhận được hai mươi tin nhắn liên tiếp từ cùng một khách hàng vào đêm trước ngày kết thúc quý, quy luật đó đã tự nói lên tất cả.

Siêu dữ liệu tiết lộ các quy luật hành vi: ai có mối quan hệ với ai, thời gian biểu của mỗi người là gì, khi nào họ thức, khi nào họ ngủ, khi nào họ đi du lịch, khách hàng nào tích cực nhất, mối quan hệ chuyên môn nào mật thiết nhất. Một máy chủ thu thập siêu dữ liệu có thể xây dựng hồ sơ chi tiết về đời sống cá nhân và nghề nghiệp của bất kỳ người dùng nào mà không cần đọc một từ nào trong những gì họ viết.

Có một ví dụ lịch sử minh họa điều này một cách nghiệt ngã. Cựu giám đốc NSA, Michael Hayden, đã phát biểu thẳng thắn vào năm 2014: *"We kill people based on metadata"*. Khẳng định này đề cập đến các hoạt động quân sự của Hoa Kỳ chống lại các mục tiêu được xác định duy nhất dựa trên quy luật liên lạc của họ. Không một tin nhắn nào được đọc. Chỉ có biểu đồ liên lạc và thời gian biểu.

Việc một dịch vụ thu thập siêu dữ liệu không nhất thiết có nghĩa là họ sẽ sử dụng chúng để chống lại người dùng. Điều đó có nghĩa là họ có khả năng làm như vậy, và một bên thứ ba có quyền truy cập vào dữ liệu đó – thông qua lệnh của tòa án, thông qua một lỗ hổng bảo mật hoặc thông qua việc bán cho bên thứ ba nếu các điều khoản dịch vụ cho phép – cũng có khả năng đó.

Quyền truy cập vào danh bạ

Một vectơ khác hầu như không được chú ý: danh sách liên lạc. Phần lớn các dịch vụ nhắn tin yêu cầu quyền truy cập vào danh bạ điện thoại khi đăng ký. Họ tải tất cả các số lên máy chủ của mình để hiển thị những ai khác đang sử dụng dịch vụ. Kể từ thời điểm đó, công ty có một bản đồ đầy đủ về các mối quan hệ của người dùng, ngay cả khi người đó chưa từng viết một tin nhắn nào cho bất kỳ ai.

Đối với một chuyên gia có nghĩa vụ giữ bí mật nghề nghiệp – luật sư, bác sĩ, nhà tâm lý học, tư vấn viên – danh bạ đó chứa đựng các khách hàng. Nếu danh bạ đã được tải lên máy chủ của bên thứ ba, tên của các khách hàng nằm trong một cơ sở hạ tầng mà quyền tài phán và các chính sách của nó chuyên gia không kiểm soát được. Bí mật nghề nghiệp không bị phá vỡ vào ngày ai đó làm rò rỉ một cuộc trò chuyện: nó đã bị phá vỡ từ trước đó rất lâu, ngay vào lúc đồng ý tải lên.

Sự khác biệt giữa mã hóa và không thu thập

Mã hóa là bảo vệ nội dung. Riêng tư là không thu thập những gì không cần thiết. Đây là những thứ khác nhau, và sự khác biệt này mang tính quyết định về mặt vận hành. Một dịch vụ có thể mã hóa hoàn hảo tất cả các tin nhắn và đồng thời biết hầu hết mọi thứ về người dùng thông qua siêu dữ liệu. Cả hai điều này hoàn toàn tương thích. Trên thực tế, đó là mô hình kinh doanh thống trị trong ngành.

Câu hỏi đúng để đánh giá quyền riêng tư thực sự của một dịch vụ không phải là *"nó có mã hóa nội dung không?"*. Câu hỏi đó đã có lời giải từ nhiều năm nay. Câu hỏi đúng là: *"nó tạo ra những siêu dữ liệu nào và chúng được lưu trữ ở đâu?"*. Và trên hết: *"nó không cần tạo ra những siêu dữ liệu nào?"*.

Một kiến trúc giảm thiểu siêu dữ liệu ngay từ khâu thiết kế (privacy by design) – không phải bằng lời hứa, không phải bằng chính sách nội bộ – về mặt cấu trúc sẽ riêng tư hơn một kiến trúc thu thập và mã hóa chúng. Bởi vì dữ liệu không tồn tại thì không thể bị rò rỉ, không thể bị bán, không thể bị bàn giao theo lệnh của tòa án hay bị mất trong một vụ vi phạm bảo mật.

Dành cho độc giả chuyên nghiệp

Nếu hoạt động nghề nghiệp của bạn liên quan đến bí mật, sự bảo mật hoặc đơn giản là sự tôn trọng thông tin của bên thứ ba, bạn nên đặt câu hỏi theo thứ tự này:

1. Ứng dụng tôi sử dụng để liên lạc có mã hóa nội dung không? (Có lẽ là có.)
2. Nó có mã hóa siêu dữ liệu không? (Có lẽ là không.)
3. Nó có tạo ra siêu dữ liệu mà nó *không cần thiết* để vận hành không? (Gần như chắc chắn là có.)
4. Những siêu dữ liệu đó được lưu trữ ở đâu và thuộc quyền tài phán nào? (Có khả năng là bên ngoài Khu vực Kinh tế Châu Âu.)
5. Khách hàng hoặc bệnh nhân của tôi có biết rằng dữ liệu của họ ở đó không?

Câu hỏi cuối cùng là câu hỏi khó chịu. Bởi vì câu trả lời trung thực trong hầu hết các trường hợp là: không.

Bài viết này là bài đầu tiên trong chuỗi bài về cách vận hành thực sự của các công cụ liên lạc chuyên nghiệp. Các số tiếp theo sẽ đề cập đến việc tuân thủ GDPR trong nhắn tin và khái niệm bí mật nghề nghiệp trong kỷ nguyên số.

Nguồn tham khảo và đọc thêm

- Hayden, M. – Phát biểu tại Đại học Johns Hopkins, 2014 ("We kill people based on metadata"). Bản ghi chép công khai có sẵn.
- GDPR (Quy định EU 2016/679), Điều 4 và 5 – định nghĩa về dữ liệu cá nhân và các nguyên tắc xử lý (siêu dữ liệu là dữ liệu cá nhân).

- EDPS và EDPB – ý kiến về việc xử lý dữ liệu lưu lượng và siêu dữ liệu trong thông tin liên lạc điện tử (chỉ thị ePrivacy).

← TrướcLịch sử tóm tắt về dấu sápTiếp theo →Bí mật nghề nghiệp trong kỷ nguyên số

Các bài đọc gần đây

- CUADERNOS_LIST_PREGUNTAS_TITLE
- CUADERNOS_LIST_SELFHOST_TITLE
- CUADERNOS_LIST_IDENTIDAD_TITLE

Tải bài viết này để sử dụng ở bất cứ đâu bạn cần.

↓ Markdown ↓ Văn bản thuần túy ↓ PDF

Tệp sẽ được tải xuống thiết bị của bạn. Từ đó, bạn có thể lưu trữ, nhập vào Solo2 hoặc chia sẻ ở bất cứ đâu. Cuadernos không quyết định điểm đến thay cho bạn.

Dấu sáp · SHA-256 8889565d9eb0d2e69613c1a31de9cd5ee1fa801a2e79c4d668e3c776cc11ceaa

Cuadernos Lacre · Một ấn phẩm của Menzuri Gestión S.L. ·
viết bởi R.Eugenio · được biên tập bởi đội ngũ Solo2.