INVESTIGATION · DIGITAL INVESTIGATION

Mobile Device Forensics and App Data Extraction

Smartphones hold messages, photographs, locations, contacts, health data and app records that can be central to an investigation. This article explains how mobile device evidence is preserved and extracted, the levels of extraction, where app data are stored, the challenges of encryption and the cloud, the legal limits on searching phones, and how to keep extraction reliable and proportionate.

ModQu research article · 2,399 words · 22 references (APA 7th edition)

For many people, the smartphone is the most complete record of their lives. It holds their messages and calls, photographs and videos, contacts and calendars, browsing history, location history, health and fitness data, banking and payment records, and the data of hundreds of apps. In criminal, regulatory and workplace investigations alike, a phone can show who communicated with whom, where someone was, what they knew and when. It can also exonerate. Mobile device forensics is the discipline of preserving, extracting, analysing and reporting this evidence in a way that is reliable and lawful. This article explains how it is done, what it can and cannot reveal, and the safeguards that make it trustworthy.

Why phones are different

Mobile devices differ from traditional computers in ways that make forensic work harder. They come in thousands of models, run frequently updated operating systems, store data in proprietary formats, and are connected to networks that can change their contents remotely. Most modern phones encrypt their storage by default and tie the encryption keys to the user's passcode and to secure hardware within the device. Barmpatsalou et al. (2013), reviewing seven years of research, noted that the rapid pace of change in devices and operating systems means that forensic methods quickly become outdated, and that no single tool or method works for all devices.

The guidance published by the United States National Institute of Standards and Technology describes mobile device forensics as the science of recovering digital evidence from a mobile device under forensically sound conditions using accepted methods, and emphasises that the variety of devices and the volatility of their data require particular care (Ayers et al., 2014).

Preservation

The first priority is to prevent the data on a seized phone from being changed or lost. A phone that remains connected to a network can receive new messages that overwrite deleted data, be remotely wiped by its owner, or have its location history updated. Standard practice is to isolate the device from all networks, for example by placing it in a shielded container or enabling flight mode, and to keep it charged, because a phone that powers off may return to a state in which its data are fully encrypted and inaccessible without the passcode (Ayers et al., 2014).

Each step must be documented: who seized the device, when, in what condition, what was visible on the screen, and how it was stored and transferred. International standards on digital evidence set out requirements for identifying, collecting, acquiring and preserving potential evidence in ways that maintain its integrity (International Organization for Standardization [ISO], 2012). United Kingdom guidance similarly requires that no action should change data that may later be relied on in court, and that where this is unavoidable, the person doing it must be competent and able to explain the consequences of their actions (Association of Chief Police Officers [ACPO], 2012). With phones, some change is often unavoidable, because the device must usually be switched on and interacted with to extract data. The key is to minimise and record it.

Levels of extraction

Extraction methods are commonly arranged in levels, from simple and widely available to complex and specialised.

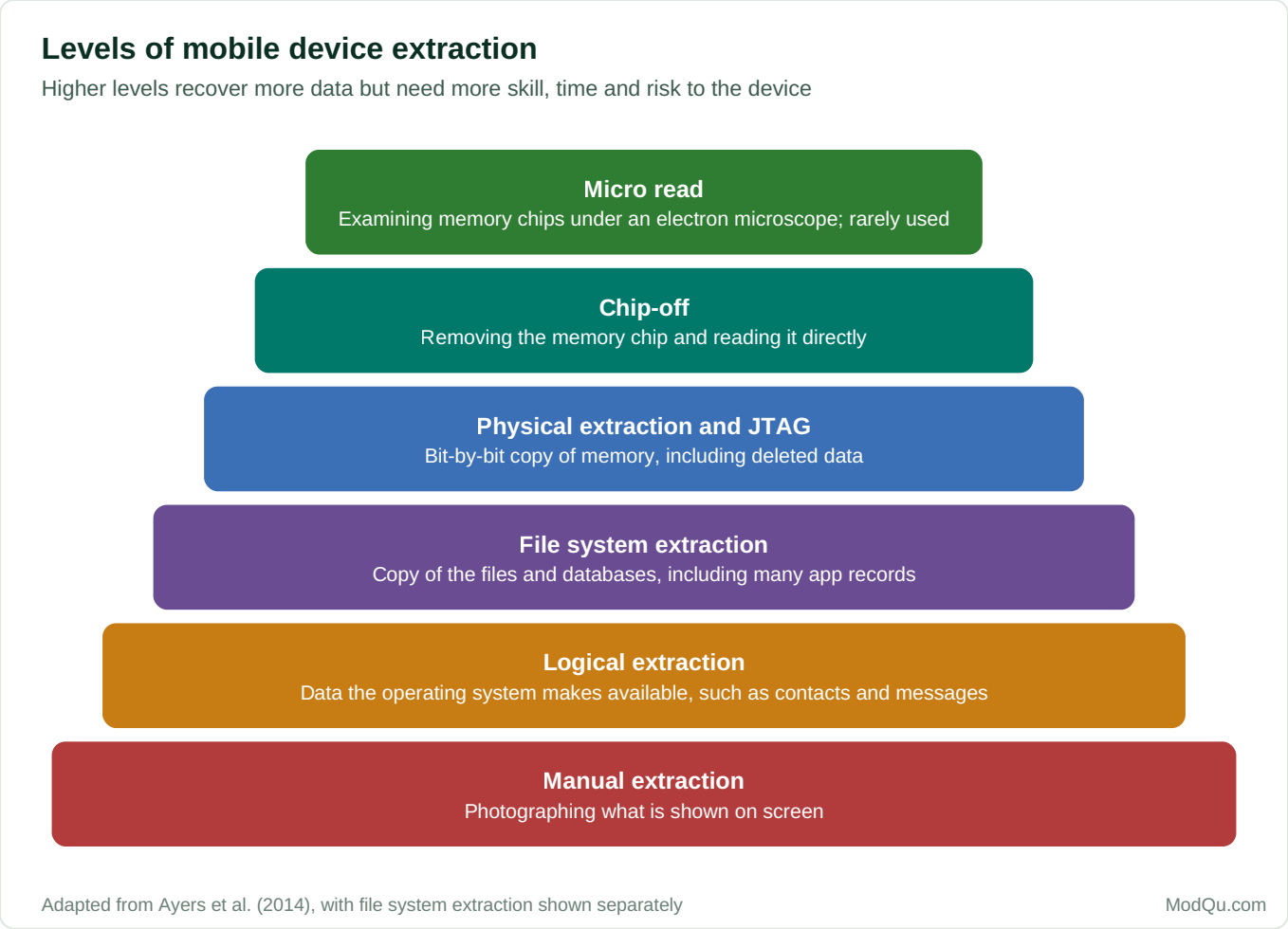


Figure 1. The classification of extraction methods used in forensic guidance, with file system extraction added as a separate level. Most examinations use logical or file system extraction; physical and chip-off methods are used when lower levels cannot reach the data needed. Modern encryption limits what physical methods can recover. Diagram by ModQu.

Manual extraction involves an examiner operating the phone and photographing what appears on the screen. It is simple and works on almost any unlocked device, but it captures only what is displayed, takes time, and risks changing data.

Logical extraction uses the device's own interfaces to obtain data that the operating system makes available, such as contacts, call logs, messages and media. It is fast and widely supported but omits deleted items and much app data.

File system extraction obtains a copy of the device's file system, including the databases in which apps store their records. This is the level at which much app data becomes available, and it can include deleted records that remain within databases.

Physical extraction obtains a bit-by-bit copy of the device's memory. In principle it captures everything, including unallocated space where deleted data may remain. Hardware methods such as JTAG, which uses test connections on the circuit board, and **chip-off**, which involves removing the memory chip and reading it directly, can achieve this when software methods cannot (Breeuwsma et al., 2007). On modern phones with full-disk or file-based encryption, however, a physical copy is often unreadable without the keys held in the device's secure hardware.

Micro read, the examination of memory cells under an electron microscope, is described in guidance but is impractical for almost all investigations (Ayers et al., 2014).

In practice, examiners choose the least intrusive method that will obtain the data needed, and record why they chose it.

Where app data live

Most evidence on a modern phone comes from apps. Messaging apps, social media, email clients, browsers, maps, ride-hailing, payment, fitness and dating apps all keep records on the device, typically in databases, property lists, caches and media folders. Analysts must understand how each app stores its data: which database tables hold messages, how timestamps are encoded, how attachments are linked to messages, and what happens when a user deletes something. Practical guides describe the storage structures of common apps on Android and iOS (Hoog, 2011; Tamma et al., 2020), but these structures change whenever an app is updated, so examiners must regularly test their tools and methods against new versions.

Common sources of evidence on a phone

Each source answers different questions and has typical pitfalls

	What it can show	Typical pitfalls
Messages and calls	Who communicated, when and what was said	Deleted or disappearing messages; shared devices
Photos and videos	Events, places, people; metadata on time and location	Edited files; metadata stripped by apps
Location data	Where the device was, from GPS, Wi-Fi and cell data	Precision varies; device may not be with its owner
Browser and search history	Interests, knowledge and intentions	Automatic or synced entries the user never typed
Health and activity data	Steps, movement and sleep at particular times	Device not worn; estimates rather than measurements
Cloud and sync records	Data from other devices on the same account	May reflect another device or user

ModQu analysis based on Ayers et al. (2014), Hoog (2011) and Tamma et al. (2020) ModQu.com

Figure 2. Phone data can answer many questions, but each source can be misread. Findings should always be interpreted in context and, where possible, confirmed by independent evidence. Diagram by ModQu.

Interpretation is as important as extraction. A timestamp may record when a message was sent, received, read or synchronised, and may be stored in a different time zone or format from what a tool displays. A location record may reflect a Wi-Fi network estimate accurate to hundreds of metres rather than a precise satellite position. An entry in browser history may have been created by a link opened automatically, or synchronised from another device logged into the same account. Analysts must test these meanings rather than accept a tool's display at face value.

Encryption and locked devices

Encryption is the greatest barrier to mobile forensics. Modern phones protect their data with keys derived from the user's passcode and protected by secure hardware that limits the number of guesses. Without the passcode, extraction may be impossible or may depend on exploiting security vulnerabilities, which commercial forensic vendors develop and sell. End-to-end encrypted messaging apps protect messages in transit, so they cannot be obtained from the service provider, although they are usually readable on an unlocked device. The Signal protocol, which underlies several such apps, has been formally analysed and found to provide strong security properties (Cohn-Gordon et al., 2020).

The tension between encryption and investigation became public in 2016, when the United States government obtained an order requiring Apple to help unlock an iPhone used by one of the perpetrators of a shooting in San Bernardino (*In re Search of an Apple iPhone*, 2016). Apple resisted, arguing that creating the software would undermine the security of all

its users. The government withdrew the case after obtaining access by another method. Abelson et al. (2015), a group of leading security experts, had argued that requiring exceptional access for government would create vulnerabilities that others could exploit.

The cloud

Much of the data associated with a phone is no longer on the phone. Messages, photographs, notes, backups and location histories are often stored in the cloud and synchronised across devices. This can help investigators, because cloud data may survive the loss or wiping of the phone, but it raises legal questions: data held by a provider are usually obtained through legal process directed at the provider, not by extraction from the device, and may be stored in another country. Examiners must also be careful not to access cloud accounts from a seized phone without lawful authority, because doing so may go beyond what a search of the device permits.

Legal limits

Courts have recognised that searching a phone is more intrusive than searching a physical object. In *Riley v. California* (2014), the United States Supreme Court held that police generally need a warrant to search the digital contents of a phone seized during an arrest, noting that phones contain "the privacies of life". In *Carpenter v. United States* (2018), it held that historical location records from a mobile network generally require a warrant. Kerr (2005) argued that rules developed for physical searches needed to be adapted to computers, because a single device holds so much information unrelated to the investigation.

In England and Wales, a review by the Information Commissioner's Office found that police practices in extracting data from the phones of victims and witnesses were inconsistent and often excessive, with people asked to hand over entire devices without a clear justification (Information Commissioner's Office [ICO], 2020). Legislation now sets conditions for extracting information from devices voluntarily provided by victims and witnesses, requiring that the extraction be necessary and proportionate (Police, Crime, Sentencing and Courts Act 2022). At the same time, failures to examine phones properly have contributed to miscarriages of justice; a parliamentary inquiry into disclosure found that problems in handling digital evidence had led to cases collapsing when relevant messages emerged late (House of Commons Justice Committee, 2018).

These developments point in the same direction: phone extraction should be targeted to what the investigation needs, but must be thorough enough to find evidence that points away from a suspect as well as towards one.

Volume and triage

A single phone can hold hundreds of thousands of messages and images, and a case may involve several phones. Quick and Choo (2014) found that the growth in the volume of digital evidence had outpaced forensic capacity, producing long backlogs, and Garfinkel (2010) predicted that examiners would increasingly need to select and prioritise rather than examine everything. Many forces now use triage: a quick, limited examination at an early stage to decide whether a device is relevant and which data to extract in full. Triage saves time but carries the risk that relevant material is overlooked, so its criteria should be recorded and, where a case depends on the device, followed by a fuller examination.

Casey (2011) stresses that the same principles of preservation and documentation apply to triage as to a full examination. Guidance on integrating forensic techniques into incident response makes the same point for organisations: decide in advance which data sources matter most, and how they will be collected without damaging them (Kent et al., 2006).

Tool reliability and human factors

Most mobile forensic work relies on commercial tools that automate extraction and decoding. These tools are essential, but they can misinterpret data, omit records they do not recognise, or present inferred information as fact. Horsman (2019) reviewed concerns about the testing of digital forensic tools and argued that validation is often inadequate given how quickly devices and apps change. Examiners should verify important findings by examining the underlying data directly or with a second tool, and should know which version of a tool produced each result.

Human factors matter too. Sunde and Dror (2019) showed that digital forensic examiners, like other forensic practitioners, are vulnerable to cognitive biases, such as being influenced by information about the suspect or by the expectations of investigators. Their recommendations include limiting examiners' exposure to irrelevant case information, documenting the reasoning behind interpretations, and using peer review.

Good practice

Taken together, guidance and research suggest a set of practices for mobile device examinations:

- **Isolate and preserve** the device promptly, keep it powered and record its condition.
- **Obtain proper authority** and keep the extraction within its scope, including any limits on dates, apps or categories of data.
- **Choose the least intrusive method** that will obtain the data needed, and record the reason.
- **Use validated tools**, record versions and settings, and verify key findings manually or with a second tool.
- **Interpret with care**, testing what timestamps, locations and other records actually mean on that device and app version.
- **Look for exculpatory evidence**, not only material that supports the investigation's hypothesis.
- **Protect the data**, which will include highly personal information about the owner and many other people, and delete what is not needed when the case ends.
- **Report clearly**, explaining in plain language what was extracted, how, what it shows and what its limitations are.

A short example

The following hypothetical case illustrates these practices. A company investigates whether an employee shared confidential tender documents with a competitor using a company phone. The employee consents to an examination limited to the company phone and a defined three-month period. The examiner isolates the phone, performs a file system extraction with a validated tool, and records the tool version and the hash value of the extraction file. The tool shows that a document with the tender's file name was sent through a messaging app to an unknown number. Before reporting, the examiner checks the app's database directly and finds that the message was never delivered: it remained in the outbox because the phone was offline. Further analysis shows that the employee deleted it the next morning. Location records

and the employee's calendar confirm that they were travelling without network access at the time. The report explains the finding, the evidence that the message did not leave the device, and the limits of the examination. The investigation then turns to other routes by which the information may have reached the competitor.

Conclusion

Mobile devices are among the richest sources of evidence in modern investigations, holding messages, media, locations and the data of countless apps. Extracting that evidence reliably requires prompt preservation, careful choice of extraction method, detailed knowledge of how apps store data, and awareness of the limits imposed by encryption and the cloud. Courts and regulators have recognised that phone searches are uniquely intrusive and require proper authority and proportionality. Tools are indispensable but fallible, and examiners are subject to bias. Investigators who preserve devices properly, extract only what they need, validate their tools, interpret data in context and report honestly can make phone evidence both powerful and trustworthy.

References

Abelson, H., Anderson, R., Bellovin, S. M., Benaloh, J., Blaze, M., Diffie, W., Gilmore, J., Green, M., Landau, S., Neumann, P. G., Rivest, R. L., Schiller, J. I., Schneier, B., Specter, M. A., & Weitzner, D. J. (2015). Keys under doormats: Mandating insecurity by requiring government access to all data and communications. *Journal of Cybersecurity, 1*(1), 69–79. https://doi.org/10.1093/cybsec/tyv009

Association of Chief Police Officers. (2012). *ACPO good practice guide for digital evidence* (Version 5).

Ayers, R., Brothers, S., & Jansen, W. (2014). *Guidelines on mobile device forensics* (NIST Special Publication 800-101 Revision 1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-101r1

Barmpatsalou, K., Damopoulos, D., Kambourakis, G., & Katos, V. (2013). A critical review of 7 years of mobile device forensics. *Digital Investigation, 10*(4), 323–349. https://doi.org/10.1016/j.diin.2013.10.003

Breeuwsma, M., de Jongh, M., Klaver, C., van der Knijff, R., & Roeloffs, M. (2007). Forensic data recovery from flash memory. *Small Scale Digital Device Forensics Journal, 1*(1), 1–17.

Carpenter v. United States, 585 U.S. 296 (2018).

Casey, E. (2011). *Digital evidence and computer crime: Forensic science, computers and the internet* (3rd ed.). Academic Press.

Cohn-Gordon, K., Cremers, C., Dowling, B., Garratt, L., & Stebila, D. (2020). A formal security analysis of the Signal messaging protocol. *Journal of Cryptology, 33*(4), 1914–1983. https://doi.org/10.1007/s00145-020-09360-1

Garfinkel, S. L. (2010). Digital forensics research: The next 10 years. *Digital Investigation, 7*(Suppl.), S64–S73. https://doi.org/10.1016/j.diin.2010.05.009

Hoog, A. (2011). *Android forensics: Investigation, analysis and mobile security for Google Android*. Syngress.

Horsman, G. (2019). Tool testing and reliability issues in the field of digital forensics. *Digital Investigation, 28*, 163–175. https://doi.org/10.1016/j.diin.2019.01.009

House of Commons Justice Committee. (2018). *Disclosure of evidence in criminal cases* (Eleventh Report of Session 2017–19, HC 859).

Information Commissioner's Office. (2020). *Mobile phone data extraction by police forces in England and Wales: Investigation report*.

In re Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203, No. ED 15-0451M (C.D. Cal. Feb. 16, 2016).

International Organization for Standardization. (2012). *Information technology — Security techniques — Guidelines for identification, collection, acquisition and preservation of digital evidence* (ISO/IEC 27037:2012).

Kent, K., Chevalier, S., Grance, T., & Dang, H. (2006). *Guide to integrating forensic techniques into incident response* (NIST Special Publication 800-86). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-86

Kerr, O. S. (2005). Searches and seizures in a digital world. *Harvard Law Review, 119*(2), 531–585.

Police, Crime, Sentencing and Courts Act 2022, c. 32 (UK).

Quick, D., & Choo, K.-K. R. (2014). Impacts of increasing volume of digital forensic data: A survey and future research challenges. *Digital Investigation, 11*(4), 273–294. https://doi.org/10.1016/j.diin.2014.09.002

Riley v. California, 573 U.S. 373 (2014).

Sunde, N., & Dror, I. E. (2019). Cognitive and human factors in digital forensics: Problems, challenges, and the way forward. *Digital Investigation, 29*, 101–108. https://doi.org/10.1016/j.diin.2019.03.011

Tamma, R., Skulkin, O., Mahalik, H., & Bommisetty, S. (2020). *Practical mobile forensics* (4th ed.). Packt.