The provenance of Apple Health data: A timeline of update history

By:
Luke Jennings, Matthew Sorell, Hugo G. Espinosa

From the proceedings of
The Digital Forensic Research Conference
DFRWS APAC 2024
Oct 22-24, 2024

Contents lists available at ScienceDirect

Forensic Science International: Digital Investigation

journal homepage: www.elsevier.com/locate/fsidi

DFRWS APAC 2024 - Selected Papers from the 4th Annual Digital Forensics Research Conference APAC

The provenance of Apple Health data: A timeline of update history

Luke Jennings[a,*], Matthew Sorell[a], Hugo G. Espinosa[b]

[a] *The University of Adelaide, Adelaide, 5005, South Australia, Australia*
[b] *Griffith University, Nathan, 4111, QLD, Australia*

ARTICLE INFO

Keywords:
Health
Apple
Smart watch
iPhone
Provenance
Timeline
Cyber hygiene
Pattern of life
Hardware
Firmware
Update history
Operating system
Device

ABSTRACT

Fitness tracking smart watches are becoming more prevalent in investigations and the need to understand and document their forensic potential and limitations is important for practitioners and researchers. Such fitness devices have undergone several hardware and software upgrades, changing the way they operate and evolving as more sophisticated pieces of technology. One example is the Apple Watch, working in conjunction with the Apple iPhone, to measure and record a vast amount of health information in the Apple Health database, *healthdb_secure.sqlite*. Over time, an end user will update their devices, but their health data, uniquely, carries over from one device to the next. In this paper, we investigate and analyse the hardware and software provenance of a real 5+ year Apple Health dataset to determine changes, patterns and anomalies over time. This provenance investigation provides insights in the form of (1) a timeline, representing the dataset's history of device and firmware updates that can be used in the context of investigation validation, (2) anomaly detection and, (3) insights into cyber hygiene. Analysis of the non-health data recorded in the health database arguably provides just as much insightful information as the health data itself.

1. Introduction

> ”*While this is essential for digital forensics to align with other forensic science fields, without this formalisation being complemented by peer-reviewed technical work including techniques that allow data to be extracted from data sources, and an understanding of artefacts that allow the interpretation of this data in the context of investigating crime, the technical capabilities within the field could formalise, but stagnate, risking missing important evidence as technology rapidly changes*” - Breitinger et al. (2024).

Devices may retain synchronisation artefacts of other devices that are out of reach. Can those synchronisation artefacts be exploited to inform an investigator about these devices? Garfinkel (2010) and Luciano et al. (2018) discuss ongoing and future challenges in digital forensics. The volume of digital data, and that it comes from disparate sources, cloud systems, and a multitude of devices is a significant challenge in digital forensics. The overwhelming volume of devices, the emergence of new devices, and updates to device firmware can slow down investigations. Tools can be deployed to assist investigators with the triage process to identify where to focus limited resources.

BankMyCell (2023) estimate that Apple Watches are one of the most popular brands of consumer wearables at a market share of 26 %. Apple Watches are already appearing in high-profile cases such as the Nilsson murder Opie (2019) and the Ladenburger murder BBC (2018). Edwards (2016) discusses the potential artefacts that may be useful in Apple Health, van Zandwijk and Boztas (2019) validate the accuracy of Apple Health step counts as digital evidence, and Jennings et al. (2023) interprets the time and location data in the context of workout activities. Our paper demonstrates that Apple Health data can be used to assist investigations in the triage stage as a validation, intelligence and analysis tool. The effect of firmware updates on Apple Health data is also discussed.

Carrier et al. (2003) defines the purpose of digital forensics tools to “include the acquisition of data from a source, analysis of the data and extraction of evidence, and preservation and presentation of the evidence”.

He quotes the agreed definition of Digital Forensic Science from DFRWS Workshop I Palmer et al. (2001, Fig. 5) which addresses eight specific areas as “The use of scientifically derived and proven methods towards the preservation, collection, *validation*, *identification*, *analysis*, *interpretation*, documentation and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal, or helping to

[*] Corresponding author.
E-mail address: luke.jennings@adelaide.edu.au (L. Jennings).

https://doi.org/10.1016/j.fsidi.2024.301804

Available online 18 October 2024

anticipate unauthorized actions shown to be disruptive to planned operations."

Addressing the provenance of Apple Health data in this paper builds on Edwards (2016), van Zandwijk and Boztas (2019) and Jennings et al. (2023) noted above. We address the issues of validation, identification, analysis and interpretation, as noted in Carrier's seminal paper, as being intrinsic requirements of digital forensic science as a practice.

Casey et al. (2009) defines three levels of forensic examination as.

1. Survey/Triage forensic examination;
2. Preliminary forensic examination; and
3. In-Depth forensic examination.

Overill et al. (2013) expands on Casey's three levels of examination, and suggests that Survey/Triage and Preliminary forensic examination form a feedback loop. That is; an initial preliminary examination on a device can identify more key devices on which to perform Survey/Triage.

Hargreaves and Marshall (2019) take the concepts discussed by Carrier, Casey and Overill and presents an overall approach for inferring the existence of, and partial content of other devices, dubbed SyncTriage. This approach involves the exploitation of synchronisation artefacts to infer the content of devices that have not been accessed, or not have been identified and seized.

In this paper we demonstrate that the provenance of Apple Health data can be used as a synchronisation artefact to construct a timeline of events to infer the existence of other devices linked through a common Apple ID. The timeline can be exploited to validate the metadata of other artefacts, such as photographs, and as a tool to analyse a person's pattern of life, due to the nature of health data associating physical activity data with timestamps.

The opening quote above by Breitinger et al. (2024) states that "complementing research … without an understanding of the interpretation of this data in the context of investigating crime", the field of digital forensics risks becoming "stagnant". This paper interprets the provenance data of a real Apple Health dataset in the context of scenarios or case studies to reveal its potential.

This paper makes the following contributions.

1. Demonstrates the specific synchronisation artefacts, and analyses them in the form of scenarios, or case studies, revealing how such artefacts can be used in a real investigation as a validation, intelligence or analysis tool (Section 4);
2. Demonstrates how timeline constrained synchronisation artefacts from a longitudinal real dataset can be used to analyse and interpret a person's pattern of life (Cyber Hygiene as an example) (Section 4.4);
3. Makes available a real dataset for the research community and practitioners (Upon request to authors)
4. Identifies and discusses potential artefacts and anomalies, showing how device behaviour can change with firmware updates (Section 3.5);
5. Presents the analysis of synchronisation artefacts in a temporal domain in the context of a timeline (Tables 5 and 6);
6. Provides an informative schematic overview of a section of the Apple Health database (Figs. 1 and 2).

The remainder of the paper is structured as follows: section 2 provides a summary of related work, section 3 describes the method for the research, section 4 presents the results, and section 5 provides the conclusion and future work.

2. Related work

This section focuses on the existing work related to iOS databases, synchronisation artefacts, timelining and pattern of life in digital forensics.

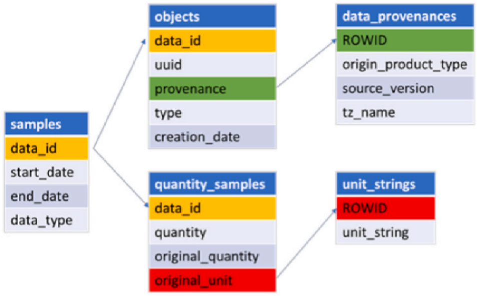


Fig. 1. Database structure for data provenances + timestamps.

2.1. iOS databases/artefacts

Morrissey and Campbell (2011) and Zdziarski (2013) discuss iOS forensics including overviews of the various databases found within iOS and potential artefacts. Edwards (2016) discusses specific iOS databases and artefacts that can be exploited to analyse and interpret a person's pattern of life. Historically, the more artefact-rich databases in iOS digital forensics can be considered to be the Cache, PowerLog, SMS, Call History, Calendar and Photos databases.

Stanković et al. (2021) explores UAV forensics, relying on location data that is stored in the iOS Cache.sqlite database noted above. They note that a specific challenge with the Cache database is that it only holds data up to one week. This limitation can be noticed with other prominent iOS databases with varying lengths of data retention. Jennings et al. (2023) explores the location data of the Apple Health database, discovering that the health database retains its full data in the order of years. Additionally, the health data is tied to the Apple ID and not bound to a particular device and survives across device upgrades.

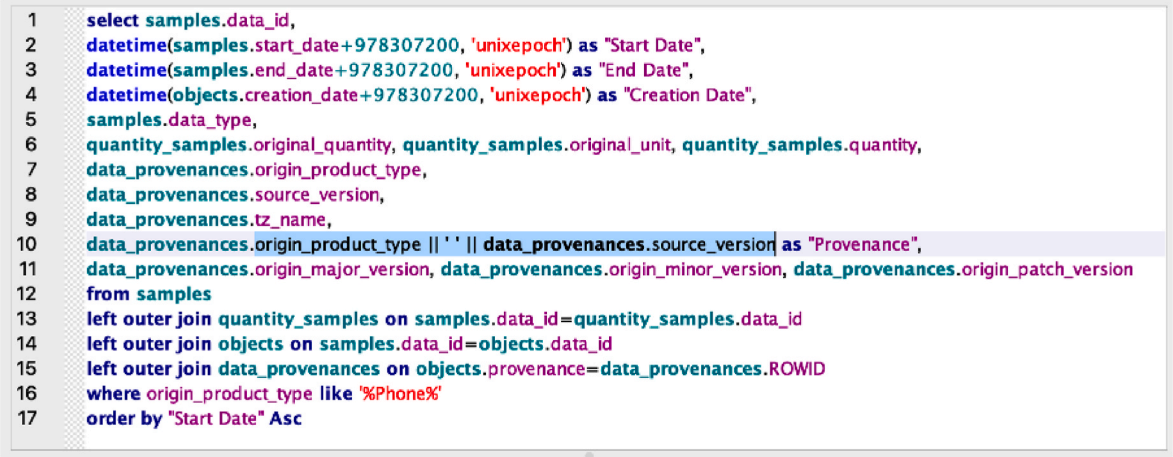
2.2. Synchronisation artefacts

Friedman et al. (2012) explores the synchronisation between iOS devices and iCloud. Friedman found that there were artefacts to show whether iCloud was enabled on devices, but found little evidence showing whether two devices were connected to each other through iCloud. Boucher and Le-Khac (2018) suggest that in modern applications, developer's attempts to integrate a person's data across device upgrades seamlessly, present a challenge to forensic investigators to determine the original source of artefacts. Boucher proposes a framework to determine if artefacts on a device are local or synced elsewhere. Hargreaves and Marshall (2019), however, describe how synchronisation artefacts can instead be exploited as a generalisable approach to infer content of a device from another. Hargreaves also presents an overview of other work focusing on synchronisation related artefacts.

2.3. Timelines in digital forensics

Timelines are vital in digital forensic investigation. Årnes (2017) states that "A timeline of a chain of events can include physical events, as well as digital events …. The activity revealed on the cell phone, such as call logs, can provide relevant evidence related to the timeline of the crime.".

There are research papers dedicated to timeline construction processes and tools in digital forensics. For example, Olsson and Boldt (2009) created a tool called Cyber Forensic Time Lab, to assist investigators through the visualisation of evidence by indexing with time variables and plotting on a timeline. Hargreaves and Patterson (2012) present a framework to produce high-level, human-readable events based on one or more low-level events. Henseler and Hyde (2019) combine timelines with link analysis through a graph database and



```sql
select samples.data_id,
datetime(samples.start_date+978307200, 'unixepoch') as "Start Date",
datetime(samples.end_date+978307200, 'unixepoch') as "End Date",
datetime(objects.creation_date+978307200, 'unixepoch') as "Creation Date",
samples.data_type,
quantity_samples.original_quantity, quantity_samples.original_unit, quantity_samples.quantity,
data_provenances.origin_product_type,
data_provenances.source_version,
data_provenances.tz_name,
data_provenances.origin_product_type || ' ' || data_provenances.source_version as "Provenance",
data_provenances.origin_major_version, data_provenances.origin_minor_version, data_provenances.origin_patch_version
from samples
left outer join quantity_samples on samples.data_id=quantity_samples.data_id
left outer join objects on samples.data_id=objects.data_id
left outer join data_provenances on objects.provenance=data_provenances.ROWID
where origin_product_type like '%Phone%'
order by "Start Date" Asc
```

Fig. 2. Query to extract initial timeline.

graph query language to assist with artefact analysis. A discussion on commercial tool timeline analysis is also presented. Bhandari and Jusas (2020) introduce an abstraction-based methodology for the reconstruction of timelines efficiently for investigators. These papers focus on processes, methods, and tools to improve the timeline reconstruction in investigations.

This paper focuses on utilising a timeline reconstructed with Apple Health provenance artefacts in the context of scenarios or case studies.

2.4. *Pattern of life*

Edwards (2016), as noted above, explores various iOS databases that have the potential to model pattern of life, including Apple Health data. For example, the battery life can be exploited to analyse application usage. Edwards also discusses typical data retention time for the databases discussed.

In this paper, we use cyber hygiene as an example of analysing pattern of life by comparing the timeline reconstructed from the Apple Health provenance artefacts with official Apple update release documentation.

Maennel et al. (2018) define cyber hygiene as “Cyber hygiene is a set of practices aiming to protect from negative impact to the assets and human life from cyber security related risks.”. Maennel describes “good” online hygiene practices include updating applications, software and operating systems within 48 h of patches becoming available.

It must be made clear that with the term ”cyber hygiene” used throughout this paper, we are not inferring an attitude or state of mind of the user when they update. We are only interested in the behaviour of the user, that is; when did the update occur? It can be determined that if patches to iOS and WatchOS are made within 48 h, this either constitutes “good” or “bad” cyber hygiene behaviour.

3. Method

3.1. *Background to the method*

The Apple health database *healthdb_secure.sqlite* is an SQLite database that incorporates tables of data, and also describes the overall structure through what is known as the Schema. The Schema describes which tables are interconnected through which indices and may also include pre-programmed queries as Views and Triggers. The detailed structure of an SQLite file is outside the scope of this paper but for the interested reader, is discussed in Tutorial (2015). The specific understanding of the tables and codes in this database are a consequence of previous work Edwards (2016) Jennings et al. (2023), with both systematic and ad-hoc experimental validation of the database by the authors over time. Our analysis does not consider the entire database structure and is limited only to those tables necessary for the timeline creation and device provenance analysis.

The health dataset under investigation is provided by one of the authors. This dataset was captured between 2017 and 2022, capturing everyday health and provenance data for one user with various iPhones and Apple Watches. This dataset can be made available to interest readers on request to the authors, subject to usage policy.

3.2. *Ground truth and validation*

While the use of a real 5+ year health dataset is valuable, The ground truth must be established before it can be analysed and the results found trustworthy. The devices and timestamps must first be confirmed separately. Previous research by the authors in van Zandwijk and Boztas (2019) investigate the validity of step counts as forensic artefacts in digital forensics. The calibration of timestamps was part of their experimental discussion. We have also validated the timestamps with our own experiments, the topic of a separate publication in preparation, and found that our results are consistent with van Zandwijk and Boztas (2019). Specifically in regards to timestamps, we found that.

1. Start times have a mean lag time of approximately 7–8 s.
2. End times have a mean lag time of approximately 2–3 s.

We consider the timestamps for step counts to be valid and reliable for the purpose of creating a device timeline in this paper. Granularity of the timestamps for timeline creation are limited to the official update release documentation supplied by Apple, which only give a date, and not by the timestamps of the health data. If timestamps were given by Apple, then the accuracy of the step count time stamps become increasingly important.

In order to evaluate the accuracy of the device provenance timeline, the actual devices owned by the dataset owner must be documented. The primary method of confirmation for the ground truth was the inspection of receipts for the device and date of purchase. In the absence of a receipt, email documentation (such as support tickets), other service bills or interview questioning were used to confirm the dates of the devices. The established device ground truth is summarised in Table 1.

Through an interview with the dataset owner, it was confirmed that automatic updates on the iPhone and Apple Watch were switched off. In some aspects, automatic updates can be a measure of cyber hygiene, but these are an inference of an attitude or state of mind, which is not what we are attempting. Since we are only interested in the behaviour, that is the binary nature of was the device updated or not and consequently when, automatic updates are not considered to be a requirement for

Table 1
Ground truth of device ownership.

Device	Receipt Date
iPhone 6S Plus Gold 128 GB	12 November 2015
Apple Watch S1 38 MM	May 22 2017
iPhone 7 Plus Gold 256 GB	24 June 2017
Apple Watch S3 42	1 March 2018
iPhone XS Max	16 June 2019
Apple Watch S4 44	6 September 2019
Apple Watch S7	20 January 2022
iPhone 13 Pro Max 512 GB	1 June 2022

analysis.

3.3. Extraction

The information from the health database can be extracted in several ways from the iPhone, discussed by the authors in Jennings et al. (2023). The most direct method is a direct export from the application interface. This process supplies the extractor with export.xml files detailing the user's various health and activity information and workout route.GPX files. However, this method presents some limitations. With the direct export Apple Health takes a shortcut, exporting all data in the local time zone at the place and date of the export. This means that data recorded, for example, in winter, will be exported during summer in daylight savings time; and data recorded in Europe but exported in Singapore will be exported in the Singapore time zone (UTC+8). The export time zone is included in the exported times, but the source time zone is not.

Apple Health information can also be retrieved from iCloud Jennings et al. (2023) or using one of the many dedicated forensic tools, for example MSAB's XAMN. The raw SQLite database can also be extracted from the encrypted backup of an iPhone Reincubate (2024), and then the tables can be constructed with SQL queries, which is the method utilised in this study. The file pathway for accessing the database is*/private/var/mobile/Library/Health.*

The SQLite database, *healthdb_secure.sqlite*, contains various tables which contain health, activity, workout, and as we will demonstrate, device provenance information. Data within SQLite tables can be abstracted away from related fields and must be re-joined through SQLite queries using the join command to link tables that have a shared ID field Tutorial (2015). For this study we work as directly as possible with the database using an SQLite database browser (DB Browser for SQLite V3.11.2) and SQL commands similar to tools such as those provided in the iLEAPP suite Brignoni (2024) recognizing that browsers have their own limitations. By working directly with the database we maintain as complete oversight of the database structure and schema as possible rather than relying on a third party's analysis which we would then need to validate separately.

3.4. Apple Health SQLite database structure

In the Apple Health database, the **data_provenances** table contains the device provenance information for the Apple account. The columns of particular interest to us in the **data_provenances** table are.

- *ROWID*
- *origin_product_type*
- *source_version*
- *tz_name*
- *origin_major_version*
- *origin_minor_version*
- *origin_patch_version*

The hardware provenance data that the health information is recorded by is stored in the *origin_product_type* column. The *local_product_type* column is the device that the *origin_product_type* is paired with, for example if the origin device is an Apple Watch, the local device will be the paired iPhone. When there is more than one device used by the person, the provenance is identified separately. The firmware provenance is stored in the *source_version* and *origin_major_version*, *origin_minor_version* and *origin_patch_version* columns. An example is demonstrated in Table 2 which illustrates how the provenance information is stored in the database.

For the rest of this paper, we concatenate *origin_major_version*, *origin_minor_version* and *origin_patch_version* into a single column and refer to it as "origin version". For the device provenance in Tables 3–8 below, we combine origin product type and source version into a single column, and do the same for origin product type and origin version, for analysis. It is important to note that the firmware version stored in *source_version* may not always be identical to the origin version stored in the other columns. These differences and anomalies are explored in section 3.5. *Tz_name* stores the timezone that the health activity was performed in. The *ROWID* column is what connects this table to other tables in the database through queries. The structure is shown in Fig. 1, and a query to join them is shown in Fig. 2.

The **data_provenances** table does not store timestamps for health activity. Instead, these are stored in the **samples** table. In addition to that, the health measurements themselves (i.e. number of steps, heart rate, etc.) are stored in the **quantity_samples** table. These tables are shown in Fig. 1. The *data_types* column in the **samples** table describe what activity is being performed, such as steps or distance, as an integer Edwards (2016) Jennings et al. (2023).

3.5. Data types and provenance considerations

The data type matters in the construction of the timeline and in this paper, we are only considering step counts for the creation of the timeline. This section will explore the other data types and justify why they are eliminated from our provenance timeline analysis, and why we believe step counts to be most suitable. When considering health data, the most typical and prominent data types are considered to be.

- Heart Rate;
- Basal and Active Energy Burnt;
- Distance;
- Step Count;
- Other.

On first glance, it may appear to the reader that more data types would provide more data points to provide a more insightful timeline but this is not the case. In order to provide the most accurate version of the provenance timeline, we need the data types that we include to sufficiently cover these three specific categories.

1. Time Granularity;
2. Availability;
3. Reliability.

Time granularity in this context refers to the precision of the timestamps. In Apple Health, the time stamps record the date, hours, minutes and seconds. For time granularity, step counts and distances can give precision in seconds. However, heart rate (BPM) and energy burnt (cal) are summary measurements, with a time granularity of 1 min.

Availability in this context refers to what devices can record which data types. Heart rate and energy burnt can only be recorded on an Apple Watch, while distance and step count can be recorded on both iPhone and Apple Watch.

For these time granularity and availability reasons, we exclude the heart rate and energy burnt from the provenance timeline. This leaves only step counts and distances, however, this leads into the third category of reliability.

It is generally expected that distance, along with flights of stairs

Table 2
Device Provenance in *healthdb_secure.sqlite*

origin_product_type	origin_major_version	origin_minor_version	origin_patch_version	source_version
iPhone9,4	11	0	2	11.0.2

Table 3
Distance and iOS 12 anomalies in the database.

Start Date	End Date	Creation Date	data_type	Provenance - Origin Version	Provenance - Source Version
5/5/2018 23:31	5/5/2018 23:40	6/5/2018 0:22	7 (Steps)	iPhone9,4 11.3.1	iPhone9,4 11.3.1
27/5/2018 5:50	27/5/2018 5:56	27/5/2019 18:48	8 (Distance)	iPhone9,4 12.2.0	iPhone9,4 4.2.3
18/6/2018 15:55	18/6/2018 16:03	18/6/2018 16:39	8 (Distance)	iPhone9,4 11.4.0	iPhone9,4 11.4
20/7/2018 22:51	20/7/2018 22:57	20/7/2018 23:21	7 (Steps)	iPhone9,4 11.4.1	iPhone9,4 11.4.1
19/9/2018 9:13	19/9/2018 9:20	19/9/2018 10:14	8 (Distance)	iPhone9,4 12.0.0	iPhone9,4 12.0
22/9/2018 5:19	22/9/2018 5:24	23/9/2019 9:14	8 (Distance)	iPhone11,6 13.0.0	iPhone11,6 5.0
8/10/2018 8:42	12/10/2018 5:35	14/10/2018 20:10	8 (Distance)	iPhone9,4 12.0.1	iPhone9,4 12.0.1

Table 4
Origin version anomaly.

Data Type	Provenance - Origin Version	Provenance - Source Version	Date
7 (Steps)	iPhone8,2 10.3.0	iPhone8,2 10.3.1	29/04/2017
8 (Distance)	iPhone8,2 0.0.0	iPhone8,2 10.3.2	24/05/2017
8 (Distance)	iPhone9,4 0.0.0	iPhone9,4 10.3.2	24/06/2017

Table 5
Device timeline.

Date	Provenance - Source Version	Device Name
29/04/2017	iPhone8,2 10.3.1	iPhone 6s Plus
24/06/2017	iPhone9,4 10.3.2	iPhone 7 Plus
27/11/2017	Watch2,6 3.2	Watch Series 1
01/03/2018	Watch3,2 4.2	Watch Series 3
17/06/2019	iPhone11,6 12.3.1	iPhone Xs Max
12/07/2019	Watch4,4 5.2.1	Watch Series 4
20/01/2022	Watch6,9 8.1.1	Watch Series 7
01/06/2022	iPhone14,3 15.5	iPhone 13 Pro Max

Table 6
User's iPhone Timeline.

Provenance - Source Version	User Update	Apple Release Date	Days to Update
iPhone9,4 12.0	19/09/2018	17/09/2018	2
iPhone9,4 12.0.1	08/10/2018	08/10/2018	0
iPhone9,4 12.1	11/11/2018	30/10/2018	12
iPhone9,4 12.1.2	03/01/2019	17/12/2018	17
iPhone9,4 12.1.4	27/02/2019	07/02/2019	20
iPhone9,4 12.2	12/04/2019	25/03/2019	18
iPhone9,4 12.3.1	04/06/2019	24/05/2019	11
iPhone11,6 12.3.1	17/06/2019	24/05/2019	24
iPhone11,6 12.4	28/07/2019	22/07/2019	6
iPhone11,6 12.4.1	12/09/2019	26/08/2019	17
iPhone11,6 13.0	22/09/2019	19/09/2019	3
iPhone11,6 13.1.1	28/09/2019	27/09/2019	1

Table 7
Photo metadata summary.

Photo	Device	Metadata Timestamp	Provenance Timeline
1	iPhone XS Max	26 December 2020	17 June 2019
2	iPhone 7 Plus	26 May 2018	24 June 2017
3	iPhone 13 Pro Max	13 July 2022	1 June 2022
4	iPhone 6s Plus	3 May 2017	29 April 2017

Table 8
Watch outliers.

Provenance - Source Version	User Date	Apple Release Date	Days to Update
Watch2,6 3.2	27/11/2017	27/03/2017	245
Watch2,6 4.1	28/11/2017	31/10/2017	28

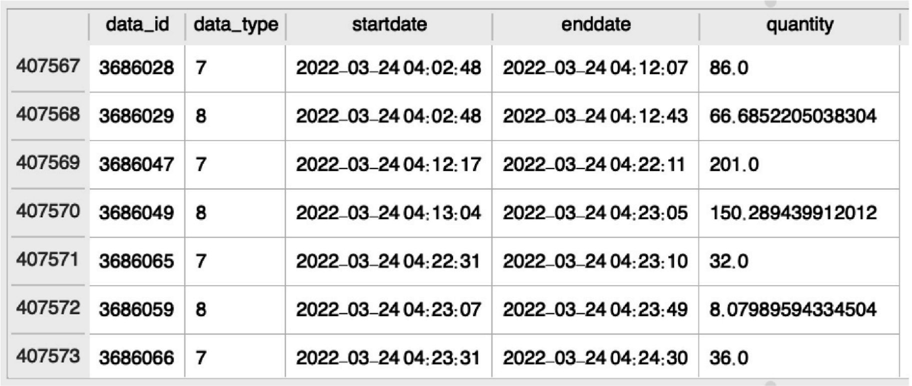
climbed, are derived from steps van Zandwijk et al. (2023). This means it is expected that they share the same timestamps. However, this is only true most of the time, and through our own case work, research, and experimentation have seen instances where this is does not hold true.

An example from the database where step counts (data type 7) and distances (data type 8) do not have matching time stamps is shown in Fig. 3.

It is demonstrated that *data_id* 3686028 (steps) and 3686029 (distance) have the same start time, but not the same end time. The distance measurement continues for 36 s after the step count interval had already ended. This raises the question of what is happening during this 36 s interval that the user is not stepping, but still moving.

The mismatched end date timestamps create a level of uncertainty around the actual activity interval of the distance timestamps. This is only one example where the supposedly derived distance is unaligned with the data type measurement it is based on, but we have seen many instances of these anomalies. For now this level of uncertainty is still within the accuracy required for the timeline, which is constrained by Apple's official release documentation, which is only a date.

However, this is also not the only issue around the accuracy of the timestamps regarding distance measurements. Besides the *start_date* and



	data_id	data_type	startdate	enddate	quantity
407567	3686028	7	2022-03-24 04:02:48	2022-03-24 04:12:07	86.0
407568	3686029	8	2022-03-24 04:02:48	2022-03-24 04:12:43	66.6852205038304
407569	3686047	7	2022-03-24 04:12:17	2022-03-24 04:22:11	201.0
407570	3686049	8	2022-03-24 04:13:04	2022-03-24 04:23:05	150.289439912012
407571	3686065	7	2022-03-24 04:22:31	2022-03-24 04:23:10	32.0
407572	3686059	8	2022-03-24 04:23:07	2022-03-24 04:23:49	8.07989594334504
407573	3686066	7	2022-03-24 04:23:31	2022-03-24 04:24:30	36.0

Fig. 3. Distance times unaligned.

end_date columns in the **samples** table, there is the *creation_date* column in the **objects** table, which is assumed to be the timestamp that the activity is logged in the database. Consider Table 3, which is a snapshot of the timeline created when using both step counts and distances.

With iOS version 12, there are instances where the distances recorded have a significant delay in being logged in the database (*creation_date* timestamp). In Table 3, we can see for some of the distance records (rows 2 and 6), this delay was almost a year after the actual activity interval (*start_date* and *end_date* timestamps). When it is finally logged in the database it records the provenance information at the time it is logged and not the provenance for when the activity was actually performed.

The timestamps are not the only unreliability or uncertainty around distance measurements in the database. There are also inconsistent and missing provenance records for distance measurements. From section 3.4 above, it was shown that the firmware provenance is recorded in the *source_version* column, as well as the *origin_version* columns. In Table 3 it is the source version that is incorrect and shows a much lower version number than what is expected to appear in the database, i.e., “iPhone9,4 4.2.3” in row 2 and “iPhone11,6 5.0” in row 6. It can be observed that these source version provenance anomalies can be linked to the delayed *creation_date* timestamps in the distance records.

In Tables 4 and it is demonstrated that there is another provenance mismatch between origin version and source version, but it is origin version that is incorrect with a version 0.0.0 appearing instead.

Due to the uncertainties of the timestamps for distance measurements, and the inconsistencies around provenance records, it has been determined that for the purpose of creating an accurate provenance timeline that distance records are unreliable and should not be used.This does not mean that distances are unreliable in other analyses, they are just not fit for the purpose of this paper in the creation of the provenance timeline. The use of distances, requires further experimentation and validation. The distances need to be evaluated in terms of accuracy versus reliability, and the sequencing of their orders (*data_id*) within the database evaluated.

This leaves step counts as the only prominent data type that satisfies the three categories needed for a reliable timeline, that is; the step counts have the best time granularity, they are available on both types of devices, and their timestamps and provenance records are reliable. While we have only discussed the typical or prominent data types of interest recorded in fitness devices (steps, distance, heart rate and energy burned), it is appropriate to raise the question about less interesting or atypical data types of a lesser focus or lesser interest.

In a yet to be published paper, the author undertook a process of identification, attribution, evaluation and correlation of the (as of 2022, iOS 16.0) 43 data types within the Apple Health database. Of these 43 data types, the prominent data types discussed above provide the most context to a user’s actions and activity, and are sufficient for the purpose of this paper.

There is, however, a possibility that some unexpected data types may address outlier and edge case scenarios in the provenance timeline, such as when the user is using neither device. An example of this is discussed in section 4.5 limitations. This data type selection process revealed unexpected anomalies while allowing for an enhanced and deeper analysis on the timeline. Through this process it became clear that the creation of an accurate and reliable timeline was possible, through the use of step counts. This in itself is an unexpected contribution to the deeper understanding of the potential of Apple Health data in digital forensics.

3.6. Timeline creation

The timestamps for the provenance of the user account are stored in the **samples** (as *start_date* and *end_date*) and **objects** (as *creation_date*) tables. The *start_date* and *end_date* are the start and end of the health activity measurement. The *creation_date* is the timestamp for when the activity is logged in the database. The timestamps are stored as Apple Cocoa Core UTC. Distinguishing these three dates is important because, as with the example described above, a health activity measurement could start one day, finish the next day, and then log in the database potentially hours after resulting in a delayed timeline by 24 h or more.

The initial creation of the timeline is performed by the query in Fig. 2. In this query we sort the results by the *start_date* in ascending order to get the first instances of a particular hardware and firmware combination. For ease of analysis and simplicity we created the timelines separately for iPhone and Apple Watch using the “where” clause in the SQLite query. We then filter by unique instances of hardware + firmware to create the timeline showing the first instances in the database for specific devices and OS versions.

The identification and analysis of the anomalies in section 3.5 above is not an indication that the health data is unreliable, but rather a snapshot of a complex and evolving database. To a newer and less experienced researcher or practitioner analysing a database, these anomalies serve as potential false flags that might be misinterpreted. By shining the spotlight on these intricacies we can gain a better understanding of the potential and limitations of the Apple Health database.

4. Results and discussion

4.1. User timeline

By following the process described in section 3 we have successfully extracted a provenance timeline of the user using only their Apple Health database. For the user’s iPhones, there are 65 unique timestamped health database entries for devices and corresponding firmware. For the user’s Watches, there are 53 database entries. These entries illustrate when the user either purchases (more specifically, sets up) a new device or updates an existing device. The timeline for new device acquisition for the user is shown in Table 5 for the time period between 29 April 2017 and 1 June 2022. The second column is how the database describes the device names which we translate using Every-Mac.com (2024) into the more common market names into column three.

It is important to make the distinction between purchase and set up as there are scenarios where these can be quite different. One example is the purchase of a phone or watch as a gift and leaving it boxed up for a period of time. Another example is in the case of initial factory setup and configuration which acts as an indicator for Apple's inventory and logistics management. Both of these examples can impact the potential delay between user patch download dates. In the following section we compare the user's provenance timeline with the official release dates as posted by Apple Apple (2024).

4.2. *Apple documentation for OS releases*

Table 6 demonstrates a snapshot of the user's iPhone timeline including regular firmware updates, a device upgrade and the days it took the user to update compared to Apple's release dates Apple (2024).

As demonstrated in Table 6 there is sufficient provenance data in the Apple Health database to establish a detailed and long term timeline. With this timeline there are several insights that can be made to help investigators.

- It can now be determined which devices are connected to a particular suspect's Apple account.
- No serial number – if the health data can show that the owner has used a previous model of iPhone, this may also assist the analysis of attribution of network records which typically include the International Mobile Equipment Identifier (IMEI).
- If law enforcement seizes a latest generation iPhone which does not match the network records for neither the IMEI or the International Mobile Subscriber Identity (IMSI) of a suspect, the health data might identify the previous iPhone make and model which does align with communication records for the previous device.
- Camera roll metadata might attribute a photograph taken from a different device other than that seized by law enforcement. Analysis of the health database can link the provenance of the older make and model phone to the suspect's Apple account and validate the photo's provenance.

Such a rich and extensive provenance history is invaluable as a resource for investigators and researchers while not being particularly difficult to extract. As a consequence of having a detailed timeline available we can establish and measure the user's cyber hygiene; that is what their update pattern behaviour looks like. Following Table 6 we can calculate the number of days to update for every single iPhone, Watch and firmware version over the user's 5+ years of health data activity.

4.3. *Device validation and photograph attribution*

This section explores examples of how a device provenance timeline synthesised from Apple health data can be utilised by investigators in the context of scenarios. The scenarios that will be explored are Validation; Attribution; and Intelligence/Discovery.

4.3.1. *Validation*

Photographs extracted from the camera roll of a suspect's iPhone contain Exif metadata in which the authenticity/validity may come under questioning during an investigation or trial. This metadata can be validated using Apple Health data.

The owner of this Apple Health dataset's current phone is the iPhone 13 Pro Max. The four photographs in Fig. 4 were extracted from that device's camera roll. Inspecting the Exif metadata of the four photographs reveals details about the digitisation date, geo-location and device make and model. It was confirmed through interview questioning with the health dataset owner that these photos were indeed taken by the user on their device.

Table 7 summarises the device makes and models identified for the

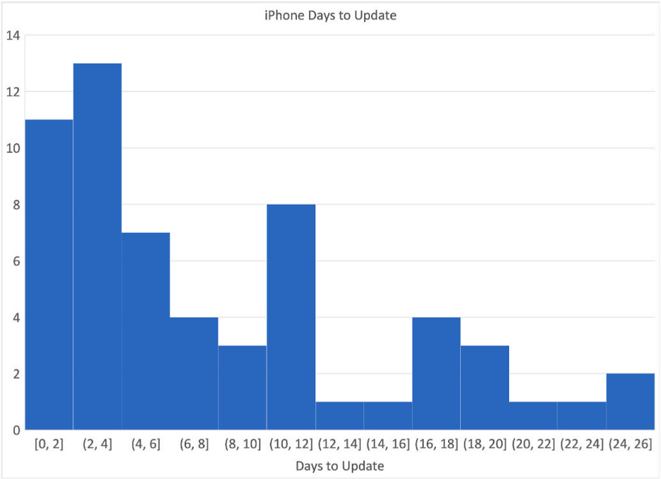


Fig. 5. iPhone update pattern histogram.

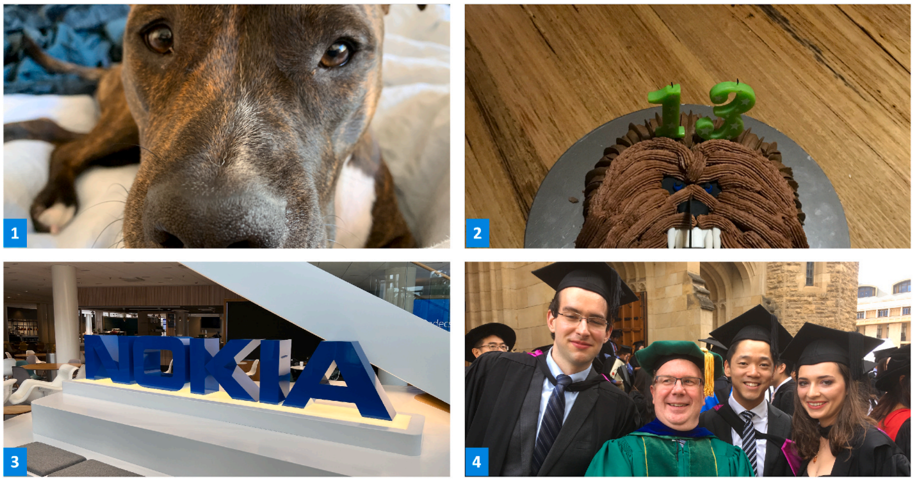


Fig. 4. Extracted photos.

photographs extracted, being an iPhone 6s Plus, iPhone 7 Plus, iPhone XS Max and iPhone 13 Pro Max. The dates that the photos were taken are presented in the third column, and the timeline dates generated from the health data provenance are presented in the fourth column. It can be observed that for each of the photos, both the date that the photo was taken, and the device used to take the photo are consistent with provenance information extracted from the user's Apple Health data, hence validating the metadata and timestamps observed.

4.3.2. Attribution

In an investigation multiple phones or smart watches may be present, and a challenge in digital forensics is attributing the devices discovered to a particular person, suspect or victim. Consider the scenario where at a crime scene an iPhone XS Max, iPhone 6s Plus, iPhone 7 Plus, and an iPhone 13 Pro Max were seized. The question that investigators will then ask is who did these devices belong to. By analysing the provenance timeline generated from this Apple Health dataset, it can be identified that these four devices belong to the same Apple account.

4.3.3. Intelligence and discovery

This scenario considers the initial triage stage discussed in section 1. Consider that a single iPhone 13 Pro Max was seized by investigators at a crime scene, and a preliminary forensic examination is performed. The Apple Health data can reveal that there are three other iPhone devices associated with the person of interest that have been in considerable use in recent years. This can prompt investigators to expand their search and investigation to other devices, which if the timeline of the crime aligns with the period of use identified in the provenance timeline, may provide even more relevant forensic artefacts to the investigation.

4.4. Pattern of life: cyber hygiene modelling and analysis

While the primary contribution of this timeline is to serve as a tool for validation, the level of detail found in the timeline presents an extended, additional use case: modelling of pattern of life. In this section, we demonstrate how the generated provenance timeline derived from Apple health data can be used to measure a user's pattern of life in the context of cyber hygiene as one example.

To measure and analyse the cyber hygiene of the user through their health data provenance there are some outliers in the data that need to be discussed. These outliers are present when the user upgrades their device and the time at which they do so can have a substantial effect on the following analysis.

Consider Table 8 which shows the first provenance instance of the user's Apple Watch in the database in 2017. The entry suggests that the user took 245 days to update their first Apple Watch. What is more likely is that the device came pre-installed with an OS that was out of date by the time of acquisition. Therefore, it is unfair to include this entry in the cyber hygiene measure. Consequently, the following entry is also unfair to include in the analysis. The user updated their device 24 h later but because of the time of the device purchase the days to update is larger than if they had just purchased the device earlier. Hence for the following analyses we produce a histogram for both the user's iPhones and Apple Watches to demonstrate their update behaviour over the duration of the whole health dataset and we remove the outliers introduced by new device acquisition as described above.

4.4.1. iPhone histogram

After removing the outliers we have 59 remaining iPhone provenance data points, presented in Fig. 5. The user most commonly updates their device within 0–4 days, accounting for 40.68 % of the total device updates.

Interestingly, there is a spike in the 10–12 day range which accounts for 13.56 % of total device updates.

On average the user updates their iPhones within 8–9 days with a standard deviation of approximately 7 days. The quickest time the user updated their phone was 0 days and the longest was 26 days.

4.4.2. Watch histogram

After removing the outliers we have 45 remaining Apple Watch provenance data points presented in Fig. 6. The user most commonly updates their devices within 1–3 days accounting for 60 % of the total device updates.

On average, the user updates their Watches in 4 days with a standard deviation of 3 days. The quickest time the user updated their watch was 1 day and the longest was 17 days.

4.4.3. Insights

In the context of cyber safety this information is not typically obtainable for potential threat actors. However, with the rising popularity of these fitness trackers the likelihood of people sharing their workouts and fitness information with friends, family or even online may also rise. Even in the basic export method discussed above in section 3, and referenced in Jennings et al. (2023), there would be sufficient device provenance information available in those records. A potential threat actor may use this sort of process to profile their potential target, discovering that they take a certain amount of days to update their device (and with it apply the security patches), and leverage certain vulnerabilities in the devices and to focus their attack on one particular device or the other.

Do you have to update your iPhone to update your Apple Watch? No.

4.5. Limitations

The timeline analysis relies heavily on accurate release date information from Apple and that we have to assume that it is correct. Additionally, it is unknown if iOS update times are region and timezone dependent. It is plausible that iOS updates have staggered release times. With the current precision of the timeline, this could shift the ”days to update” in the pattern of life analysis by 24 h or more.

Another assumption is that the user has set up their Apple Health app. While a less active user would generate fewer data entries, a timeline could still be constructed. In this case, the more data available the more precise the constructed timeline will be.

In our analysis we limited our investigation to step counts, however, there is potential in using other data types to further refine the constructed timeline. One such data type is Apple Stand Hours. This data type records if the user has stood and moved for at least a minute within each hour of the day. A fringe case scenario in which this could be

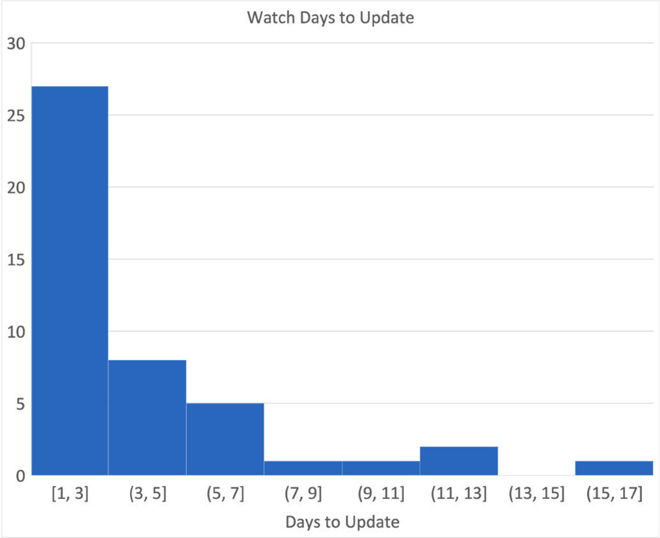


Fig. 6. Watch update pattern histogram.

applied is one of a lazy Sunday, where the user has their watch on the bedside table, selects update, and leaves it there the rest of the day and overnight before wearing it again on Monday morning. Apple Stand Hours would record that the user was not standing throughout that day and capture the change in firmware provenance whereas focusing only on steps counts would miss this.

The timestamps in the Apple Health database go down into hours, minutes and seconds whereas Apple's security release documentation EveryMac.com (2024) only gives the date. If there were timestamps associated with Apple's documentation we could potentially refine the precision of the timeline into hours, minutes and seconds.

5. Conclusion

Using only the Apple Health database, *healthdb_secure.sqlite*, we have extracted the device provenance information (hardware and firmware) for one user's dataset over 5+ years. This database provides provenance information associated with every single health activity entry in the database allowing us to create a timeline of every major hardware and software upgrade for the user. This timeline is useful in validating other forensic records such as photos and call records when older devices are present in place of the currently seized device. The timeline itself provides several key insights into the user's cyber safe practices; that is if they update quickly or not. For the user's iPhone they typically (40.68 %) update within 0–4 days. For the user's Apple Watch they typically (60 %) update within 1–3 days. This information is potentially dangerous in the hands of a threat actor who can use it to tailor a targeted attack against a person. The analysis of device provenance from such an unexpected source provides many key insights which are invaluable to investigators and future researchers alike.

5.1. Future work

There are many aspects that could be addressed in future research to improve the depth of understanding of device provenance. First, the consideration of other data types besides steps can provide more context into fringe case scenarios in which the device records the lack of activity. Second, it is possible to expand the investigation into other datasets such as those by FitBit and Garmin, or even those same devices but incorporated into the Apple Health App. Third, is the consideration of other aspects of Pattern of Life in which a provenance timeline could help explore. The current significant challenge is finding datasets that are recorded over several years by relatively active users during ordinary day to day activities, and even more so those that are willing to supply it for study.

References

Apple, 2024. Apple Security Releases. https://support.apple.com/en-us/HT201222. (Accessed 26 September 2023).

Årnes, André, 2017. Digital Forensics. John Wiley & Sons. (Accessed 27 September 2023).

BankMyCell, 2023. Top wearable device companies by shipments market share. https://www.bankmycell.com/blog/global-smartwatch-market-share/. (Accessed 18 September 2023).

BBC, 2018. Apple health data used in murder trial. https://www.bbc.com/news/technology-42663297. (Accessed 9 April 2024).

Bhandari, Sandeepak, Jusas, Vacius, 2020. An abstraction based approach for reconstruction of timeline in digital forensics. Symmetry 12 (1), 104. (Accessed 27 September 2023).

Boucher, Jacques, Le-Khac, Nhien-An, 2018. Forensic framework to identify local vs synced artefacts. Digit. Invest. 24, S68–S75.

Breitinger, Frank, Hilgert, Jan-Niclas, Hargreaves, Christopher, Sheppard, John, Overdorf, Rebekah, Scanlon, Mark, 2024. Dfrws eu 10-year review and future directions in digital forensic research. Forensic Sci. Int.: Digit. Invest. 48, 301685.

Brignoni, Alexis, 2024. https://github.com/abrignoni/iLEAPP/blob/master/scripts/artifacts/Health.py. (Accessed 18 September 2023).

Carrier, Brian, et al., 2003. Defining digital forensic examination and analysis tools using abstraction layers. International Journal of digital evidence 1 (4), 1–12.

Casey, Eoghan, Ferraro, Monique, Nguyen, Lam, 2009. Investigation delayed is justice denied: proposals for expediting forensic examinations of digital evidence. J. Forensic Sci. 54 (6), 1353–1364.

Edwards, S., 2016. The Ios of Sauron: How Ios Tracks Everything You Do. Sans Digital Forensics and Incident Response Summit. (Accessed 18 September 2023).

EveryMac.com, 2024. https://everymac.com/ultimate-mac-lookup/. (Accessed 26 September 2023).

Friedman, Rachel, Brunty, Josh, Fenger, Terry, 2012. A digital forensic analysis on the icloud® and its synchronization to apple® devices. Res. Pap. SO 1–45.

Garfinkel, Simson L., 2010. Digital forensics research: the next 10 years. Digit. Invest. 7, S64–S73.

Hargreaves, Christopher, Marshall, Angus, 2019. Synctriage: using synchronisation artefacts to optimise acquisition order. Digit. Invest. 28, S134–S140.

Hargreaves, Christopher, Patterson, Jonathan, 2012. An automated timeline reconstruction approach for digital forensic investigations. Digit. Invest. 9, S69–S79. (Accessed 27 September 2023).

Henseler, Hans, Hyde, Jessica, 2019. Technology assisted analysis of timeline and connections in digital forensic investigations. In: LegalAIIA@ ICAIL, pp. 32–37. (Accessed 27 September 2023).

Jennings, Luke, Sorell, Matthew, Espinosa, Hugo G., 2023. Interpreting the location data extracted from the apple health database. Forensic Sci. Int.: Digit. Invest. 44, 301504. (Accessed 18 September 2023).

Luciano, Laoise, Baggili, Ibrahim, Topor, Mateusz, Casey, Peter, Breitinger, Frank, 2018. Digital forensics in the next five years. In: Proceedings of the 13th International Conference on Availability, Reliability and Security, pp. 1–14.

Maennel, Kaie, Mäses, Sten, Maennel, Olaf, 2018. Cyber hygiene: the big picture. In: Secure IT Systems: 23rd Nordic Conference, NordSec 2018, Oslo, Norway, November 28-30, 2018, Proceedings 23. Springer, pp. 291–305. (Accessed 27 September 2023).

Morrissey, Sean, Campbell, Tony, 2011. iOS Forensic Analysis: for iPhone, iPad, and iPod Touch. Apress.

Olsson, Jens, Boldt, Martin, 2009. Computer forensic timeline visualization tool. Digit. Invest. 6, S78–S87. (Accessed 27 September 2023).

Opie, Rebecca, 2019. Smartwatch reliability to be scrutinised by tech experts in alleged murder case. https://www.abc.net.au/news/2019-04-11/technology-experts-to-give-evidence-about-smartwatch-data/10995210. (Accessed 18 September 2023).

Overill, Richard E., Silomon, Jantje AM., Roscoe, Keith A., 2013. Triage template pipelines in digital forensic investigations. Digit. Invest. 10 (2), 168–174.

Palmer, Gary, et al., 2001. A road map for digital forensic research. In: First Digital Forensic Research Workshop, Utica, new york, pp. 27–30.

Reincubate, 2024. Iphone backup extractor. https://www.iphonebackupextractor.com/. (Accessed 18 September 2023).

Stanković, Miloš, Mirza, Mohammad Meraj, Karabiyik, Umit, 2021. Uav forensics: dji mini 2 case study. Drones 5 (2), 49.

SQLite Tutorial, 2015. A visual explanation of sqlite joins. https://www.sqlitetutorial.net/sqlite-join/. (Accessed 18 September 2023).

van Zandwijk, Jan Peter, Boztas, Abdul, 2019. The iphone health app from a forensic perspective: can steps and distances registered during walking and running be used as digital evidence? Digit. Invest. 28, S126–S133. https://doi.org/10.1016/j.diin.2019.01.021. ISSN 1742-2876. https://www.sciencedirect.com/science/article/pii/S1742287619300313.

van Zandwijk, Jan Peter, Lensen, Kim, Boztas, Abdul, 2023. Have you been upstairs? on the accuracy of registrations of ascended and descended floors in iphones. Forensic Sci. Int.: Digit. Invest. 47, 301660.

Zdziarski, Jonathan, 2013. Ios Forensic Investigative Methods.