Beneath the bonnet of online privacy

Article *in* EU Research · September 2023

1 author:

Huma Shah
Coventry University
66 PUBLICATIONS   **620** CITATIONS

SEE PROFILE

Beneath the bonnet of online privacy

The GDPR regulation was designed to give internet users a greater degree of control over their personal data, yet not all websites and apps are fully compliant and many still have trackers embedded. The CSI-COP project aimed to heighten awareness of what information is being gathered about us online, as **Dr Huma Shah** and **Professor Ian Marshall** explain.

The General Data Protection Regulation (GDPR) came into force across the European Union in 2018, regulating the way that cookies and other digital trackers can be used on websites to gather personal data. The regulation was designed to give individual internet users a greater degree of control over their personal data, which is reflected in its core principles. "For example, under GDPR websites need to be transparent when personal data is going to be collected, and the internet user has to give their informed consent, so tracking shouldn't be done by default. The reason behind collecting users' personal data also has to be clearly stated. What is the purpose?" outlines Dr Huma Shah, Assistant Professor in AI Ethics in the research centre for Computational Science and Mathematical Modelling at Coventry University. Despite this, some tracking cookies are still embedded in many websites, and internet users are not always notified. "For example, Google analytics and Facebook login are ubiquitous across different websites, such as some healthcare websites. So Facebook will know when you are using that healthcare website," explains Dr Shah. "Web development environments or platforms could have trackers embedded."

CSI-COP project

As the Science lead on the EU-funded CSI-COP project, Dr Shah is investigating the extent to which citizens are still being tracked online and looking at overall levels of compliance with GDPR. The wider aim in the project is to raise the level of scientific literacy amongst European citizens and heighten awareness of the privacy implications of unthinkingly accepting cookies when browsing the internet. "People may not really understand what they're accepting when they accept cookies and what they're giving away," says Dr Shah. The project is tackling the issue by training citizen scientists to investigate how cookies are used on different websites and apps, engaging the general public across each of the countries represented in the CSI-COP consortium. "We're an international consortium, with partners in eight countries, each of which has been looking to recruit a diverse group of citizens. In previous citizen science projects there have been relatively few women and young people, so we worked with youth organisations and women's groups, aiming to recruit a balanced cohort from the general public," continues Dr Shah.

The common characteristics among these citizens are that they all use the internet and share a concern about online privacy. An important initial step was to informally educate these citizen scientists through a short five-step course - 'Your Right to Privacy Online' - addressing key topics around online privacy. "What is privacy? What is personal data? How is our data harvested on the internet? What rights do we as citizens have? What free tools are available on the internet to help us manage our data?" outlines Dr Shah. Once they had completed the course, the interested citizens were asked if they would like to join the project and investigate the websites they visit and the apps they use, to find if they are clear about what cookies or tracking technologies exist beneath. The citizens recorded their investigations using a bespoke tool. "We didn't want any third-party tracking. So we're also showing our citizen scientists how to comply with GDPR," says Dr Shah. "The citizen scientists then provided their opinion - was the cookie notice clear about what was underneath? Was the privacy policy transparent? Then citizens used the free website and app audit tools from the internet to essentially go beneath the bonnet of websites and apps to see what exactly was under there. Our citizen scientists have so far looked at over 850 websites and 340 apps."

CSI-COP have already produced a 'taxonomy of cookies and trackers' from the citizen scientists' website and app investigations. This followed the project's work on data cleaning and classifying the different types of cookies. Dr Shah says "For example, e-commerce sites have session cookies to record what you put in the basket. Then there are persistent cookies, that are necessary for a website to work, while there will be others, like performance and analytics cookies. We went through the citizen scientists' findings to build that classification."

The way websites are monitored and analysed is evolving, and today most websites outsource the analytics. "People today want more information about what's happening on their website, such as the dwell time, and which pages are more interesting than others. Historically, you used to do that by putting little bits in your code, but nowadays people just outsource it," outlines Ian Marshall, PI in CSI-COP as well as Chief Operating Officer at Coventry University,

Behavioural information

A lot of website owners use Google analytics for example because it gives them behavioural information, such as which page on their website a user landed on and how long they stayed there.

The problem is that it's not only the website owner which gets the information, but also Google or whatever other company may be doing the analytics, which may not be fully compliant with GDPR.

Jaimz Winter (left) Coventry University's CSI-COP Research Assistant; Professor Maayan Zhitomirsky-Geffet (right), PI on CSI-COP at Bar Ilan University.

Citizen scientists (sitting: Lindsey Birnsteel -left; Sue Rowe - right) and Coventry University's CSI-COP team (Professor Ian Marshall and Dr. Huma Shah standing, Matthias Pocs - sitting), and stakeholder, Dr. Matthew England checking permissions in mobile apps in CSI-COP main Dissemination event in Brussels 23 May 2023.

CSI-COP citizen scientists' website and app investigations are available to look through from the project website, and from the Zenodo open-access platform:

Website investigations:
CSI-COP project site:
https://csi-cop.eu/project-results/citizen-scientists-website-investigations/

Zenodo:
https://zenodo.org/record/7472957

App investigations:
CSI-COP project site:
https://csi-cop.eu/project-results/citizen-scientists-app-investigations/

Zenodo:
https://zenodo.org/record/7472879

The document classifying cookies and trackers is available to download from the project website and from the Zenodo platform:

CSI-COP website:
https://csi-cop.eu/project-results/taxonomy-of-cookies-and-online-trackers/

Zenodo:
https://zenodo.org/record/7801846

The citizen scientists and the CSI-COP researchers' investigations were combined in the innovation of an open-access repository of online trackers. CSI-COP Repository was launched in the main project dissemination event in Brussels in May 2023 and is now freely available to search via the project website here: https://csi-cop.eu/repository/.

Seated: Coventry University CSI-COP technical sub-contractor, Mr Venkatesh Nanneboina, of Xcel Resources Ltd.; Dr. Huma Shah, lead scientist on CSI-COP.

"The Austrian data protection authority has decided that Google analytics is not compliant with GDPR. The regulation is very strict about where the data of European citizens goes," says Dr Shah. The bigger picture here is concern about the potential for companies to develop behavioural profiles of internet users by combining data from multiple sources. "When you start to build up a profile of the individual and identify behavioural traits, it moves from being marketing-led to behavioural-led," says Professor Marshall. "When you start taking information from Facebook, Google and other apps, and combining them, you get the potential to behaviourally profile people and then influence them."

Many people may be happy to accept cookies in certain circumstances, for example where it allows companies to learn about their personal preferences for certain products, but the possibility of profiling individual behaviour marks a significant step beyond this. Many of us automatically accept cookies without fully understanding what's being tracked, but there's no disadvantage to taking a bit longer to dig deeper and then make a decision, believes Professor Marshall. "You can then look at what's being tracked and whether you think it's acceptable. Most websites then give you the option to switch off the marketing stuff," he outlines.

The project offers a short MOOC (Massive Open Online Course) in English from the project website, which has been translated into twelve languages. The aim is to heighten awareness of these topics amongst the wider public through the MOOC and other events hosted by CSI-

COP partners across Europe and in Israel. "Over 600 people have completed the course in the available languages. We've held citizen science stakeholder cafes and parent/teacher roundtables, to increase reach," says Dr Shah. "Some of the project's citizen scientists have emerged as privacy champions were invited to attend CSI-COP's main event in Brussels, in May 2023, in person or online to share their experience of web and app investigations with a variety of stakeholders."

This can then help inform the regulation of data protection and online privacy as personal data gathering technology continues to evolve. While many companies have been responsible in terms of complying with GDPR, reducing the number of trackers and providing clear statements and policies, Professor Marshall says that others haven't. "We still come across websites which don't allow you to choose whether or not to accept cookies, or are designed in such a way that it's easiest to click accept and any other option is difficult to find. We've also found that some of the tools we have been using to try and identify trackers are being blocked," he outlines. There is currently no standardised way of writing cookie banners either, another topic which is being addressed in the project's policy recommendations. "During some CSI-COP events we have been asking attending members of the public, including school teachers in school visits, to play a short cookie game which involves timing yourself in avoiding cookies. It brings up real cookie banners to show how confusing it is," continues Dr Shah. "It can be difficult to avoid cookies, because the wording in a cookie banner or cookie notice can be ambiguous. You think you've rejected cookies, but actually you haven't, so there needs to be a standardised way of writing cookie banners."

PICASSO awards

The project's work in highlighting these kinds of issues was recognised at the inaugural Picasso Privacy Awards in London in December 2022, where CSI-COP won in the Best Innovative Privacy Project category. This recognised the project's novel approach to investigating online privacy, which is an increasingly prominent issue. "Engaging citizen scientists was a way to do this in a bottom-up way," explains Dr Shah. The aim of engaging the general public in this way was to inform citizens and enable them to make informed decisions about online privacy as they navigate the internet and deal with a rapidly changing technology landscape. "Technology has always accelerated beyond

society's ability to understand and to govern it. We're trying to educate as many people as we can reach and want to learn, so we can help them to understand the position we are currently in with respect to online privacy," outlines Professor Marshall. "Technology companies, governments and other organisations will continue to use the internet to deliver more and more services in future, many of which will be automated. We want to help people understand cookies and equip them with the knowledge they need to make an informed judgment."

A key step in this respect is standardising cookie banners, which is one of the main policy recommendations arising from the project's work, alongside the standardisation

of privacy policies. Another recommendation is that websites on EU-funded projects should be privacy-by-design, so ensuring that they are compliant with GDPR. "There should be limited third-party tracking in these websites, and when there is, their purpose should be made transparent," says Dr Shah. This is a highly dynamic field, and as a researcher in AI ethics and computer science, Dr Shah plans to continue working in this area beyond the conclusion of CSI-COP in August 2023. "There are data protection and privacy concerns from AI technologies, as well as issues around biases and misinformation," she continues. "I'm currently preparing a proposal which carries forward the work that has been done in CSI-COP."

Under GDPR websites and apps **need to be transparent** when personal data is going to be collected so the user is enabled to give their **informed consent**, so **tracking shouldn't be done by default.** The reason behind collecting users' personal data also has to be clearly stated. **What is the purpose?**

From left to right: Professor Yannis Gialelis, PI on CSI-COP at University of Patras; Mrs. Anna Michelle Adimakopoulou, a Member of the European Parliament (MEP); Dr. Huma Shah, and Jessica Mitchell, European Research Policy Manager, Coventry University Brussels Hub (Europe).

Dr Huma Shah

Dr Huma Shah is an Assistant Professor in the Research Centre for Computational Science and Mathematical Modelling at Coventry University. She gained her PhD in 2011 and is a Member of the European Platform for Women Scientists.