INVESTIGATION · DIGITAL INVESTIGATION

Investigating Online Identities and Digital Footprints

People leave traces wherever they go online. This article explains how investigators use those traces to establish who is behind an account, how identities can be linked across platforms, how reliable such links are, and the legal and ethical limits on this kind of work.

ModQu research article · 2,636 words · 25 references (APA 7th edition)

Every time someone creates an account, posts a comment, uploads a photograph, buys something or uses an app, they leave a trace. Taken together, these traces form a digital footprint: a record of a person's activities, associations, locations and interests spread across many services. Investigators use digital footprints to find out who is behind an anonymous account, whether two accounts belong to the same person, whether someone is who they claim to be, and what a person has said and done in public.

The same traces that make online identity investigation possible also make it risky. Footprints are incomplete, easily misread and sometimes deliberately faked, and a mistaken identification can cause serious harm to an innocent person. This article explains what digital footprints consist of, the methods used to attribute accounts and link identities, how to judge the strength of a link, and the legal and ethical limits on the work.

What a digital footprint contains

Digital footprints include information that people publish deliberately and information that is recorded about them as a side effect of using services. Clarke (1994) introduced the idea of a "digital persona": a model of an individual built from data, which can be used to make decisions about the person whether or not it is accurate. Social network sites, in particular, encourage users to construct public profiles and to display their connections to others (boyd & Ellison, 2007).

Investigators distinguish several kinds of trace:

- **Self-published content:** profiles, posts, photographs, videos, comments and reviews.
- **Identifiers:** usernames, email addresses, telephone numbers and account numbers, which are often reused across services.
- **Connections:** friends, followers, groups and interactions, which show who a person associates with.
- **Metadata:** information about content, such as when a post was made, from what device, and sometimes from where.
- **Records held by organisations:** registration details, login records, payment data and IP addresses, which are not public but may be obtained by legal process.

Footprints reveal more than people realise. Kosinski et al. (2013) showed that the pages people "liked" on Facebook could be used to predict sensitive attributes such as sexual orientation, political views and religion with considerable accuracy. Research on mobility data has shown that a few location points are enough to identify most individuals in a large dataset (de Montjoye et al., 2013), and that even datasets stripped of names can often be re-identified by linking them with other data (Narayanan & Shmatikov, 2008; Rocher et al., 2019). Sweeney (2000) found that a combination of date of birth, sex and postal code was enough to identify a large majority of the United States population uniquely.

People also disclose more online than they might in person. Suler (2004) described an "online disinhibition effect", in which the sense of anonymity, invisibility and distance leads people to say and reveal things they would not say face to face. For investigators, this means that online traces can be unusually revealing; for the people concerned, it means that material posted casually may later be read in a very different context.

Methods of attribution

Attribution means establishing who is responsible for an account or a piece of content. Investigators approach it from several directions, building up a picture from many small pieces of evidence (Bazzell, 2023).

Identifiers and reuse

The most direct links come from identifiers. A username used on one platform may be used on others; an email address or telephone number linked to one account may appear in another context, such as a company registration or an old forum post. Profile photographs can be searched with reverse image tools to find the same image elsewhere. Reuse is common because people find it convenient, and a single reused identifier can connect an anonymous account to a named one.

Content and context

What a person posts often reveals who they are: photographs of their home, their workplace, their pets or their car; references to events they attended; comments about their job or their town. Each detail narrows the range of possible authors. Location can be inferred from landmarks, shadows, weather and signs in photographs, as open-source investigators have shown in verifying images from conflict zones (Higgins, 2021; Dubberley et al., 2020).

Behaviour and timing

Patterns of activity are also informative. The hours at which an account posts suggest a time zone and daily routine. Two accounts that are always active at the same times, never at the same moment, and post about the same topics may be operated by the same person. Accounts that post in coordinated bursts may be part of a network, and automated accounts can often be detected by the regularity and volume of their activity (Ferrara et al., 2016).

Writing style

Every writer has habits of vocabulary, spelling, punctuation and sentence structure. Authorship attribution uses these features to estimate who wrote a text, by comparing it with writing samples from candidate authors (Juola, 2006; Stamatatos, 2009). The methods can be powerful when there are enough text and a small set of plausible authors, but they are far less reliable with short texts, large numbers of candidates or writers who deliberately disguise their style.

Records held by platforms

The strongest evidence of who controls an account is usually held by the platform: registration details, login IP addresses, device identifiers and payment records. Investigators can obtain these only through lawful process, such as a court order or a request by law enforcement under applicable law. In civil cases, courts can order platforms to disclose information identifying anonymous users who are alleged to have committed wrongs, but they generally require evidence that the claim is well founded and weigh the user's interest in anonymity.

Judging the strength of a link

Linking two accounts, or an account and a person, is an inference. Some links are strong; others are weak and easily coincidental. A good investigation assesses each link separately and considers how the links combine.

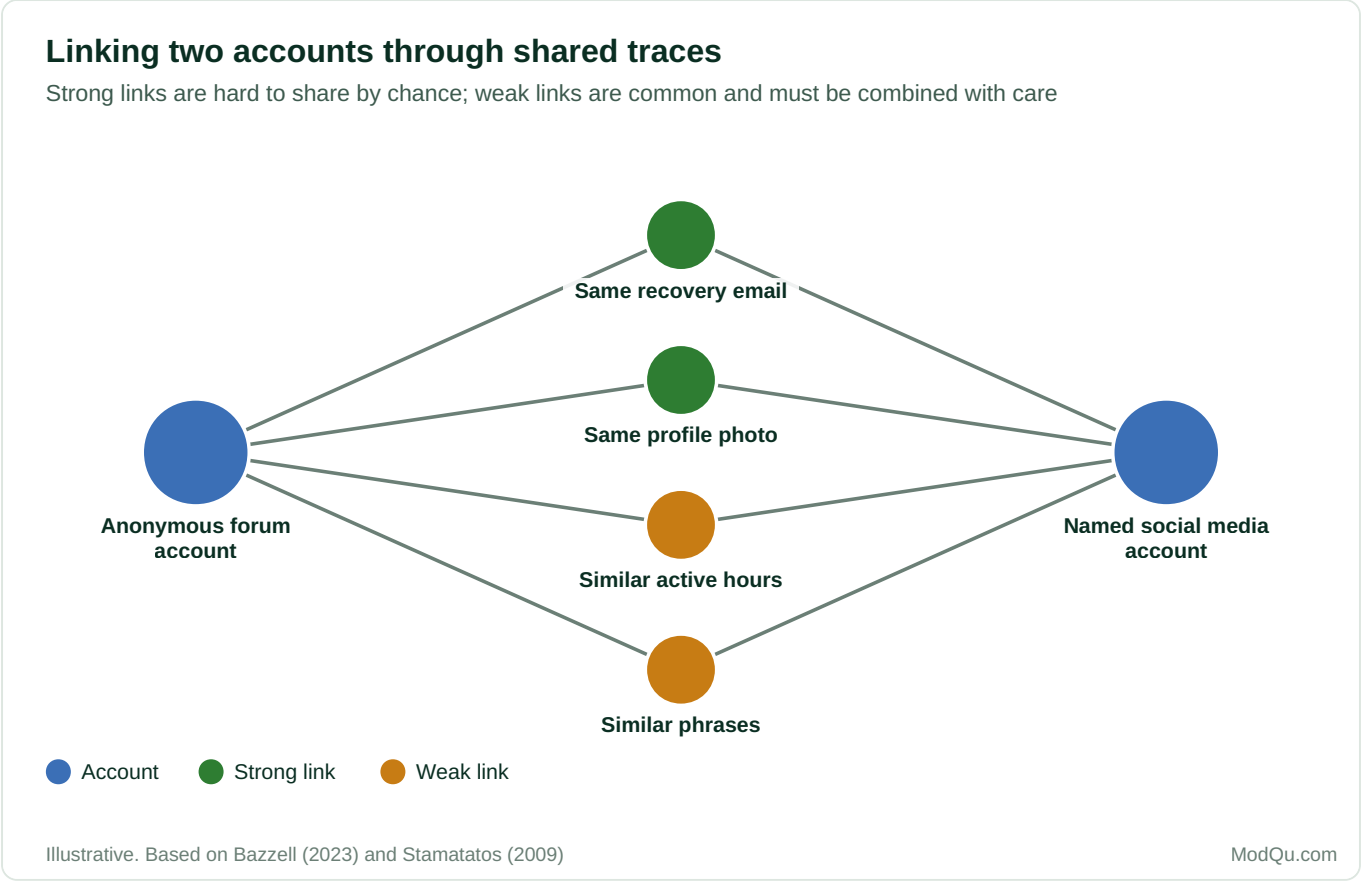


Figure 1. Two accounts can be linked by several kinds of shared trace. A shared recovery email or original photograph is strong evidence; similar hours of activity or turns of phrase are weak and may be shared by many people. Diagram by ModQu.

Several questions help to judge a link:

- **How many people could share this trace?** A unique email address is shared by one person; a common username, a popular stock photograph or a time zone is shared by many.
- **Could the trace have been copied or faked?** Photographs can be taken from other people's profiles, writing styles can be imitated, and identifiers can be used to impersonate someone or to frame them.
- **Are the links independent?** Two links that rest on the same underlying fact add little. A photograph and a caption copied together from one source count as one piece of evidence, not two.
- **What alternative explanations exist?** Shared accounts, family members using the same device, impersonation and coincidence should all be considered.

The danger of weak links is shown by cases of crowd-sourced investigation that went wrong. After the Boston Marathon bombing in 2013, users of online forums tried to identify the bombers from photographs and wrongly named several innocent people, one of whom had been missing and was later found dead; the misidentifications spread widely before they were corrected (Nhan et al., 2017). Trottier (2017) describes how "digital vigilantism" can expose people to harassment on the basis of online accusations that have not been tested.

Investigating networks of accounts

Some investigations concern not a single person but a network of accounts acting together. Influence operations, harassment campaigns, review fraud and scams often use many accounts that appear independent but are controlled by a small group. Investigators look for signs of coordination: accounts created at around the same time, with similar names or profile images; accounts that post the same or nearly the same content within minutes of each other; accounts that follow and amplify one another; and accounts whose activity starts and stops together. Automated accounts can often be identified by the volume and regularity of their activity, although more sophisticated operations mix automation with human operators (Ferrara et al., 2016).

Network analysis makes these patterns visible. By mapping which accounts interact, share content or use the same links, investigators can identify clusters that behave as a unit. The same caution applies as for individual attribution. Genuine communities also share content and react to the same events, and coordination does not by itself show who is behind a network or why. Strong conclusions usually require evidence from the platforms themselves, such as shared devices, payment methods or login patterns, which only they or authorities with legal powers can obtain.

Preserving online evidence

Online content can disappear at any time. Posts are deleted, accounts are suspended, pages are edited and websites go offline. Studies of web links cited in legal opinions and journals have found that a large share no longer lead to the material originally cited, a problem known as link rot (Zittrain et al., 2014). Investigators therefore capture evidence as soon as they find it, recording the full web address, the date and time of capture, the method used and, where possible, a hash value of the captured file.

Screenshots are useful but limited, since they can be edited and do not capture the underlying page. Better practice combines screenshots with saved copies of the page, archived versions from independent archiving services and notes describing how the material was found. The Berkeley Protocol recommends that online investigators document each step so that another person could follow the same path and verify the result, and that captured material be stored securely with a record of who has handled it (Human Rights Center, University of California, Berkeley, School of Law, & Office of the United Nations High Commissioner for Human Rights, 2022). These practices apply the principles of the chain of custody to material that has no physical form.

Investigators should also take care not to alert the subject. Visiting a profile can, on some platforms, notify its owner or appear in visitor lists. Searching from an account linked to the investigator's real identity, or clicking links sent by the subject, can reveal the investigation or expose the investigator to tracking. Organisations that carry out online investigations regularly use dedicated equipment and accounts, set up in accordance with the law and their own policies, to separate investigative work from personal use.

Verifying identity claims

Investigators are often asked not who is behind an account, but whether a person is who they say they are: a business contact, a job candidate, an online seller or a source. Verification involves checking that the person's claimed history is consistent across sources, that their profile has a plausible history rather than appearing suddenly, that photographs are original rather than taken from elsewhere, and that contact details lead to the person and organisation claimed. Verification methods developed by journalists, such as tracing images to their earliest appearance and contacting sources through independently obtained details, apply directly (Silverman, 2014). Fake personas are used in fraud, espionage and influence operations, and can be convincing on the surface (Wardle & Derakhshan, 2017).

Reporting attribution findings

Attribution findings should be reported with care, because readers tend to treat a name in a report as a conclusion. A good report sets out each link between accounts or between an account and a person, the evidence for it, how strong it is and what alternative explanations were considered. It then states an overall conclusion in calibrated language, such as "the evidence strongly supports", "is consistent with" or "is insufficient to establish", and explains what further evidence would change it.

Some investigators use a scale of confidence borrowed from intelligence analysis, in which each level corresponds to a stated range of probability, so that readers understand how much weight the conclusion can bear. Whatever scale is used, it should be applied consistently and defined in the report. It is also good practice to separate the facts found, such as "the same photograph appears on both accounts and no earlier copy was found", from the inference drawn, such as "the two accounts are probably operated by the same person". A reader who disagrees with the inference can then see exactly where the disagreement lies.

Legal and ethical limits

Online identity investigation raises serious privacy questions, even when it uses only public information. Nissenbaum (2004) argued that privacy concerns the flow of information in context: information shared publicly in one setting, such as a hobby forum, may still be misused when collected and combined with other information for a different purpose. Solove (2006) identified aggregation, the combining of individually innocuous pieces of information into a revealing profile, as a distinct privacy harm.

The law reflects these concerns. In the European Union and the United Kingdom, collecting and processing personal data, including data that are publicly available, requires a lawful basis and must be necessary and proportionate (Regulation (EU) 2016/679, 2016). Individuals can ask search engines to remove links to information about them that is inadequate or no longer relevant (*Google Spain SL v Agencia Española de Protección de Datos*, 2014). Creating false accounts to obtain information, accessing private content without authorisation or obtaining records by deception may be unlawful, and evidence obtained in these ways may be excluded or may expose the investigator to liability.

Ethical practice goes beyond legal compliance. The Berkeley Protocol on digital open source investigations recommends that investigators minimise harm to the people whose information they collect, consider the risks to sources and subjects, protect their own security, and document their work so that it can be reviewed (Human Rights Center, University of California, Berkeley, School of Law, & Office of the United Nations High Commissioner for Human Rights, 2022). Commercial data about individuals can be wrong or outdated, and relying on it without checking can cause real harm (Lageson, 2020).

A short example

The following hypothetical case illustrates the approach. A company receives a series of anonymous posts on an industry forum alleging misconduct by one of its managers. The posts contain details that suggest an insider. The company's investigator, working under legal advice, examines only public information.

The forum account was created on the day of the first post. Its username does not appear elsewhere, but its profile picture, a photograph of a distinctive mountain lake, appears through reverse image search on a public photography site under a different username. That username is used on a public code repository, where the account's commits list an email address at a personal domain. Domain registration records are privacy-protected. The posting times match a time zone different from the company's headquarters but consistent with one of its regional offices.

The investigator's report sets out each link and its strength. The photograph is a strong link between the forum and photography accounts if it is original, and searches found no earlier copy. The connection to the code repository is strong. Nothing yet links any of these accounts to a named employee, and the time zone evidence is weak, since many people live in that region. The company decides to investigate the substance of the allegations rather than the identity of the poster, recognising that the posts may be a protected disclosure and that trying to unmask the author could amount to retaliation.

Conclusion

Digital footprints allow investigators to attribute accounts, link identities and verify claims with a speed that would have been impossible a generation ago. But each link is an inference, and inferences built on weak, copied or coincidental traces can identify the wrong person. Sound practice assesses every link on its own merits, considers alternative explanations, records the reasoning, and respects the legal and ethical limits on collecting and combining personal information. The question is not only whether a person can be identified, but whether the evidence justifies the identification and whether there is a legitimate reason to make it.

References

Bazzell, M. (2023). *OSINT techniques: Resources for uncovering online information* (10th ed.). IntelTechniques.

boyd, d. m., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication, 13*(1), 210–230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

Clarke, R. (1994). The digital persona and its application to data surveillance. *The Information Society, 10*(2), 77–92. https://doi.org/10.1080/01972243.1994.9960160

de Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific Reports, 3*, Article 1376. https://doi.org/10.1038/srep01376

Dubberley, S., Koenig, A., & Murray, D. (Eds.). (2020). *Digital witness: Using open source information for human rights investigation, documentation, and accountability*. Oxford University Press.

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), 96–104. https://doi.org/10.1145/2818717

Google Spain SL v Agencia Española de Protección de Datos (C-131/12) EU:C:2014:317.

Higgins, E. (2021). *We are Bellingcat: An intelligence agency for the people*. Bloomsbury.

Human Rights Center, University of California, Berkeley, School of Law, & Office of the United Nations High Commissioner for Human Rights. (2022). *Berkeley protocol on digital open source investigations: A practical guide on the effective use of digital open source information in investigating violations of international criminal, human rights and humanitarian law*. United Nations.

Juola, P. (2006). Authorship attribution. *Foundations and Trends in Information Retrieval, 1*(3), 233–334. https://doi.org/10.1561/1500000005

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences, 110*(15), 5802–5805. https://doi.org/10.1073/pnas.1218772110

Lageson, S. E. (2020). *Digital punishment: Privacy, stigma, and the harms of data-driven criminal justice*. Oxford University Press.

Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy* (pp. 111–125). IEEE. https://doi.org/10.1109/SP.2008.33

Nhan, J., Huey, L., & Broll, R. (2017). Digilantism: An analysis of crowdsourcing and the Boston marathon bombings. *The British Journal of Criminology, 57*(2), 341–361. https://doi.org/10.1093/bjc/azv118

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review, 79*(1), 119–157.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). (2016). *Official Journal of the European Union, L 119*, 1–88.

Rocher, L., Hendrickx, J. M., & de Montjoye, Y.-A. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. *Nature Communications, 10*, Article 3069. https://doi.org/10.1038/s41467-019-10933-3

Silverman, C. (Ed.). (2014). *Verification handbook: A definitive guide to verifying digital content for emergency coverage*. European Journalism Centre.

Solove, D. J. (2006). A taxonomy of privacy. *University of Pennsylvania Law Review, 154*(3), 477–564. https://doi.org/10.2307/40041279

Stamatatos, E. (2009). A survey of modern authorship attribution methods. *Journal of the American Society for Information Science and Technology, 60*(3), 538–556. https://doi.org/10.1002/asi.21001

Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior, 7*(3), 321–326. https://doi.org/10.1089/1094931041291295

Sweeney, L. (2000). *Simple demographics often identify people uniquely* (Data Privacy Working Paper 3). Carnegie Mellon University.

Trottier, D. (2017). Digital vigilantism as weaponisation of visibility. *Philosophy & Technology, 30*(1), 55–72. https://doi.org/10.1007/s13347-016-0216-4

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking* (Council of Europe Report DGI(2017)09). Council of Europe.

Zittrain, J., Albert, K., & Lessig, L. (2014). Perma: Scoping and addressing the problem of link and reference rot in legal citations. *Harvard Law Review Forum, 127*, 176–199.