Digital Investigation 24 (2018) 72–78

Contents lists available at ScienceDirect

Digital Investigation

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

Alexa, did you get that? Determining the evidentiary value of data stored by the Amazon® Echo

Douglas A. Orr [a, *], Laura Sanchez [b]

[a] University of North Georgia, United States
[b] Champlain College, United States

ARTICLE INFO

Article history:
Received 2 June 2017
Received in revised form
28 November 2017
Accepted 11 December 2017
Available online 1 February 2018

In 2014, Amazon spearheaded the hands free digital assistant trend with the release of the Amazon Echo. Powered with the Alexa Voice Service, an intelligent voice-activated cloud-based service created by Amazon, the Echo can fulfill a variety of tasks such as playing music, ordering pizza, and turning off lights, all with a simple verbal prompt from the user. Additionally, one can utilize the free Alexa App on phones and tablets to connect to and task the Echo or other third-party services and devices.

With the seeming ability to easily manage one's life and home by streamlining tasks, the Amazon Echo made its way into millions of homes since its release. According to a recent report by the Consumer Intelligence Research Partners (CIRP), currently 8.2 million Echo devices are owned in the United States. This is a 60% increase from 5.1 million in 2016 (Soper, 2017). The device also spurred competition in the Internet of Things (IoT) market with companies such as Google creating and releasing their own hands free digital assistants.

As the popularity of the Amazon Echo and other hands-free devices continues to increase, questions regarding user privacy have emerged. Upon its release, concerns were raised about whether the Echo is always *on*, actively listening in on conversations and other activities taking place in one's household. If always on, is the information being recorded? If so, how is this information being used?

Amazon addressed these concerns by providing details about how the device works. According to Amazon, the Echo is always on and listening, but is listening only for specific key words such as *wake* and *Alexa*(*Alexa and Alexa Device FAQs*, n.d.) If these key words are not used, Alexa does not pay attention to or record what is being said around it. However, once these words are spoken by the user, Alexa becomes attentive, listening for a specific request or question (*Alexa Voice Service UX Design Guidelines*, n.d.). Once awakened by the key words, the device also begins to record the user's speech which is processed in the cloud to generate a response for the user (*Alexa Terms of Use*, 2016). The recording is then saved to an Amazon server, which according to Amazon, is stored to provide a more personalized service to the user.

As more information has been made available about the workings of the Amazon Echo and the Alexa Voice Service, questions regarding user privacy and data have continued. Some now question if the data collected by the device can be utilized by law enforcement during an investigation (Rossow, 2016; Sauer, 2017; Williams and Ambrogi, 2017). This can be analyzed in two ways. One interpretation makes this a Fourth Amendment issue. Specifically, is law enforcement allowed to obtain this data? The other interpretation takes into consideration whether the data collected by the device possess any evidentiary value. Accordingly, if the data are obtained by law enforcement, what kind of information will they find and will it indeed be helpful?

The authors focus on the latter interpretation – whether data collected by the device possess any evidentiary value. We will first explain how the Amazon Echo and Alexa Voice Service function and what type of information is collected and how it is stored. Next, the use of the Amazon Echo and Alexa in criminal investigations will be explored by reviewing the digital footprint of the device as well as current legal cases of law enforcement seeking to utilize Echo generated data in an investigation. Finally, we conduct a forensic analysis of the Amazon Echo and Alexa and conclude with an opinion regarding the data's evidentiary value.

The Amazon Echo and Alexa

What is the Amazon Echo? Alexa?

The Amazon Echo belongs to a family of voice-controlled hands-free wireless and Bluetooth-enabled speakers developed by

* Corresponding author.
E-mail addresses: Douglas.orr@ung.edu (D.A. Orr), laura.sanchez@protonmail.com (L. Sanchez).

https://doi.org/10.1016/j.diin.2017.12.002

Amazon. The original Echo, a 9.3″ x 3.3″ x 3.3″ cylinder device, weighing at 37.5 oz., was initially available in 2014 to Amazon Prime members or by invitation-only, and, later, to the general U.S public in 2015. Since Echo's release, Amazon has designed a smaller version called the Echo Dot, measuring 1.3″ x 3.3″ x 3.3″ and weighing 5.7 oz, which was released in early 2016.

Both Echos are powered by Alexa, a cloud-based voice service also created by Amazon. Alexa enables the Echo and Echo Dot to carry out a variety of tasks such as play music, connect to smart home devices, get information, order food or a car, make purchases on Amazon.com, or check scheduled calendar events. These tasks can be initiated with the user's voice or manually through the free Alexa app.

To initiate a task using one's voice, the device first needs to be activated or *woken up*. This is accomplished by speaking one of the four currently available *wake* words: *Alexa, Echo, Amazon,* or *Computer* (*Change the Wake Word*, n.d.). Once awake, the device listens for a question or a command. The device also begins recording, capturing what is said following the wake word. The clip of speech is then sent to an Amazon server where a response is processed and returned to the user. Responses generally consist of fulfilling the user's request such as playing a song or reading the news. On occasion, may politely challenge the user with a question (e.g., what city?) for clarification purposes.

The Amazon Alexa app which is available for a variety of platforms and devices can also be used to initiate more complex tasks. With the app, a user can add items to their shopping carts or to-do lists; start, pause, shuffle, and stop music; set alarms and timers; and control smart home devices. This is particularly helpful when away from one's home. For example, while on vacation, one can use the Alexa app to turn lights on/off or make certain a door is locked. With the application, a user can also manage a variety of device settings such as location services and Wi-Fi network.

Regardless of how one opts to initialize tasks, all actions are synced with the Alexa app. This means that voice-controlled questions and commands made to the Echo or Echo Dot will appear in the Alexa app and that changes made in the Alexa app will equally be synced to the Echo or Echo Dot. Essentially, if an Echo or Echo Dot is used in conjunction with the Alexa app, they will remain in sync.

How does Alexa work?

The responses received from the Amazon Echo or Echo Dot and the ability to initiate tasks and make changes via the Amazon Alexa app are dictated by Alexa's *skill* set. Like humans, Alexa has a certain skill set, or capabilities, that allow it to complete tasks. However, these skills are created by developers and must be enabled by the user to function. Consequently, Alexa can only comply with and function within the range of skills that have been created and enabled for its use. If a skill does not exist or has not been enabled, Alexa will not be able to process the request submitted by the user.

Skills. As the driving force behind the Echo and Echo Dot, Amazon's Alexa Voice Service now has over 10,000 skills available (Isbitski, 2017). Varying in topic from food to games to sports, each of these skills allows users to create a more personalized experience suited to their needs and interests. Users can enable skills with the Alexa app. They may do this online, in the Amazon Alexa Skills Department, or by simply asking the Echo or Echo Dot to do it. Once enabled, users can access skills by speaking a certain set of words or phrases defined for the skill set, such as *Alexa, play Julie London station on Pandora*. If recognized by the device, the user will have their question answered or their command fulfilled.

While there are thousands of skills that a user can access, there are only three distinct types of skills that can be created by developers: *custom skills*, *smart home skills*, and *flash briefing skills* (*Getting Started with the Alexa Skills Kit,* n.d.). Each type requires a different level of work from developers when designing and building skills with one requiring developers to create skills from scratch while the others require developers to utilize a specific Application Program Interface (API). Additionally, each type requires a different level of effort from users when invoking a skill.

Custom skills. Custom skills are the most complex and require developers to design and build them from scratch. To build a custom skill, developers must define and implement a host of components, including the following:

- *Intents*: These are the actions or requests that the skill can realize; a skill's functionality. Essentially, an intent defines what a user can do with a skill. For example, an intent can be defined as *look up bank account information.*
- *Utterances*: These are the words or phrases spoken by a user to call upon a particular action, or intent. For example, Get balance for savings account.
- *Invocation Name*: This is the name given to the skill and it is applied by users when making a request. For example, the name *ABC Bank* could be given to the skill for looking up bank account information.
- *Cloud-Based Service*: The Internet accessible endpoint for a skill. It is here where skills are sent to be processed.

Together, intents and utterances make up what is called the *interaction model*. The interaction model is what allows users to communicate with a skill and for the same skill to process the request made by a user. When a skill is invoked by a user, the request is sent to the cloud-based Alexa Voice Service; the defined intent is recognized by Alexa; the request is routed to the service for the skill; the skill service receives the request and acts to fulfill the request; and, finally, the skill sends the response to the Alexa Voice Service, which is then voiced to the user via their Alexa-enabled device.

Although the most complicated skill type to create, it is also the most flexible as it can manage any variety of requests so long as all the necessary components are included and align with the development code. As explained by Amazon, it also provides developers with the most control over the user's experience (*Understanding the Different Types of Skills*, n.d.). Based on their design, developers determine how a skill is presented and how it will operate, affecting how a user can utilize and interact with the skill.

Smart home skills. As the name implies, smart home skills provide users with the ability to control their Internet of Things (IoT) devices, such as lights, locks, thermostats, and cars. Utilizing these skills is quite simple for the user. Users only need to enable a skill and connect their smart device or service to Alexa. This may require the syncing of accounts. However, once these accounts are initiated, a user can speak to their Echo or Echo Dot, or any other Alexa-enabled device, to initiate a task, such as turning on a light. When invoking skills, the name of the skill does not need to be included, making the user experience easier and more efficient.

Building smart home skills is also easier for developers as not every single component required for the skill needs to be defined and implemented by the developer. This is because smart home skills are built using Amazon's Smart Home Skills Application Program Interface (API). The API simplifies the building process by having an interaction model already in place, defining the following:

- *Device Directives*: Similar to the intents of custom skills, device directives are the actions or requests that can be handled by a skill. For example, a device directive may be *lock/unlock door.*

- *Utterance*: Like custom skills, smart home skills requests can be made by using a specific set of words or phrases. For example, *lock the front door.*

With the interaction model already built and in place, the only thing required of the developer is to define how a skill responds to a directive. This particular development code is known as the *skill adapter* (*Steps to Create a Smart Home Skill*, n.d.). If, however, there is a particular skill that the developer would like to implement, but is not available in the Smart Home Skill API, they will need to build it. In other words, the developer will have to build a custom skill.

Flash briefing skills. Alexa-enabled devices, such as the Echo or Echo Dot, can provide users with content such as news and interviews in summary form known as *flash briefings.* To receive flash briefings, users need to enable a content skill, such as *Up First* by NPR, and utter a phrase such as *Alexa, what's in the news?* Depending on the skill, *Alexa either reads text content or plays audio content*, ("Understanding the Flash Briefing Skill API," n.d) in response.

The creation of a flash briefing skill, also known as a *content skill*, may be among the easiest to create. This, however, requires more maintenance on the part of the developer. Like developers creating a smart home skill, those creating a content skill do so by using Amazon's Flash Briefing Skill API. The API builds the interaction model, defining the words or phrases that can be uttered by users to make a request while the developer provides the information necessary to display the skill in the Alexa Skills Department on Amazon's website. All of this it does in addition to content feeds. As users generally expect up-to-date information, it will be necessary for developers to update content feeds regularly.

As Alexa's popularity continues to grow, more and more companies are investing resources to develop skills and/or Alexa-compatible devices. Many companies, such as Capital One (Collier, 2016), Domino's Pizza (Domino's Pizza, 2016), and 1-800-Flowers (1-800-FLOWERS.COM, Inc., 2016), have created custom skills as it allows them to provide customers with a wider range of services on a greater number of channels advancing their reach as a business. Other companies, such as Honeywell (Honeywell, 2016) and Ford (Ford Motor Company, 2017), have opted to integrate their own products with the Alexa Voice Service, recognizing that such connectivity provides a level of efficiency sought by today's consumers.

What type of data is collected and stored by the echo?

As an interactive device that is compatible with other Alexa-enabled devices and third-party services and is accompanied by an app, the Amazon Echo collects quite a bit of data. The type of data varies and includes the following:

- All the skills enabled by a user
- All recordings of questions or requests made by the user directly to the Echo
- All requests and changes made with the Alexa app
- All responses given by the Echo
- Basic Amazon subscriber information, including payment options and shipping and billing address
- Information provided by third-party services, such as from a linked account

Having such data collected is part of the user experience when subscribing to any service offered by a service provider. To receive services, it is generally expected that users supply providers with information about themselves and that they allow providers to continue collecting information about them. Often, service providers and retailers such as Amazon, Google, and Apple state that collecting such data allows them to analyze user trends, enabling them to provide users with more customized services and features such as recommendations and advertisements (Amazon, 2016; Google, 2017; Apple, 2016).

Additionally, much of this data, if not all, is central to the Amazon Echo's functionality and necessary for the continued improvement of the Alexa Voice Service. For example, for certain skills to work, such as Spotify, the device must have access to the user's Spotify account to obtain information such as settings, preferences, stations, and playlists. This information is then stored in the cloud and made ready for the next time the user wants to invoke the Spotify skill. This stored data allows the device to not only to respond to a request but also to do so efficiently and effectively.

Alexa's aptitude for voice recognition is also improved by the collection of data. Built with artificial intelligence capable of natural learning, recordings of questions and other requests made by the Amazon Echo user help Alexa to develop a better understanding of language nuances such as accents and dialects. As explained by Cao and Bass (2016), "the more a speech recognition engine consumes, the better it gets at understanding different voices and the closer it gets to the eventual goal of having a natural conversation in many languages and situations," (para. 8). The benefit of this is enhanced communication with users, more accurate responses, and an overall improved user experience.

So, where exactly is all this data stored? Given that the Amazon Echo is powered by the cloud-based Alexa Voice Service, all collected data is stored on Amazon servers which may be located outside of the country in which a user resides (*Alexa Terms of Use*, 2016). Users can access and view this data by logging onto https://alexa.amazon.com/on their desktop or the Alexa app on their smart device (i.e., phone or tablet) with their Amazon account information. Once signed in, users can view *home cards* documenting their interaction with the Echo. The home cards display what the device heard and how it responded. Also, users will be able to listen to the snippet of speech that was recorded by the device, giving users the opportunity to hear exactly what Alexa heard and recorded. Both the home cards and the voice recordings can be deleted by users should they choose to do so.

Use of the Amazon Echo and Alexa in criminal investigations

As they continue to grow in popularity and become integrated into more devices and services, the Amazon Echo and Alexa Voice Service have become of interest to the law enforcement community. Like other types of technology, law enforcement has begun to view the Amazon Echo and its Alexa service as potential sources of *digital evidence*; "any information or data of value to an investigation that is stored on, received by, or transmitted by an electronic device," (Law Enforcement Cyber Center, para. 1). As data pertaining to the user and the device's usage are collected and stored by Amazon, it is entirely reasonable that law enforcement would consider this data valuable to their investigation. This information, when analyzed, could provide law enforcement with details regarding a user's interests, purchases, location, and engagements. Accordingly, these details could then be used to confirm or disprove a user's involvement in criminal behavior.

Of particular interest to law enforcement are the clips of recorded speech stored by Amazon. Mentioned earlier, the Amazon Echo is activated by a wake word, such as *Alexa*, and a process is invoked with a phrase like *tell me the news*. When the wake word is used, the device not only becomes alert for a question or request but also begins to actively record. The recording of the user's question or command is sent to a server to obtain a response.

According to Amazon, these recordings are then stored on a server for the purpose of improving the user experience and Alexa's voice recognition capabilities.

For purposes relating to a criminal investigation, these recordings may also provide something more valuable – sounds captured during the commission of a criminal act. While relatively short, some in law enforcement believe that such clips could contain relevant evidence pertaining to a crime. In Bentonville, Arkansas the local Police Department moved, by way of a search warrant, to obtain such recordings from Amazon while investigating the 2015 murder of Victor Collins. Although the Bentonville Police Department could extract data from the Amazon Echo located in the home of suspect James Andrew Bates, it believes "the Echo … may have recorded some of what transpired in the house on Southwest Elm Manor Avenue, during the early morning hours of Nov. 22, 2015 (Augenstein, 2016, para. 7)." It is possible that these recordings may reveal a material fact not identified in the data extracted directly from the Echo.

Until recently, Amazon declined to provide the police department with these recordings, even when presented with a search warrant, claiming that the demand was overbroad. However, on March 3, 2017, the suspect provided a waiver to the Police Department allowing them to obtain the recordings from Amazon. In turn, Amazon obliged, turning the recordings over to prosecutors that same day (Ars Technica, 2017; McLaughlin, 2017).

In addition to the data collected by Amazon's server, it may be possible to obtain evidentiary data from the cache of a device on which Alexa app is installed. The cache, which stores browsing and application data for faster loading, may contain data related to the Echo and Alexa. For example, the cache may include information regarding a user's utterances and the responses generated by the device. Having access to a device and its cache may be beneficial should an investigator be unable to extract data from the Echo or have access to the home cards or audio recordings.

Research study

The purpose of this paper is to determine whether or not data generated while using the Amazon Echo and its related Alexa Voice Service app could be of evidentiary value to an investigation. This will be accomplished by acquiring and analyzing data located 1) online at https://alexa.amazon.com/, 2) in the folder structure of the Alexa app, and, 3) in the cache of a mobile device.

Methodology

Due to the various types of data storage and their locations, two different methods were utilized in this study to acquire and analyze the data. The first method consisted of *manual extraction* which is defined in *Guidelines on Mobile Device Forensics* as "recording information brought up on a mobile device screen when employing the user interface," (Ayers et al., 2014, pg. 15). Except for the use of a digital camera for notetaking purposes, manual extraction occurs without the assistance of forensic tools. An investigator simply utilizes the screen or keyboard on a mobile device to view the contents of the device.

Manual extraction has its disadvantages. Among them is the potential to destroy or modify data. Indeed, this can pose challenges for an investigation and the admissibility of evidence in court. Additionally, it is not possible to recover deleted files which can be a source of evidentiary data. An investigator will only be able to see what is immediately available on the screen.

While manual extraction is a less sophisticated method for acquiring and analyzing data, it was the best option for examining data in two different areas. The first area was on a desktop web browser while visiting https://alexa.amazon.com/. Although generally considered a method appropriate for mobile devices, it was also a suitable option in this particular instance because it was necessary to interact with the computer's user interface to navigate the website. Furthermore, manual extraction was also selected as no other areas of the computer were of interest in the study.

Manual extraction was also utilized to analyze the contents associated with the Alexa app found on a mobile device, specifically a smart phone. In this instance, however, there was concern about altering data as the smartphone and the Alexa app were intended to undergo another type of forensic analysis in the latter part of the study. To remedy this concern, and in alignment with forensic best practices, a backup of the phone was created and stored. Should the manual extraction alter the data, the phone would be restored with the backup.

During the manual extractions, active documentation procedures were conducted. This was accomplished a chronological stepwise methodology. Information pertaining to data was also documented in both note and picture form. All images were taken with a digital camera and saved to a computer.

The second method utilized to acquire and analyze data was a *logical backup extraction*, which entails mining artifacts from the backup of a device. Generally, "[a] backup would be retrieved from the workstation (Windows or MacOS) the device typically connects to for updates or syncing music, movies, and applications," (Proffitt, 2012, pg. 6). Once obtained, the backup could be read and parsed by one of several forensic tools available, allowing the analyst to view installed applications, photographs, call logs, text messages, and calendar events, among other things.

A more recent method of acquiring and analyzing evidence, logical backup extraction has several benefits – the most important being that it is forensically sound. As data is not modified during the backup process or during analysis, the integrity of the data remains intact. Also, logical backup extraction provides another avenue for obtaining data from a device should the device be missing or damaged.

Even with access to a user's workstation and a device backup, logical backup extraction may not be an option. Like other methods of digital evidence acquisition, logical backup extraction can be prevented with sound encryption protocols. In addition to encrypting their devices, users may also opt to encrypt backups of their devices. Without a password to decrypt a backup, data located within the backup may be rendered inaccessible and unreadable.

Although the mobile device used for this study was undamaged, given the device's make and model (Apple iPhone 6), a logical backup extraction was the most viable option available for acquiring and analyzing data from a device that was not jailbroken. Furthermore, we assumed, for purposes of this study, that the mobile device and login information for https://alexa.amazon.com/ were unavailable. Such is the case in real-life scenarios.

Data collection

The initial phase of the study commenced on February 5, 2017 with the setup of an Amazon Echo Dot (2nd Generation, software version 571207720) in a communal area of a home occupied by three individuals. The Echo Dot was utilized in a single-family dwelling by all three individuals for the duration of ten weeks. During this time, the Echo Dot remained powered on and was never muted. This particular setup provided the Echo Dot with a periodic but conventional stream of data to collect in the form of voice commands.

Additionally, the complementary Alexa app was installed on an Apple iPhone 6 (software version 10.2.1). The Alexa app was required for the Echo Dot setup, specifically with a connection to

the Internet. Throughout the initial phase of the study, the app was utilized to search for and enable/disable Alexa *skills* and to write to Alexa's task lists, further contributing to data generation.

Analysis

Upon completion of the data collection phase, the examination process was initiated with manual extraction of the contents available online at https://alexa.amazon.com/and on the Alexa app installed on the mobile device. This was accomplished with the use of a laptop equipped with Windows 10, version 10.0.14393; Mozilla Firefox, version 52.0.2; an Apple iPhone 6 with iOSX 10.2.1; and the Amazon Alexa app, version 1.16.65.0. As manual extraction simply requires interacting with the user interface to view contents, forensic tools were not utilized.

The first area to be analyzed was the web-based Alexa interface, which required user login. Using the username and password, access was granted to the account associated with the Amazon Echo Dot and Alexa. The page to be opened immediately following login was the home page, which displayed a series of *home cards*. As discussed previously, home cards document and store data regarding user interactions with the Echo.

Screenshot 1: https://alexa.amazon.com login.

Screenshot 2: Sample home cards.

In total, 190 home cards were counted in the home page. The home cards covered a significant spectrum of skills from weather to music. As expected, the typical display of home cards was found to consist of the following:

- A title-generally the name of a skill or an utterance made by the user (e.g., *Alexa, tell me a joke*)
- A response provided by the device which may include an image
- Voice feedback-includes an audio clip recorded by Alexa and printed text stating what was heard by Alexa (e.g, *Alexa heard: Alexa play Jeopardy*), followed by the question *Did Alexa do what you wanted?*
- Remove card option

Screenshot 3: Home card layout.

Generally, timestamps are not located on home cards meaning that home cards do not document the date and time a question was asked or a request made. There are exceptions, however. Among them are *flash briefings* which provide the date and time the skill was invoked; *weather* which provides the date; and, *alarms* which also provide date and time.

Screenshot 4: Flash briefings and weather timestamps.

Screenshot 5: Alarm timestamps.

Home cards also document interactions initiated with the mobile Alexa app such as adding an item to a task list (shopping or to-do lists). These cards, however, do not include audio feedback.

Screenshot 6: Task lists modified with Alexa app.

After analyzing the home cards, other areas of the user account were examined. Among the sections analyzed were:

- **Now Playing** displays music currently in play, as well as the history of played music (songs, artists, albums, stations)

Screenshot 7: Now Playing.

- **Music and Books** displays a list of enabled skills related to music and books. If a skill is clicked on and happens to be a linked account, a new page will open, providing a list of stations or books related to the account.

Screenshot 8: Music and Books.

Screenshot 9: List of Pandora stations after clicking on Pandora in Music and Books.

- **Lists** displays shopping and to-do lists composed by the user either by speaking to the Echo or utilizing the Alexa app. In addition to viewing the lists, there is an option to print them.

Screenshot 10: Shopping list.

Screenshot 11: To-do list.

- **Timers and Alarms** displays a list of currently set timers and alarms.

Screenshot 12: Alarms.

- **Skills** allows users to search for and enable a variety of skills

Screenshot 13: Skills.

- **Smart Home** displays information regarding connected smart devices being controlled within the home and enabled smart home skills

Screenshot 14: Smart Home.

- **Settings** provides information regarding any Echo devices set up in the area, accounts associated with the device(s), and general information.

Screenshot 15: Settings.

Once these areas were examined for potential sources of data, the Alexa app installed on the iPhone 6 was also analyzed using manual extraction. The examination indicated that the mobile Alexa app had a similar setup as the web-based Alexa interface. Furthermore, the information provided in the mobile app was identical to that provided in the web-based interface.

The next phase of analysis was to examine the data obtained from the mobile phone by logical backup extraction using the forensic tool *Blacklight*, version 2016.3.1, by BlackBag Technologies. The process to acquire the backup was relatively simple requiring only the following steps:

- **Connect** device
- **Create new case** followed by selecting **Add Evidence**
- Select **iPhone** under **Attached/Mounted De**vices
- Select type of ingestion which in this case was **triage**
- Click on **Start**

When processing was complete, examination of the backup started by clicking on the Browser icon in the tool bar and navigating the path**/owner's iPhone/mobile/Applications/com.amazon.echo**. Here, two folders were located: **Documents** and **Library**. The Documents folder contained one file, **LocalData.sqlite**. The **Library** folder contained two other folders: **Cookies**, which contained the file **Cookies.binarycookies**, and **Preferences**, which contained the files **com.amazon.echo.plist** and **com.apple.Emojicache.plist**.

Screenshot 16: Device information displayed in BlackLight.

Screenshot 17: Amazon Echo data displayed in BlackLight.

Results

Web-based Alexa interface and Alexa app

Analysis of the web-based Alexa interface and the mobile Alexa app yielded artifacts that may be of evidentiary value in an

investigation. Of particular interest was information located in **History**, under **Settings**. This is where information pertaining to vocal interactions with Alexa are stored and organized in a series of lines, each of which contains a title and a timestamp. Titles reflect commands made by users or responses provided by Alexa such as *Text unavailable … Click to play recording* while timestamps reflect the date and time that the interaction occurred.

Information located in these listings may be pertinent to an investigation. For example, the timestamps associated with interactions may be used in the development of a timeline and to confirm or disprove the timeframe of one's alibi. Of course, these timestamps should serve as additional data to other material sources of data.

Additionally, **History** listings provide many more clips of recorded speech. While **home cards** also provide this service, recordings are generally limited to commands made by the user. In the **History** listings, however, recordings for innocuous utterances, unintended voice requests (e.g., recording dialogue from a movie or an on-going conversation), and dead air (e.g., silence when a user takes too long to respond to Alexa) were located. Furthermore, the recordings located in **History** can be a few seconds longer than those in the **home cards**, as was discovered when comparing the recordings associated with the phrase *is before he die*. In the home cards, the recording was approximately 3.55 s long, whereas in the History listings recording, it was approximately 9.46 s long.

In this study, it was possible to closely examine brief conversations of recordings in **History** and distinguish user voices. In several of the analyzed recordings, the topic of conversation and the identity of those involved could be determined. This may be beneficial during the course of an investigation to establish if certain individuals were present or if certain topics were discussed. This of course would be highly dependent upon the clarity of the recordings.

Screenshot 18: History listings under Settings.

Screenshot 19: Expanded History listing.

Also of interest in **Settings** was **Household Profile**. This allows a user to add an adult with their own Amazon account to the Alexa account associated with the Echo. By joining the household group, users can access their own content, as well as shared content. In an investigation, such a profile containing different linked user accounts could very well provide other sources of information, such as the names and email addresses of other individuals in a household. This information may be obtained from the music, audiobook, and calendar accounts associated with a user added to a household profile. This may be of benefit, for example, if a potential suspect has their own account set up as part of the profile. It may provide investigators with the opportunity to establish their presence at the location, confirm their identity, and possibly analyze some of their activities as well.

Calendars in **Settings** also serves as a potential source of data. Although calendar events are not listed in this section, email addresses associated with the calendar(s) in use are displayed. This can be particularly helpful when the email address of a person of interest is not yet known or requires confirmation.

Settings also contains information related to the user's commute. In **Traffic**, address entries are established to create a starting point and destination. When the Traffic skill is invoked, Alexa will refer to these addresses to provide commute information. Generally, users will input their home address and work address. There is also the possibility of adding additional addresses between the starting point and the destination. Although there is typically little variance between the locations, this information may assist in confirming or disproving one's daily routine or whereabouts.

Screenshot 20: Calendar accounts in Settings.

Screenshot 21: Traffic input in Settings.

In addition to several of the options in **Settings**, **Lists** may also serve as a source of evidentiary data. When users add items to their to-do or shopping list, either by voice or through the Alexa app, these items are displayed in **Lists**. It is possible that an item or action related to the investigation may appear. Although it is highly unlikely that a user will include *purchase anthrax* or *kill Bobby* on their shopping or to-do list, it is possible that something as simple as *pick up Bobby after work* could provide a valuable lead.

Finally, **home cards** on the web-based interface and Alexa app also contain information relevant to an investigation. Like the **History** listings, **home cards** provide information regarding commands made by a user, invoked skills, and voice feedback. While **home cards** do not contain timestamps, or capture every interaction between the user and device, they are visually appealing and are easy to navigate quickly, which may be beneficial to the user in emergency situations.

Mobile device cache

Due to security measures implemented by the maker of the mobile device, a limited amount of data was discovered during analysis of the logical backup extraction. Among the items found were:

- *Cookies.binarycookies*, which documents the app's persistent cookies
- *com.amazon.echo.plist,* which stores application information and, often, user settings
- *com.apple.Emojicache.plist,* which stores information related to Emojis used by the iOS device
- *LocalData.sqlite,* which stores localized data (i.e., information pertaining to direct use of the app on the mobile device) in a database

Of these files, two provided some potentially evidentiary data. The *com.amazon.echo.plist* file, for example, provided the path for accessing the Alexa app cache. The path is*/var/mobile/containers/Data/Application/E8BAD8A8-88E6-456B-9A3D-E4C4D5CCB755/Library/Caches.* Although the cache was inaccessible during the course of this examination, it is possible to access it by jailbreaking the phone and utilizing a program such as the Filza file manager. If accessed, the cache could provide a wealth of evidentiary data.

Screenshot 22: com. Amazon.echo.plist file with path to Alexa app cache.

The *LocalData.sqlite* file, on the other hand, was found to contain information related to **Lists**, an enabled Alexa skill. The database file listed items added to the shopping and to-do lists. These items are stored and visible in the database because they can be directly modified by the user on the Alexa app. Storage in the database provides users with access to the lists.

As discussed previously, **Lists** may be beneficial as they may contain content related to an investigation. Should the web-based Alexa interface or the Alexa be unavailable for analysis, this database may provide some limited, but relevant details, including added items and Epoch timestamps.

Conclusion

As the Amazon Echo grows in popularity, questions regarding whether data collected by the device has value as evidence in an investigation, holds weight as evidence (exculpatory or otherwise), and has standing in a Fourth Amendment issue will continue to emerge. The authors hope that this cursory acquisition and analysis contributes to this discussion.

Furthermore, the results of this study seem to suggest that the Amazon Echo does indeed possess data of evidentiary value.

Varying in form from timestamps to recordings, the data collected by the device has the potential to provide law enforcement with the ability to further their investigation. Used in conjunction with other collateral sources of evidence, the data collected by the Amazon Echo may be used to confirm or disprove theories established by law enforcement.

Perhaps the greatest shortcoming of this study is that the authors never asked the question *Alexa, who murdered the victim?*

References

1-800-FLOWERS.COM, Inc, 2016, April 26. 1-800-FLOWERS.COM® to Offer Floral Gifting through Amazon Alexa. Retrieved from: http://www.prnewswire.com/news-releases/1-800-flowerscom-to-offer-floral-gifting-through-amazon-alexa-300257108.html.

Amazon - Alexa Terms of Use., 2016. Retrieved January 19, 2018. https://developer.amazon.com/public/solutions/alexa/alexa-voice-service/support/terms-and-agreements.

Amazon, 2016, September 30. Amazon Privacy Notice. Retrieved from: https://www.amazon.com/gp/help/customer/display.html/ref=hp_left_v4_sib?ie=UTF8&nodeId=201909010.

Apple, 2016, September 12. Privacy Policy. Retrieved from: https://www.apple.com/privacy/privacy-policy/.

Ars Technica, 2017, March 6. Press Release: Amazon Echo Subpoena in Arkansas Murder Case. Retrieved from: https://arstechnica.com/wp-content/uploads/2017/03/echoagreement.pdf.

Augenstein, S., 2016, December 28. Could Amazon Echo Be a Witness in Arkansas Murder Case? Retrieved from: https://www.forensicmag.com/news/2016/12/could-amazon-echo-be-witness-arkansas-murder-case.

Ayers, R., Brothers, S., Jansen, W., 2014, May. Guidelines on Mobile Device Forensics (10.6028/NIST.SP.800–101r1). Retrieved from: http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-101r1.pdf.

Cao, J., Bass, D., 2016, December 13. Why Google, Microsoft and Amazon Love the Sound of Your Voice. Retrieved from: https://www.bloomberg.com/news/articles/2016-12-13/why-google-microsoft-and-amazon-love-the-sound-of-your-voice.

Collier, Z., 2016, December 16. The Story of the Capital One Alexa Skill. Retrieved from: https://developer.amazon.com/blogs/post/c70e3a9b-405c-4fe1-bc20-bc0519d48c97/the-story-of-the-capital-one-alexa-skill.

Domino's Pizza, 2016, February 3. Domino's® Adds Amazon Echo Ordering Capability in Time for Big Game. Retrieved from: http://www.prnewswire.com/news-releases/dominos-adds-amazon-echo-ordering-capability-in-time-for-big-game-300214216.html.

Ford Motor Company, 2017, January 4. Alexa in the Car: Ford, Amazon to Provide Access to Shop, Search and Control Smart Home Features on the Road. Retrieved from: https://media.ford.com/content/fordmedia/fna/us/en/news/2017/01/04/alexa-car-ford-amazon-shop-search-home.html.

Google, 2017, April 17. Privacy Policy. Retrieved from: https://www.google.com/policies/privacy/.

Honeywell, 2016, August 18. Honeywell's Entire Family of Connected Thermostats Can Make Home More Comfortable with Amazon's Alexa-Enabled Devices. Retrieved from: http://www.prnewswire.com/news-releases/honeywells-entire-family-of-connected-thermostats-can-make-home-more-comfortable-with-amazons-alexa-enabled-devices-300315359.html.

Isbitski, D., 2017, February 23. Alexa Now Has over 10,000 Skills Available [Web log post]. Retrieved from: https://developer.amazon.com/blogs/post/d5a3614a-d0ae-404c-bc7e-db5bb08a52e7/alexa-now-has-over-10-000-skills-available.

Law Enforcement Cyber Center. (n.d.) . Digital Evidence. Retrieved from: http://www.iacpcybercenter.org/officers/digital-evidence/.

McLaughlin, E., 2017, March 17. Suspect OKs Amazon to hand over Echo Recordings in Murder Case. Retrieved from: http://www.cnn.com/2017/03/07/tech/amazon-echo-alexa-bentonville-arkansas-murder-case/.

Proffitt, T., 2012, November 5. Forensic Analysis on iOS Devices. Retrieved from: https://www.sans.org/reading-room/whitepapers/forensics/forensic-analysis-ios-devices-34092.

Rossow, A.L., 2016, December 29. Amazon'S Echo May be Sending its Sound Waves into the Court Room as Our First “Smart Witness”. Retrieved from: https://www.linkedin.com/pulse/amazons-echo-may-sending-its-sound-waves-court-room-our-rossow-esq-.

Sauer, G., 2017, February 28. A Murder Case Tests Alexa's Devotion to Your Privacy. Retrieved from: http://www.geekwire.com/2017/8-million-people-amazon-echo-customer-awareness-increases-dramatically/.

Soper, T., 2017, January 25. More than 8M people Own an Amazon Echo as Customer Awareness Increases ‘dramatically’. Retrieved from: http://www.geekwire.com/2017/8-million-people-amazon-echo-customer-awareness-increases-dramatically/.

Williams, J.C., Ambrogi, B., 2017, January 20. Lawyer 2 Lawyer [Audio podcast] (Hosts). Retrieved from: https://legaltalknetwork.com/podcasts/lawyer-2-lawyer/2017/01/legalities-amazons-echo/.