HÃY TỰ BẢO VỆ MÌNH TRƯỚC CÁC CÔNG CỤ TÌM KIẾM!

Xuất bản phẩm của AOL về nhật ký tìm kiếm của hơn 650 000 người dùng nhấn mạnh thêm một điểm quan trọng: những điều bạn đưa vào các trang trực tuyến có thể không được riêng tư như bạn nghĩ. Các công cụ tìm kiếm đặt hàng tỉ đô la cơ sở hạ tầng vào tay

Xuất bản phẩm của AOL về nhật ký tìm kiếm của hơn 650 000 người dùng nhấn mạnh thêm một điểm quan trọng: những điều bạn đưa vào các trang trực tuyến có thể không được riêng tư như bạn nghĩ. Các công cụ tìm kiếm đặt hàng tỉ đô la cơ sở hạ tầng vào tay bất kì người dùng ngẫu nhiên những người thường lưu lại Website của họ. Tuy nhiên cái giá mà bạn trả là để cho các công ty giữ truy vấn tìm kiếm của bạn, và điều đó có thể cung cấp một chút thông tin về đời sống của bạn. Đưa ra một số ý kiến để giữ gìn sự riêng tư trong khi sử dụng các công cụ tìm kiếm, CNET News.com đã chuẩn bị một loạt các câu hỏi thường xuyên sau: Q: Vì sao AOL công bố nhật ký các tìm kiếm đó? A: Mục đích nghiên cứu của AOL là công bố dữ liệu với hi vọng thông tin sẽ giúp các nhà khoa học và các nhà thống kê khác học hỏi được thêm về việc mọi người dùng internet như thế nào. AOL cũng đã xin lỗi về điều này vào hôm thứ hai vừa rồi. Họ nói rằng ấn bản này không được xem xét một cách đúng đắn. Q: Làm sao để tôi tự bảo vệ được mình nếu trong tương lai, các công cụ tìm kiếm vẫn thực hiện như thế. A: Bởi AOL cũng đã nhận được phản đối và sẽ nhanh chóng ngừng thực hiện những kiểu tìm kiếm đó. Nhưng trong bốn đại gia về tìm kiếm (AOL, Google, Microsoft và Yahoo) chỉ có Google là chống lại trát hầu toà của văn phòng thẩm phán. Q: Vì sao các công ty tìm kiếm lại lưu trữ thông tin của tôi sau khi việc tìm kiếm hoàn thành? A: Không có luật nào đòi hỏi các công ty phải xoá thuật ngữ tìm kiếm của bạn. Và có một số lí do thương mại để giữ lại chúng, ít nhất là trong một khoảng thời gian dù rất ngắn. Chẳng hạn giữ lại các thông tin cụ thể có thể giúp nhận dạng các cú click gian lận (như các lần click giả mạo trên quảng cáo Web để kéo giá đối thủ tăng lên). Hay trong việc tối ưu hoá kết quả tìm kiếm tại các khu vực địa lý khác nhau. Biên dịch một profile của người dùng có thể giúp kết quả tìm kiếm gắn kết lại (Chẳng hạn như bộ tìm kiếm Google Personalized Search). Hơn nữa đĩa lưu trữ khá rẻ nên các kĩ sư có khuynh hướng thích lưu trữ dữ liệu hơn là xoá chúng. Nhưng rõ ràng là chẳng có lý do thuyết phục nào để giữ các bản ghi đó lâu hơn, thậm chí tới vài tháng. Trừ khi một khách hàng tự nguyện chọn quyền lựa chọn của mình là personalization.. Q: Có bất kì công cụ tìm kiếm nào khác không lưu trữ thông tin sử dụng của người dùng không? A: Có, Ixquick.com, một bộ khởi động được tài trợ bởi Holland Ventures của Amsterdam. Công ty này tự hào nói rằng công cụ Ixquick.com không lưu giữ lại bất kì thông tin nào về địa chỉ internet của người dùng. Mặt khác, nó cũng không ghi lại các thuật ngữ tìm kiếm. Nhưng công ty cũng nói, nó không thể liên kết chúng tới bất cứ người nào, số nique ID hay địa chỉ Internet. “Chúng tôi là một hãng tin tưởng vào lí do riêng tư”. Giám đốc của Ixquick.com Robert Beens nói, trong cuộc phỏng vấn gần đây với CNET News.com. “Tôi có thể hình dung rất nhiều người đang hài lòng với sự riêng tư của họ”. Beens cũng nói rằng “Chúng tôi xoá địa chỉ (giao thức mạng internet) của người dùng. Chúng tôi có một chương trình chạy, mở các file đăng nhập, xoá địa chỉ IP của người dùng và viết đè lên chúng. Và công ty cũng gỡ bỏ ID unique từ cookies của Ixquick.com”. Q: Có phải AOL cũng định làm như thế? A: Không ai biết! Nhưng Jason Calacanis, người đồng sáng lập nhà xuất bản Weblogs Inc. mà AOL mua năm ngoái nói rằng họ nên làm như thế. Trong một blog hôm thứ hai, Calacanis viết: “Tôi muốn chúng ta trung thực, KHÔNG GIỮ THÔNG TIN ĐĂNG NHẬP trong việc tìm kiếm dữ liệu. Vâng, các bạn đã nghe rằng điều đó đúng… chúng ta thậm chí không nên giữ loại dữ liệu này”. Q: Ixquick.com làm việc như thế nào? A: Ixquick.com được biết đến như là một

công cụ siêu tìm kiếm. Để tìm kiếm truy vấn tại Mỹ nó liên hệ tới Yahoo, AltaVista, Alltheweb, Entireweb, Amazon, Netscape, Wikipedia và một ít website khác. Nó biên dịch kết quả, quyết định website nào nhận được nhiều phiếu thích đáng nhất và đưa ra các kết quả cao nhất. "Cũng có thể đánh lừa một máy tìm kiếm bằng cách chỉnh sửa các đường link, đuôi hoặc nội dung của website. Nhưng để đánh lừa được 11 máy tìm kiếm thì rất là khó" - Been nói. Q: Vâng, Ixquick.com thì rất là tốt. Nhưng tôi muốn dùng công cụ tìm kiếm yêu thích của mình. Làm sao tôi có thể bảo vệ sự riêng tư trong khi dùng nó? A: Việc đầu tiên bạn nên làm là xoá toàn bộ cookies mà các công cụ tìm kiếm thiết lập. Các cookies đó cho phép các công ty ghi nhận mỗi lần viếng thăm của bạn. Trong Firefox, vào Preferences and select Privacy. Tại đó có tuỳ chọn cho bạn xoá cookies và thậm chí ngăn chặn công cụ tìm kiếm thiết lập lại nó. (Thật không may, không phải tất cả các Web browser đều đưa ra tuỳ chọn này). Giả sử như bạn đang dùng Google, hãy thêm "google.com" vào danh sách của Firefox ở mặt không cho phép dùng cookies. Cảnh báo: điều này ngăn cản bạn dùng các tuỳ chọn kiểu persionalization hay Gmail. Nó giải thích vì sao có thể bạn vẫn muốn giữ các đường dẫn khác như Opera, Safari hay Internet Explorer. Nếu bạn thực sự lo lắng, hãy vào Anonymizer.com và đăng kí tuỳ chọn đường dẫn nặc danh (điều này chủ yếu cho người dùng Windows) . Tor là một lựa chọn thêm. Nó rất khó dùng, nhưng để bảo vệ sự riêng tư của bạn có thể nó sẽ có giá trị. Q: Ngoại trừ Ixquick.com ra, chúng ta có thêm thông tin nào về các công cụ tìm kiếm khác không? A: Chúng tôi đã khảo sát các công ty tìm kiếm vào tháng hai năm nay, và cũng đã hỏi họ đúng câu hỏi này. Trong buổi phỏng vấn đó: Đưa ra con số thuật ngữ tìm kiếm - Họ có thể cung cấp một danh sách người dùng internet (được nhận dạng bởi địa chỉ mạng và cookies) tìm kiếm với một thuật ngữ. Thứ hai, đưa ra một tập hợp các địa chỉ Internet - Họ có thể cung cấp một danh sách các thuật ngữ tìm kiếm với người dùng có địa chỉ Internet. Điều đó tạo ra một hồ sơ cá nhân điện tử hiệu quả. Q: Ai có thể truy cập vào danh sách các thuật ngữ tìm kiếm của tôi? A: Trong trường hợp bạn phạm tội ủy viên công tố chắc chắn có thể truy cập được vào những thông tin đó. Rõ ràng luật sư trong trường hợp dân sự như là người được uỷ quyền li dị, nhân viên phụ trách vấn đề về tranh cãi và chia cắt gia đình… cuối cúng cũng sẽ yêu cầu Google, Microsoft, Yahoo, AOL và các hãng tìm kiếm khác đưa ra nhật ký tìm kiếm của người dùng. Q: Việc này đã từng xảy ra trước đây chưa? A: Hầu hết! Một người đàn ông ở Bắc Carolina bị bắt vì tội giết người. Vì các từ khoá trong Google mà anh ta đã tìm là "cổ", "cắn", "đập", và "giữ" trước khi vợ anh ta bị giết. Nhưng những cụm từ tìm kiếm đó được tìm thấy trong máy tính của Robert Petrick chứ không được lấy trực tiếp từ Google. Q: Các địa chỉ Internet được đưa ra ngoài như thế nào? Có phải mọi người luôn chỉ có một cái giống nhau? A: Điều đó còn tuỳ. Nhiều nhà cung cấp cáp và DSL định vị địa chỉ mạng chỉ khi chúng được dùng (phương thức này được gọi là DHCP và PPPoe). Những địa chỉ IP đó có thể thay đổi thường xuyên. Những địa chỉ IP khác có khuynh hướng được sửa đổi. Khả năng và nhiệm vụ của mỗi thành viên tại trường đại học, các nhân viên trong công ty sửa đổi địa chỉ Internet ngày càng nhiều. Bộ tìm kiếm AOL là trường hợp duy nhất. Bởi vì người dùng của AOL có khuynh hướng đăng nhập khi dùng. Và AOL sẽ biết bạn là ai vì bạn cung cấp các thông tin đúng đắn khi đăng nhập dịch vụ. Q: Nếu Google biết tôi đang kết nối từ một địa chỉ internet động 192.1.1.1 vào ngày thứ nhất, 192.2.2.2 vào ngày thứ hai, 192.3.3.3 vào ngày thứ ba, làm sao nó có thể liên kết truy vấn của tôi lại với nhau để tạo thành hồ sơ đó? A: Đây chính là chỗ "cookies" được dùng. Một cookies đơn giản là một thiết bị để website nào đó nhận ra mọi người trong lần trở lại tiếp theo. Google, AOL, Yahoo, Microsoft, tất cả đều thiết lập cookies mặc định. (Hạn của Microsoft là năm 1016, Yahoo là năm 2010, Google là năm 2038, AOL thiết lập một cookie 3 phần có hạn là năm 2011). Trong ví dụ trên Google.com đã cài đặt

một cookie để bất cứ ai kết nối từ địa chỉ IP 192.1.1.1 vào ngày thứ nhất, và sau đó được đoán ra là có cùng web browser khi kết nối từ địa chỉ 192.2.2.2 và 192.3.3.3 hai ngày tiếp theo. Và tất nhiên nếu một người nào đó đăng nhập vào tài khoản Google của họ thì lại càng dễ dàng. Q: Các công ty giữ thông tin tìm kiếm của tôi trong bao lâu? A: Theo khảo sát của chúng tôi, Microsoft, Google và Yahoo đều nói rằng họ giữ dữ liệu lâu đến mức cần thiết. Và nó cũng có thể nghĩa là mãi mãi. Microsoft nói thêm rằng họ đang "xem xét các cách" để cung cấp cho người dùng tuỳ chọn xoá nhật ký tìm kiếm của họ. Yahoo cũng có câu tương tự. Và không biết chính xác được AOL giữ chúng trong bao lâu.