OLI BUCKLEY

IT'S NOT WHAT YOU TYPED, IT'S THE WAY YOU TYPED IT...

Typing patterns can predict a user's name and native language.

The way we type says a lot about who we are, with the rhythm and cadence of our keystrokes as identifiable as our handwriting or signature. However, it doesn't stop there. Keystroke Dynamics — the study of typing patterns, enables researchers to identify characteristics about the person at the keyboard. This includes things such as handedness, hand size, mood or typing style.

Our work, *Collecting and Leveraging Identity Cues using Keystroke Analysis* (CLICKA), evolved this idea of user identification to derive personal characteristics unique to the individual. This work focused on determining the name and native language of an anonymous user, based solely on *how* rather than *what* they typed.

The first experiment centred on determining the name of an anonymous user by collecting typing samples from 84 users. Participants completed several typing exercises, where the timing of each keypress and release was recorded. The research hypothesised that a user would type a familiar combination of keys more quickly. As such, the data were subdivided into short phrases containing two characters (bigrams) and ranked according to their typing speed.

> **The name prediction achieved a balanced accuracy of 70% of the bigrams in a user's name.**

The second experiment used a similar approach to determine the native language of an individual. Here, 492 participants were recruited from five native languages (English, French, German, Italian and Spanish), with an event split across each group.

The research used machine learning classifiers to develop models capable of predicting both a user's name and native language. The name prediction achieved a balanced accuracy of 70% of the bigrams in a user's name. Native language prediction achieved a balanced accuracy of 71% when comparing English against everything else. When predicting based on all five language categories, the accuracy dropped to 45% — still considerably better than a random prediction.

The key takeaway of this project is that it is possible to predict identifying characteristics about a user based on their typing patterns. This often requires a small sample of data, with participants only typing 200 to 300 words.

Dr Oli Buckley is an associate professor in Cyber Security, School of Computing Sciences at the University of East Anglia.

HEATHER SHAW

THE IDENTITY IN EVERYONE'S POCKET

When people interact with their smartphones, the digital traces left behind can be used to infer their identity.

Around a quarter of an adult's daily behaviour is spent on their smartphone (Ellis et al., 2019; Shaw et al., 2020). As such, smartphone usage data can reveal important insights into a person's daily habits and can be used to infer their identity.

> **It was possible to find within a top-10 list, the person whom the application usage data belonged to 75% of the time.**

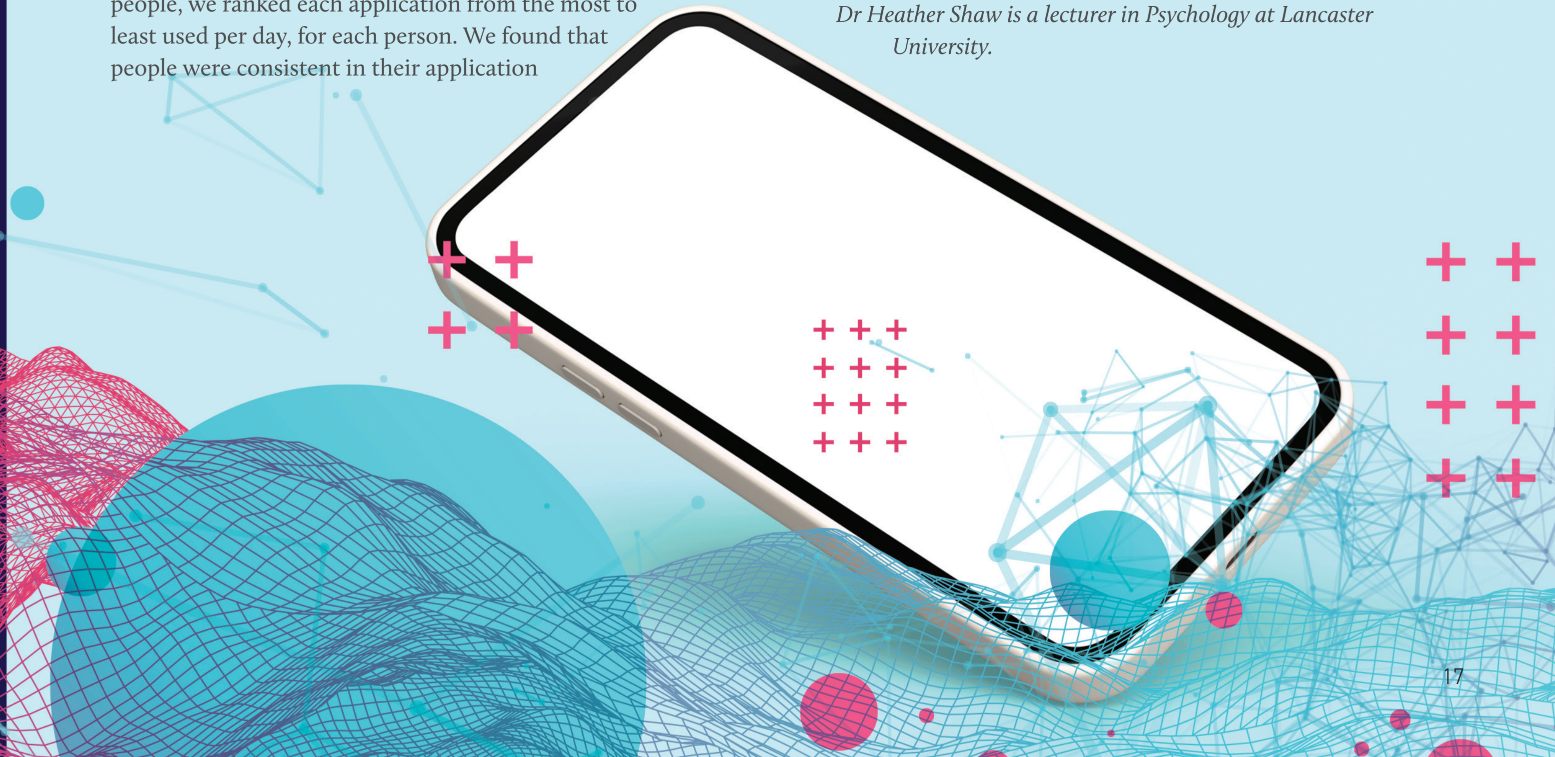
In our study of 28,692 days of smartphone data usage from 780 people, we ranked each application from the most to least used per day, for each person. We found that people were consistent in their application usage patterns on a day-to-day basis (e.g., consistently used Facebook the most and the calculator application the least). When we examined two randomly selected days from the same person, we found greater similarity in application use patterns than when we randomly selected two days that belonged to two different people.

To explore if application use could identify a single person, we fed 4,680 days of application usage data (equating to 6 days per person) into machine learning models. The models learned people's usage habits from the 6 days of application data and then tried to predict a person's identity when presented with an anonymous seventh day of data. The model was able to identify the correct person one-third of the time. Daily smartphone use can therefore act as a digital fingerprint.

The results further showed that it was possible to find within a top-10 list, the person to whom the application usage data belonged 75% of the time. In practical terms, this means that an investigation seeking to find a criminal's new phone from knowledge of their historic phone usage could reduce a pool of ~1,000 people's phones to 10 phones, with a 25% risk of missing them.

Our results suggest that access to smartphone application use data allows for a reasonable prediction about a person's identity even when they are logged-out of their account. This identification is possible with no monitoring of the conversations or behaviours within the applications themselves. Therefore, it is important to acknowledge that application usage data alone could risk our privacy if it is misused. It also questions whether usage data should be protected in the same manner as other personal identifiers.

Dr Heather Shaw is a lecturer in Psychology at Lancaster University.



LEON REICHERTS

"OK GOOGLE, SHOULD I CLICK ON THAT EMAIL?"

Designing conversational user interfaces to make us stop and think.

In recent years, data analytics tools have been given new features that enable users to query complex datasets using typed or spoken natural language. Instead of having to learn and use complex query syntax, analysts can now ask questions directly 'to' the data. Research has shown how these new ways to interact with data can improve both the user experience and task efficiency. However, central to data analysis is also knowing *what* to ask and coming up with *meaningful* questions. How can the next generation of analytics tools help users to generate more meaningful questions? This is where chatbots and voice assistants (sometimes referred to as 'conversational agents') can really come into their own, by being programmed to probe users to scaffold their questioning when using data analysis tools.

In our research group, we have begun researching how to augment human cognition by having an agent embedded in the software to proactively prompt users when looking at different data visualisations. We have found that agent prompts — even simple ones — can shift the users' attention to aspects of the data they would have missed or overlooked. It can also help them generate more exploratory questions.

> **One such interface protects against phishing attacks by helping users think more about suspicious emails.**

Our next steps are to find out whether this proactive agent approach supports more extensive data analysis and decision making in various contexts. We want to test whether such agents may, to some extent, mitigate challenges such as overconfidence or confirmation bias. We are also exploring how conversational agents can be designed to get people to 'slow down and think' when they are about to make risky decisions. Such an interface protects against phishing attacks by helping users think more about suspicious emails, enabling them to examine specific aspects of the email before deciding whether to click the potentially harmful URL.

This line of research suggests there are new opportunities for extending the reach of chatbots and conversational agents beyond their current home; instead of answering users' queries they can instead question them, encouraging people to think in new ways.

Leon Reicherts is a PhD student at the UCL Interaction Centre and part of the Ecological Brain DTP (ecologicalbrain.org). His research focuses on how to design conversational interfaces to support human cognition when performing complex data analysis and decision-making tasks.

RICHARD PHILPOT & MARK LEVINE

CCTV ANALYSIS OF VIOLENT EMERGENCIES

Systematic analysis of CCTV footage of violent and dangerous emergencies can help us understand how people behave during times of heightened security threats.

Whether it is incidents of street violence or marauding terrorist attacks, the fact that these events are invariably captured on public space CCTV means we can build a robust evidence base about behaviour in real-life emergencies.

However, CCTV data can be complex, incomplete and lacking both sound and wider contextual information. One way to get around this (and in doing so, extract the most reliable evidence from the CCTV data) is building an appropriate ethogram — a list of relevant behaviours in a particular context.

Using an ethogram approach we analysed CCTV footage of street violence in the UK, the Netherlands, and South Africa and were able to show that contrary to conventional wisdom, bystanders intervened in more than 90% of aggressive public incidents.

These tended to be coordinated interventions, with three to four bystanders working together to calm the violence. Bystanders were also at low risk of victimisation when intervening to help (Liebst et al., 2021).

> **...contrary to conventional wisdom, bystanders intervened in more than 90% of aggressive public incidents.**

In another micro-behavioural analysis of CCTV footage of an explosion in a single railway carriage (Philpot & Levine, 2021), we also showed how emergency response behaviours can be shaped by the actions of immediate others — but that the behaviours themselves can be different in different places. Proximity to the explosion site is seemingly less important than the behaviour of the people around you.

The kinds of analysis that can be done is often shaped by data availability. It's not always possible to collect data systematically, and access to CCTV footage from some incidents might be limited by ethical, legal or security concerns.

The strength of analysing CCTV data is that it not only provides a richer understanding of behaviour in emergencies (compared to research which uses self report methods), it also allows us to test the assumptions of existing models that underpin emergency preparedness. As more footage becomes available, we will continue to develop important new insights that improve security and resilience planning.

Dr Richard Philpot is a lecturer of Psychology at Lancaster University. Applying digital data, his research examines how citizens and emergency services behave and interact during spontaneous public space emergencies.

Mark Levine is a professor of Social Psychology at Lancaster University. His research explores the role of identities and group processes in pro-social and anti-social behaviour.

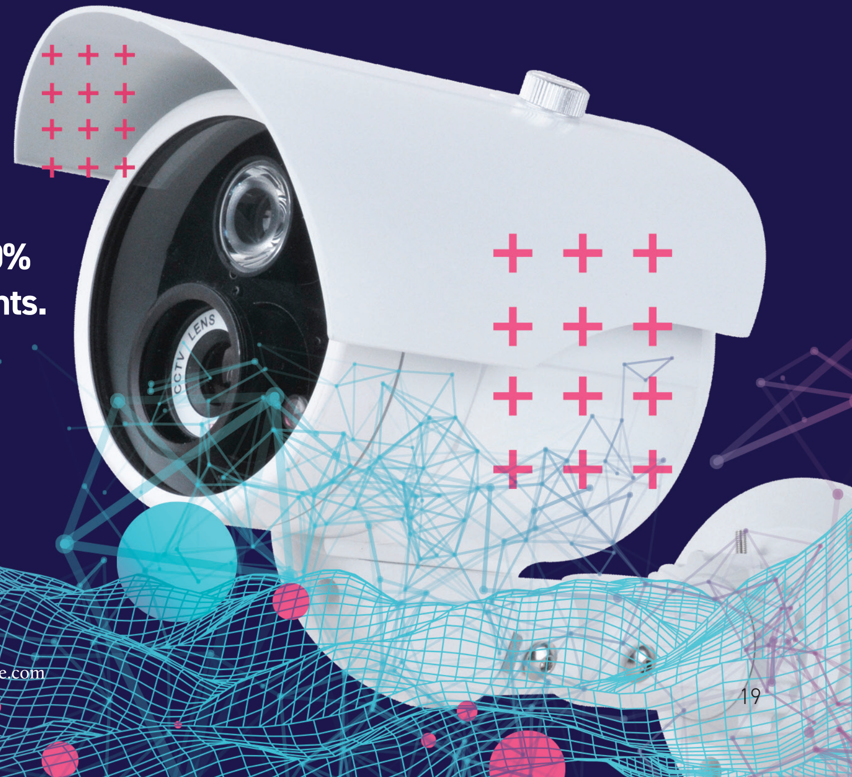


Image credit: Adapted from *Christian Horz* and other images from stock.adobe.com