

Artificial intelligence can help you protect your personal data

February 16 2018



Credit: EPFL / Alain Herzog

It's a safe bet that some of the websites and apps you use collect and subsequently sell your personal data. But how can you know which ones? An EPFL researcher has led the development of a program that can answer that question in just a few seconds, thanks to artificial intelligence.

If you're like most people, you don't always take the time to read [website](#) terms and conditions before accepting them. Not only are they extremely lengthy, they are also convoluted and written in opaque legalese.

However, they can contain surprising clauses about a website's or app's right to use the data it collects about you, such as your IP address, your age and your online preferences. To help consumers get a better grasp of what they're agreeing to, a team of researchers from EPFL, the University of Wisconsin-Madison, and the University of Michigan have developed a program that uses [artificial intelligence](#) to decipher websites' data [protection](#) policies in the blink of an eye. Called Polisis, short for privacy policy analysis, their program can be used free of charge either as a browser extension (for Chrome or Firefox) or directly on their website.

"Our program uses simple graphs and color codes to show users exactly how their data could be used. For instance, some websites share geolocation data for marketing purposes, while others may not fully protect information about children. Such clauses are typically buried deep in their data protection policies," says Hamza Harkous, a post-doc working at EPFL's Distributed Information Systems Laboratory and the project lead.

With a little help from machine learning

The researchers used artificial intelligence to teach their program how to pick apart websites' data protection policies, drawing on over 130,000 that they found online. Once the text of a policy is fed into the program, the software scours through it in just a few seconds and displays the results in easy-to-read visuals. That lets you see at a glance which data a website would be authorized to collect and for what purpose. You can then make an informed decision about whether to use the website, or, in the case of an app, download it. The program also indicates what options you have for refusing to share certain data and lists the potential

disadvantages of each one.

Polisis works hand-in-hand with another program called Pribot, which is an online chatbot where you can enter questions (for now only in English) about a website's data protection policy. For example, you can type in "Does it share my [credit card information](#)?" and get a speedy answer. While Pribot, like Polisis, is not perfect – their results are for information only and offer no legal guarantee – it gives the right answer in the top 3 in around 82% of the time. A respectable score that could make it, along with its sister Polisis, extremely useful for consumers as well as journalists, researchers and data protection watchdogs.

Giving consumers a choice

Going forward, the team's program could be used for other applications such as the Internet of Things. If you're thinking about installing a connected object in your home, then you want to make sure its data protection [policy](#) is rock-solid. "We want to show consumers that they have a choice by giving them the tools to evaluate a service and select an alternative if necessary," says Harkous. His next goals are to develop an alert system that would notify users of any unexpected use of their data, and to create a system for ranking services and connected objects according to their data protection policies.

More information: Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning. arXiv:1802.02561 [cs.CL]
arxiv.org/abs/1802.02561

Provided by Ecole Polytechnique Federale de Lausanne

Citation: Artificial intelligence can help you protect your personal data (2018, February 16)

retrieved 19 April 2024 from <https://phys.org/news/2018-02-artificial-intelligence-personal.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.