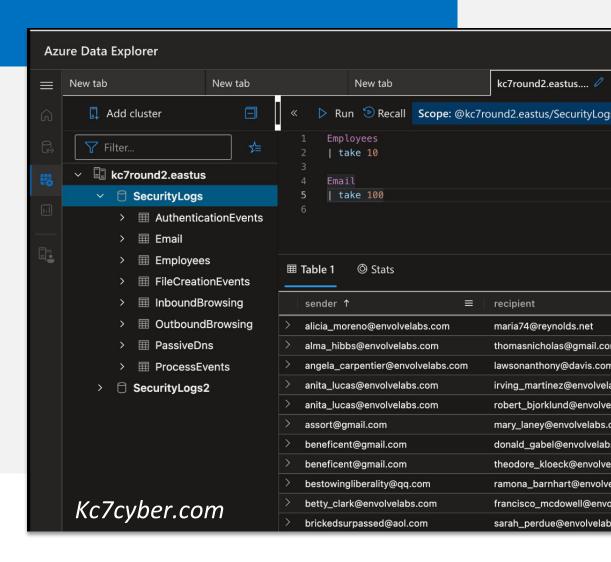


# A CYBERSECURITY GAME

Simeon Kakpovi Greg Schloemer Emily Hacker



#### **Introduction: Welcome to Envolve Labs**

Welcome to Envolve Labs Corporation! Today is your first day as a Junior Security Operations Center (SOC) Analyst with our company. Your primary job responsibility is to defend Envolve Labs and its employees from malicious cyber actors.

# **EnvolveLabs**

Envolve Labs is a med-tech startup based in the United States that was founded in 2012. Our mission is to develop a new type of flexible vaccine technology that covers many different viral strains and offers long-lasting immunity (which means no more boosters!) Our initial research has proven this technology is highly effective – we're planning to start production in Q1 2023.

EnvolveLabs has a series of key partners who contribute to the success of our business:

Partner Name	Relationship
We Sell Beakers™ (wesellbeakers.com)	A key distributor of medical-grade
	laboratory equipment, We Sell Beakers™
	provides all of the equipment for our
	world-class research facilities.
PharmaSupplies, Inc. (pharmasupplies.org)	PharmaSupplies provides Envolve Labs with
	the core ingredients used in our vaccine
	products.
Vaccine Distributors Worldwide	EnvolveLabs trusts Vaccine Distributors
(vaccinedistributors.com):	Worldwide to deliver our temperature and
	time-sensitive vaccine products to
	customers around the globe.
Research Compliance, PSC	Research Compliance, PSC is the key legal
(researchcompliance.com	partner of Envolve Labs who helps ensure
	compliance with international guidelines
	and regulations that govern vaccine
	production.

Until now, we've been laser focused on medical research and meeting production goals. But, as our work becomes more important and successful, we've realized the need to invest more in cybersecurity efforts. That's why we've hired you!

Like all good companies, Envolve Labs collects log data about the activity its employees perform on the corporate network. These security audit logs are stored in **Azure Data Explorer (ADX)** - a data storage service in Azure (Microsoft's cloud). You will use the Kusto Query Language (KQL) to parse through various types of security logs. By analysing these logs, you can help us determine whether we're being targeted by malicious actors.



You can find full documentation on ADX here: <a href="https://docs.microsoft.com/en-us/azure/data-explorer/kusto/query/tutorial?pivots=azuredataexplorer">https://docs.microsoft.com/en-us/azure/data-explorer/kusto/query/tutorial?pivots=azuredataexplorer</a>

### **Objectives**

- By the end of your first day on the job, you should be able to:
  - ✓ Use the Azure Data Explorer
  - ✓ Use multiple data sets to answer targeted questions
  - ✓ Find cyber activity in logs including: email, web traffic, and server logs
  - ✓ Use multiple techniques to track the activity of APTs (Advanced Persistent Threats)
  - ✓ Use third party data sets to discover things about your attackers
  - ✓ Build a threat intelligence report
  - ✓ Make recommendations on what actions a company can take to protect themselves

# The attackers have gotten a head start, so let's not waste any more time... time to get to work!

#### Some important links:

- The scoreboard: https://kc7cyber.azurewebsites.net/
- Azure Data Explorer: <a href="https://dataexplorer.azure.com">https://dataexplorer.azure.com</a>

### Legend

- **© Key Point** Occasionally, you will see a dart emoji with a "key point." These signal explanations of certain concepts that may enhance your understand of key cybersecurity ideas that are demonstrated in the game.
- **Question** "Thinking" emojis represent questions that will enable you to demonstrate mastery of the concepts at hand. You can earn points by entering your responses to questions from section 3 in the scoring portal.
- **Hint** "Whisper" emojis represent in-game hints. These hints will guide you in the right direction in answering some of the questions.

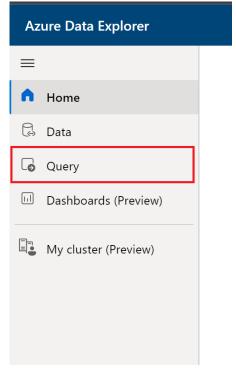
# **Section 1: The walkthrough**

#### **Getting Set Up in Azure Data Explorer (ADX)**

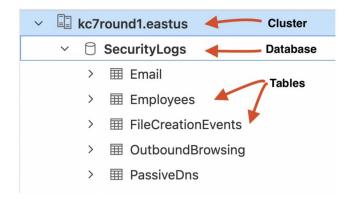
ADX is the primary tool used in the Envolve Labs SOC for data exploration and analysis. The great thing about ADX is that it is used by cyber analysts at many of the smallest and largest organizations in the world.

Let's get you logged in and started with ADX:

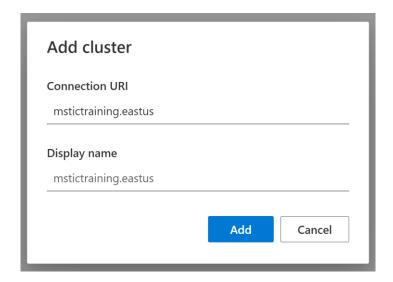
- 1. Go to aka.ms/kc7access and request access to the data
- 2. Go to <a href="https://dataexplorer.azure.com/">https://dataexplorer.azure.com/</a> and login with your Microsoft account
- 3. Click the Query tab on the left side of the screen.



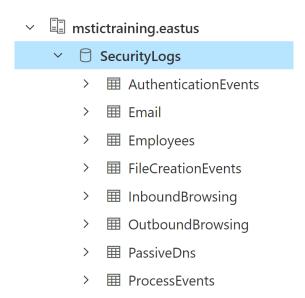
Data in ADX is organized in a hierarchical structure which consists of **clusters**, **databases**, and **tables**. All of Envolve Labs's security logs are stored in a single cluster. You'll need to add this cluster to your ADX interface so you can start looking at the log data.



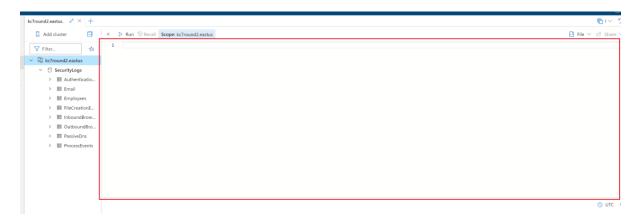
- 4. Add a new cluster using the cluster URI provided by your instructor
  - Click add cluster
  - Enter Connection URI: mstictraining.eastus



- 5. Select your database
  - Expand the dropdown arrow next to your cluster. You should then see one database, called **SecurityLogs** inside it.
  - Expend the dropdown arrow next to the SecurityLogs database.
  - Click on the SecurityLogs database. Once you've done this, you should see
    the database highlighted- this means you've selected the database and are
    ready to query the tables inside.



The big blank space to the right of your cluster list is the *query workspace*. That's where you will write the queries used to interact with our log data.



Okay, enough introductions... let's get your hands on the data.

#### First look at the data...

The **SecurityLogs** database contains eight tables. Tables contain many rows of similar data. For security logs, a single row typically represents a single thing done by an employee or a device on the network at a particular time.

We currently have eight types of log data. As you'll see in ADX, each log type corresponds to a table that exists in the **SecurityLogs** database:

Table Name	Description
Employees	Contains information about the company's employees
Email	Records emails sent and received by employees
InboundBrowsing	Records browsing activity from the Internet to devices within the company network
OutboundBrowsing	Records browsing activity from within the company network out to the Internet
AuthenticationEvents	Records successful and failed logins to devices on the company network. This includes logins to the company's mail server.
FileCreationEvents	Records files stored on employee's devices
ProcessEvents	Records processes created on employee's devices
PassiveDns (External)	Records IP-domain resolutions

**© Key Point – Over the Horizon (OTH) data**: One of the tables listed above is not like the others – **PassiveDns.** Rather than being an internal security log, PassiveDns is a data source that we've purchased from a 3<sup>rd</sup> party vendor. Not all malicious cyber activity happens within our company network, so sometimes we depend on data from other sources to complete our investigations.

You'll learn more about how to use each of these datasets in just a minute. First, let's just run some queries so you can practice using KQL and ADX.

#### **KQL 101**

Type the following query in the workspace to view the first rows in the **Employees** table. Press "run" or "shift + enter" to execute the query.

Employees
| take 10

This query has a few parts. Let's take a moment to break each of them down:



Query Component	Description
Table name	The table name specifies which table/data source the query will pull
	data from. All queries must start with a table.
Pipe character ( )	The pipe character indicates the start of a new part of the query.
	A pipe will be added automatically after typing a table name and
	pressing enter. You can also add a pipe character manually by
	holding shift and pressing the backslash (\) key. That's the one
	just below the backspace key.
Operator	The operator tells the query what exactly you want to do. The first
	operator you've learned is take, which simply takes a given number of
	rows and shows you the data there.
	You'll learn and practice using more operators soon!

The take operator is a powerful tool you can use to explore rows in a table, and therefore better understand what kinds of data are stored there.

**© Key Point – What to do when you don't know what to do:** Whenever you are faced with an unfamiliar database table, the first thing you should do is sample its rows using the take operator. That way, you know what fields are available for you to query and you can guess what type of information you might extract from the data source.

The **Employees** table contains information about all the employees in our organization. In this case, we can see that the organization is named "Envolve Labs" and the domain is "envolvelabs.com".

1. Pry it for yourself! Do a take 10 on all the other tables to see what kind of data they contain.

You can easily write multiple queries in the same workspace tab. To do this, make sure to separate each query by an empty line. Notice below how we have separated the queries for the **Employees**, **Email**, and **OutboundBrowsing** tables by empty lines on lines 3 and 6.

```
1 Employees
2  | take 10
3
4  Email
5  | take 10
6
7  OutboundBrowsing
8  | take 10
```

When you have multiple queries, it's important to tell ADX which query you want to run. To choose a query, just click on any line that is part of that query. Once you've selected a query, it will be highlighted in blue, as seen on lines 4 and 5 above.

#### Finding Out "How Many": The count Operator

We can use **count** to see how many rows are in a table. This tells us how much data is stored there.

```
Employees
| count
```

2. Plow many employees are in the company?

#### Filtering Data With the where Operator

So far, we've run queries that look at the entire contents of the table. Often in cybersecurity analysis, we only want to look at data that meets a set of conditions or criteria. To accomplish this, we apply filters to specific columns.

We can use the where operator in KQL to apply filters to a particular field. For example, we can find all the employees with the name "Linda" by filtering on the *name* column in the **Employees** table.

where statements are written using a particular structure. Use this helpful chart below to understand how to structure a where statement.

```
where field operator "value"
where name has "Linda"
```

```
Employees
| where name has "Linda"
```

The has operator is useful here because we're looking for only a *partial* match. If we wanted to look for an employee with a specific first and last name (an *exact* match), we'd use the == operator:

```
Employees
```

```
| where name == "Linda Holbert"
```

3. See Each employee at Envolve Labs is assigned an IP address. Which employee has the IP address: "192.168.0.191"?

While performing their day-to-day tasks, Envolve Labs employees send and receive emails. A record of each of these emails is stored in the **Email** table.

**© Key Point – User Privacy and Metadata:** As you can imagine, some emails are highly sensitive. Instead of storing the entire contents of every email sent and received within the company in a database that can be easily accessed by security analysts, we only capture email metadata.

Email metadata includes information like: the time the email was sent, the sender, the recipient, the subject line, and any links the email may contain. Storing only email metadata, rather than entire contents, helps protect the privacy of our employees, while also ensuring that our security analysts can keep us safe. Sometimes even metadata can reveal sensitive information, so it's important that you don't talk about log data with other employees outside the SOC.

We can find information about the emails sent or received by a user by looking for their email address in the *sender* and *recipient* fields of the **Email** table. For example, we can use the following query to see all the emails sent by "Michael Montello":

#### **Email**

```
| where sender == "michael montello@envolvelabs.com"
```

4. Parrish receive?

#### Easy as 1, 2, 3... Compound Queries and the distinct Operator

We can use the distinct operator to find unique values in a particular column. We can use the following query to determine how many of the organization's users sent emails.

```
1 Email
2 | where sender has "envolvelabs"
3 | distinct sender
4 | count
```

This is our first time using a multi-line query with multiple operators, so let's break it down:

In **line 2**, we take the Email table and filter the data down to find only those rows with "envolvelabs" in the sender column.

In **line 3**, we add another pipe character ( | ) and use the distinct operator to find all the unique senders. Here, we aren't finding the unique senders for all of the email senders, but only the unique senders that are left after we apply the filter looking for rows with "envolvelabs" in the sender column.

Finally, in **line 4**, we add another pipe character ( | ) and then use the count operator to count the results of lines 1-3 of the query.

5. Plow many users received emails with the term "vaccine" in the subject?

#### **Tracking Down a Click: OutboundBrowsing Data**

When employees at Envolve Labs browse to a website from within the corporate network, that browsing activity is logged. This is stored in the **OutboundBrowsing** table, which contains records of the websites browsed by each user in the company. Whenever someone visits a website, a record of it stored in the table. However, the user's name is not stored in the table, only their IP address is recorded. There is a 1:1 relationship between users and their assigned IP addresses, so we can reference the **Employees** table to figure out who browsed a particular website.

If we want to figure out what websites Annie Jackson visited, we can find her IP address from the Employees table.

```
Employees
| where name == "Annie Jackson"
```

The query above tells us her IP address is "192.168.3.168". We can take her IP address and look in the **OutboundBrowsing** table to determine what websites she visited.

```
OutboundBrowsing | where src_ip == "192.168.3.168"
```

6. Phow many unique websites did "Keith Mitchell" visit?

#### What's in a Name? All about Passive DNS Data

Although domain names like "google.com" are easy for humans to remember, computers don't know how to handle them. So, they convert them to machine readable IP addresses. Just like your home address tells your friends how to find your house or apartment, an IP address tells your computer where to find a page or service hosted on the internet.

**(6) Key Point – Practice Good OPSEC:** If we want to find out which IP address a particular domain resolves to, we could just browse to it. But, if the domain is a malicious one, you could download malicious files to your corporate analysis system or tip off the attackers that you know about their infrastructure. As cybersecurity analysts, we must follow procedures and safeguards that protect our ability to track threats. These practices are generally called operational security, or OPSEC.

To eliminate the need to *actively resolve* (that is- directly browse to or interact with a domain to find it's related IP address) every domain we're interested in, we can rely on *passive DNS* data. Passive DNS data allows us to safely explore domain-to-IP relationships, so we can answer questions like:

- Which IP address does this domain resolve to?
- Which domains are hosted on this IP address?
- How many other IPs have this domain resolved to?

These domain-to-IP relationships are stored in our **PassiveDns** table.

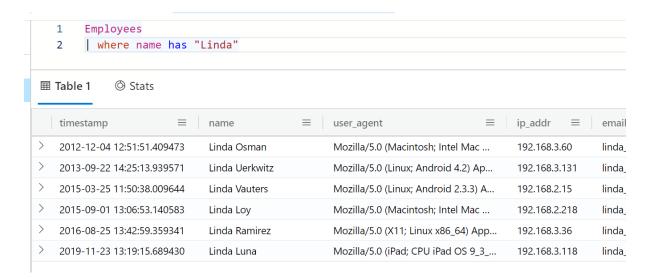
- 7. Phow many domains in the **PassiveDns** records contain the word "vaccine"? (hint: use the contains operator instead of has. If you get stuck, do a take 10 on the table to see what fields are available.)
- 8. What IPs did the domain "biotechenvolv.science" resolve to?

#### Tet statements – making your life a bit easier:

Sometimes we need to use the output of one query as the input for a second query. The first way we can do this is by manually typing the results into next query.

For example, what if we want to look at all the web browsing activity from employees named "Linda"?

First, you would need to go into the **Employees** table and find the IP addresses used by these employees.



Then, you could manually copy and paste these IPs into a query against the **OutboundBrowsing** table. Note that we can use the *in* operator to choose all rows that have a value matching any value from a list of possible values. In other words, the == (comparison) operator looks for an exact match, while the *in* operator checks for any values from the list.



Although this is a valid way to get the information you need, it may not be as elegant (or timely) if you had 100 or even 1000 employees named "Linda."

We can accomplish this in a more elegant way by using a <u>let statement</u>, which allows us to assign a name to an expression or a function. We can use a <u>let</u> statement here to save and give a name to the results of the first query so that the values can be re-used later. That means we don't have to manually type or copy and paste the results repeatedly.

```
1 let linda_ips = Employees
2 | where name has "Linda"
3 | distinct ip_addr;
4 OutboundBrowsing
5 | where src_ip in (linda_ips)
6
```

On the left of the let statement is the variable name ("linda\_ips" in this case). The variable name can be whatever we want, but it is helpful to make it something meaningful that can help us remember what values it is storing.

```
1 let linda_ips = Employees
2 | where name has "Linda"
3 | distinct ip_addr;
4 OutboundBrowsing
5 | where src_ip in (linda_ips)
```

On the right side of the let statement in the expression you are storing. In this case, we use the distinct operator to select values from only one column – so they are stored in an array – or list of values.

```
1 let linda_ips = Employees
2 | where name has "Linda"
3 | distinct ip_addr;
4 OutboundBrowsing
5 | where src_ip in (linda_ips)
```

The let statement is concluded by a semi-colon.

```
1 let linda_ips = Employees
2 | where name has "Linda"
3 | distinct ip_addr;
4 OutboundBrowsing
5 | where src_ip in (linda_ips)
```

After we store the value of a query into a variable using the let statement, we can refer to it as many times as we like in the rest of the query. The stored query does not show any output. Remember, however, that your KQL query must have a tabular statement – which means that you must have another query following your let statement.

9. Plow many unique URLs were browsed by employees named "Karen"?

**6 Key Point – Pivoting:** Part of being a great cyber analyst is learning how to use multiple data sources to tell a more complete story of what an attacker has done. We call this "pivoting." We pivot by taking one known piece of data in one dataset and looking in a different dataset to learn something we didn't already know. You practiced this here when we started in one dataset – the Employees table – and used knowledge from there to find related data in another source – OutboundBrowsing.

## **Section 2: Introducing the hackers**

🎽 Use the database SecurityLogs2 for this section of the exercise 🎽

Now that you've completed your initial round of training, you are ready to work your first case in the SOC!

A security researcher tweeted that malicious file called "ResearchBibliographyGenerator.zip" was being used by hackers. Apparently, the hackers are sending this file out via phishing emails email.



▲ NOTE! This domain and others encountered in this game are fictional and are not representative of actual malicious activity,

**© Key Point – Open Source Intelligence (OSINT):** Security researchers and analysts often use free, publicly available data, like Twitter! We call this public data OSINT, and it can be a great way to get investigative leads. Like all public data sources on the internet, you should follow up any OSINT tip with rigorous analysis, rather than blindly trusting the source.

Use the tipper above to answer the following questions:

You can optionally submit your answers to the scoreboard at https://kc7cyber.azurewebsites.net/ to get feedback and earn points.

- 1. Which user has the file named "ResearchBibliographyGenerator.zip" on their machine?
- 2. When did the user above download "ResearchBibliographyGenerator.zip"?
  - a. What domain did the user download the file from?
- 3. When did the user above receive an email containing a link to "ResearchBibliographyGenerator.zip"?
  - a. What was the subject of the email in (3)?
  - b. Who was the email in (3) sent by?
  - c. What other domains were delivered via links in emails with the same subjects as the email in (3)?
  - d. What "interview" themed subject was sent by an actor-controlled email address associated with the email in (3)?
  - e. Which "fda" themed email address sent an email with the subject from (3d)?
- 4. What other ".science" domain is closely associated via Passive DNS to this cluster of activity to the domain used to deliver "ResearchBibliographyGenerator.zip"?
- 5. Which .dll file was dropped on the victim machine shortly after the user downloaded the zip: ResearchBibliographyGenerator.zip

**Hint**: Files that are created on employees' devices are captured in the **FileCreationEvents** log. Try looking there to see which employees downloaded this file.

- a. The dll file is observed on Virustotal under what file name?
- b. Which six letter reconnaissance command was executed on the Machine of the user that loaded the implant above?

**Hint**: Try narrowing down on one particular device that downloaded the EnvolveLabs\_Research\_Tool.7z file. Then, look in both the **FileCreationEvents** and **ProcessEvents** logs to find new files and processes created <u>around the time</u> when the file was downloaded.

- c. What IP address does the dll beacon to?
- 6. What is the domain name of the legitimate service the adversary used to exfiltrate data from the victim machine?

### **Section 3: Report Analysis**

January 31, 2023 • 5 min read

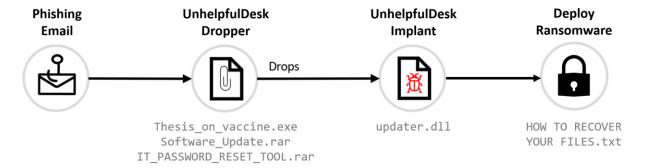
# The 10X cyber adversary: ITINIUM APT Group Leverages New 'UnhelpfulDesk' Malware In Ransomware Attacks

VulnerableArray Security Intelligence

VulnerableArray Security Intelligence (VSI) discovered a new APT group named ITINIUM. ITINIUM likely operates from Urzikstan and primarily targets pharmaceutical companies in Europe and North America.

Beginning in late 2022, ITINIUM launched a targeted phishing campaign in which it delivered emails containing malicious links. The resulting files led to implants from one new malware family that VSI calls UnhelpfulDesk. VulnerableArray Security Intelligence assesses with a high degree of confidence that the UnhelpfulDesk malware is custom-developed and use exclusively by ITINIUM.

ITINIUM uses UnhelpfulDesk in order to gain access to victim systems and encrypt their files.



#### UnhelpfulDesk

UnhelpfulDesk malware implants are dropped by files with names that resemble legitimate IT functions, such as software updates or password resets, or medical research topics, such as vaccine research. These files are delivered to victims via malicious emails containing links to download the files.

#### UnhelpfulDesk droppers:

Filename	Sha256
Thesis_on_vaccine.exe	232568cb9c5d1b3698334c504b173e637826d79074fb8fa
	23a54981578eb7dc9

ResearchBibliographyGenerator.pptx	6e4a6278077f310e69017dba9a173d9d27eddec9236231 e1717a475c26242ae6
Software_Update.rar	2f2e5f20a726e9710b9c5c7c681e66240f854acd48107e5c d193d6133297b72f
IT_PASSWORD_RESET_TOOL.rar	fe04d68b163bbf432196c0d7bb184176a4260630374c93 c916cc6b52fc9855f7

#### **Dropped implants**

Filename	Sha256
updater.dll	3666cb55d0c4974bfee855ba43d596fc6d10baff5eb45ac8b6432a7d604cb8e9
updater.dll	42a337bcec26df0130a11baf9e60179993851b88f1cabec52973f88774e903fb
updater.dll	ea05ff75fef906a60545129a7c5bea2956bfde63b8e714eb42db3ae50b99dec3
updater.dll	370ce39ba328329ff16b5ede1079f6402e68abceb34e65cb31883a3b3730b530
updater.dll	e3970346ff7fcc3665f027d7f221968087f3c42705f5799fbc1d2811ab1ca4ea

Once successfully deployed, the UnhelpfulDesk implant executes reconnaissance via the following commands:

ping 8.8.8.8
whoami
net user Administratr

Following this, the malware will encrypt files on the machine and demand a ransom to decrypt the files. The ransom note is pulled down from Pastebin as shown below:

curl https://pastebin.com/HOW%20T0%20RECOVER%20YOUR%20FILES.txt

#### Other IOCs

214.217.73[.]146 65.69.253[.]41 199.57.49[.]250 install-notice[.]com remarkablevirus[.]tech noreply\_info[@]hotmail.com vaccinejournal[@]yahoo.com

<sup>\*\*</sup> Samples of the UnhelpfulDesk implant files detected by VulnerableArray researchers are available on VirusTotal.

#### Now it's up to you...

Our CISO has asked you to evaluate this report from VulnerableArray and determine whether it is accurate. While making your assessment, consider the following questions:

- 1. Do all the reported indicators belong to the same cluster of activity? How do you know?
  - **Hint:** Use the Diamond Model (Adversary, Victim, Infrastructure, Capabilities) to help you think about clustering distinct groups of activity. It Look for similarities and differences in each of the four Diamond Model vertices.
- 2. The report claims that the UnhelpfulDesk malware is to ultimately deploy ransomware and encrypt files on an infected system. Do you agree with this assessment? Or do you see evidence of alternative actions on objectives?
  - **Hint:** Try looking for activity related to the malware-based indicators shared in the blog, then identify a few compromised systems. Do you see post-compromise activity on any of these systems that's different from the ransomware described in the blog?
- 3. What analytical mistakes, if any, were made by the authors of the blog?
  - **Hint:** Some processes are executed automatically by the malware upon execution. Other processes are run manually (hand-on-keyboard) by the operator after the command and control channel is establish.
- 4. Is the UnhelpfulDesk malware unique to the ITINIUM actor? How do you know?
  - **Hint:** The updater.dll implants appear to be dropped from files with two separate themes (IT and research). Think about why that might be the case.
- 5. Are there multiple actors targeting Envolve Labs? If so, can you describe the Tactics, Techniques, and Procedures (TTPs) of each of them? How are they similar? How are they different?
- W Hint: Compare and contrast the diamond model for each of the observed clusters of activity.
- 7. How might gaps in visibility have contributed to the conclusions of the blog author(s)? How might they contribute to your own analytical assessments?