

Title

Tuesday, July 8, 2025 11:39 AM

Title: Creating a lightweight SIEM for home use

Author: Charlie Williams

Student Number: 2235685

Course: Computing (Computer Networks and Cyber Security)

Mind Map

Tuesday, July 8, 2025 11:46 AM

Mind map: <https://mm.tt/map/3763159967?t=iHeKdBAzoa>

The above mindmap details my thought process regarding what I should look into as part of my project. After careful consideration, I have decided to create a lightweight SIEM for home use.

What is a SIEM?

A SIEM, short for security information and event management solution, aggregates typically large amounts of data from collectors installed on devices. [1][2] These devices can range from standard computers, laptops, and servers, to more obscure things like firewalls, and even IOT (internet of things) or OT (operational technology) like smart fridges or medical devices. SIEMs are integral to the continued security of a given environment. A threat hunter sifts through those massive amounts of data to look for signs of malicious activity.

Why?

Cyber security as a whole is not well understood by the public. Be it through misrepresentation through film, or deceptive advertisement (such is the case with most VPNs- the average person does not need a VPN)[2]. Often, people believe that the free version of Microsoft Defender is a fool proof piece of software that can and will stop any malicious activity. Therefore, I believe at the very least, that raising awareness and providing simple SIEM solutions for home environments is a strong step in the right direction toward helping people avoid crippling cyber attacks.

[1] <https://www.microsoft.com/en-gb/security/business/security-101/what-is-siem>

[2] <https://www.sectorlink.com/article/the-truth-about-vpn-services-benefits-misleading-claims-and-frustrations>

Aims and Objectives

Monday, July 28, 2025 7:41 PM

Aim: To set up a lightweight SIEM suitable for home use.

Objectives:

- Must be as simple to deploy and operate as possible.
- Must utilise open source tools.
- Must include log retention for construction of dashboards
- SIEM itself must be set up securely (e.g, minimising vulnerabilities, ensuring open source tools used are updated, etc.)

Research Philosophy

Monday, July 28, 2025 8:18 PM

I intend to employ a pragmatic approach to my research. I've looked into each of the suggested research philosophies and each of them to be too restrictive. I don't intend to constrain myself to a single method of thinking or interpretation. I prefer a versatile method that fits the needs of my project as I require it.

"This type of knowledge is changing, dynamic, developing, and oriented towards its consequences in human nature." [1]

This suits my project well, because it delves into both scientific aspects, where the truth is always objective, and theoretical aspects, where it is subjective.

Objective – software comparisons, including what software I choose to utilise for my project, have objective advantages over one another. I objectively have vast amounts more experience with some than others.

Subjective – this surrounds parts of my research that are not technical. This includes gauging subjective opinions on aspects such as ease of use, and practicality.

Pragmatism enables me to address both of these needs holistically.

[1]

M. Elgeddawy and M. Abouraia, *Proceedings of the 23rd European Conference on Research Methodology for Business and Management Studies*. 2024. Available:

<https://books.google.co.uk/books?hl=en&lr=&id=pj8XEQAAQBAJ&oi=fnd&pg=PA71&dq=pragmatic+research+philosophy+flexible&ots=C3IAdvQCC0&sig=7vqw0hra!NN5uGS1EhWK6JG83B4#v=onepage&q&f=false>

Search Tools

Monday, August 4, 2025 8:34 PM

The main search database I tend to utilise for my research Google Scholar. Google Scholar is a freely available and consistently helpful research database provided by Google. As it is not specialised, it is versatile and provides different materials, such as articles and books [1].

The criteria for articles Google considers "scholarly" are not known, resulting in varied quality between materials.

One source I intend to use frequently is the MITRE ATT&CK database. Most cyber security professionals will refer to it frequently as it contains a comprehensive list of tactics and techniques utilised by threat actors. It also documents information on what actors use those techniques, and how, making it a very useful tool.

This site- <https://fordham.libguides.com/Cybersecurity/Databases> - refers to several existing databases, such as the ACM digital library, computer science library, and the center for internet security.

[1] <https://atlantictu.libguides.com/googlescholar/pros>

Validity, Reliability, Generalisability, Ethical Considerations

Thursday, August 7, 2025 7:22 PM

In this section, I will discuss the validity, reliability, generalisability, and ethical considerations for my research, along with what strategies I intend to implement to ensure those values are maintained.

Validity

Validity is the most important of the considerations I intend to take with my research. It dictates whether the research itself is accurate. This is quite obviously essential as research that is not accurate is not useful for anyone. To ensure my research is valid, there are a few strategies I can implement:

- Research findings I cannot back with my own anecdotal experience must be verified by another external source.
- My tests and their results should be verified by another IT professional with relevant experience.
- Ensure any surveys/questionnaires adhere to the subject matter and are not off topic. [1]

Reliability

In research, reliability refers to the "measure of the stability or accuracy of the methods and results of an analysis." [2] Your findings need to be reproducible given set variables. For example, water will reliably boil when heated to boiling point. Poor reliability would mean being unable to produce the same result after a number of procedures [2].

- Testing must be done with multiple machines.
- Ensure findings are not contradicted, and if they are, acknowledge this, and explain why.

Generalisability

Generalisability, as the name suggests, is how much your findings can be applied to a more general context [3]. Scribbr.co.uk states that:

"Research results are considered generalisable when the findings can be applied to most contexts, most people, most of the time."[3]

I believe this will be the case for my findings. As previously stated, computing is often binary- it is either correct or incorrect.

- Findings can be applied in more than just one environment or industry.
- Findings are applicable to my own research.

Ethical Considerations

According to another article by scribbr.co.uk, ethical considerations in research are:

"a set of principles that guide your research designs and practices. These principles include voluntary participation, informed consent, anonymity, confidentiality, potential for harm, and results communication."[4]

Research as a whole should be undertaken with the end goal of bettering humanity. I strongly believe that my research aligns with this- the creation of a SIEM and its use for home computers is absolutely a positive for humanity- it can potentially help people avoid crippling cyber-attacks and the like. That said, there are still potential ethical issues to consider:

- All surveys etc. Must be done with the survey taker having full knowledge of what their answers are contributing to, and must be kept entirely anonymous. No information other than what is necessary will be extracted.
- SIEMs intake massive amounts of data from collectors, which, on home devices, could potentially contain private data or PII (personally identifiable information). Therefore, I must ensure documentation is robust enough to prevent no more data than what is necessary being collected. Also, good data security practices must be followed, including encryption, and safe data transfer. I have a responsibility to ensure any software I use is maintained, and not a security gap in itself.

[1] <https://dovetail.com/research/validity-in-research>

[2] [https://uk.indeed.com/career-advice/career-development/reliability-in-research#:~:text=What%20is%20reliability%20in%20research,research%20\(with%20types%20and%20benefits\)](https://uk.indeed.com/career-advice/career-development/reliability-in-research#:~:text=What%20is%20reliability%20in%20research,research%20(with%20types%20and%20benefits))

[3] <https://www.scribbr.co.uk/bias-in-research/generalisability>

[4] <https://www.scribbr.co.uk/faqs/what-are-ethical-considerations-in-research>

Research Strategies/Data Collection

Sunday, August 10, 2025 3:28 PM

There are two primary research strategies I would implement for this project.

- Experimental
- Quantitative Survey

Experimental

Experimental research involves implementing a scientific method for obtaining data [1]. It often is preceded by a hypothesis, and is comprised of variables, both dependant and independent. Experimental research is vital for my project. For me, this means testing a version of my SIEM in my own isolated environment. The goal is to see whether I can use it to detect threats. This experiment will be what my pilot study will be surrounding, as such I will go into details in that section.

Quantitative Survey

Because my project hinges quite heavily on usability for non IT professionals, I need to gather information on what prior knowledge most people have. To do this, I will be using a survey. I could utilise either a qualitative survey, or quantitative. Qualitative surveys usually take the form of interviews, and are a lengthy but bountiful process in terms of information. You will gain deep insights into what each survey taker understands and believes. Quantitative analysis is more centred around short, to the point questions, with pre-prepared answers. They are designed to be fast, and more reliant on large amounts of results to extract insights from the data. I believe this is a better approach for my project- I do not intend to personally tailor the software to each person, so qualitative data is less valuable- instead, I need the numbers on what prior knowledge is available. It is also considerably faster, both in planning and execution.

Pilot Data Collection Method

A pilot (or pilot study, pilot data collection) works as a trial as part of your research. In the case of healthcare research, for example, this can come in the form of clinical trials [2]. They are done to gather additional information about the feasibility or other potential issues for later experiments [2].

In my case, the pilot study will be the SIEM operating within a controlled environment. I will create a virtual machine, install my SIEM solution, preform normal activities, trigger my own attacks, then test whether they are detected.

Requirements for pilot study:

- A virtual environment, isolated from other devices to prevent unintentional spread of any attacks. It will need to be filled in with normal activity, to make detections more realistic. I will need to generate a significant amount of normal activity- I think a reasonable ratio of normal activity to malicious activity for this study is 1 malicious event per 100 benign events. In a real environment, the ratio would be much, much higher on the side of benign events, possibly going into the hundreds of thousands [3], but this is not realistic for this project. The environment itself will utilise Windows 10, with no antivirus or security controls. [6]
- SIEM solution- I will be using Logscale community edition. CrowdStrike's Logscale is industry standard and has been proven to be one of the most solid SIEM solutions in said industry.
- Collector – Collects the logs for the SIEM- I will be using LogRhythm Open Collector due to it's ease of use and open source nature.
- Attacks created by me. They will not actually be performing any malicious activity, but will be interacting with the system and modifying things that would, in any other environment, signify malicious activity. This will likely include things like proxy execution via MITRE attack tactic T1218.005 [4] or persistence via tactic T1053 [5][6].

[1] <https://resources.pollfish.com/survey-guides/what-is-experimental-research-how-is-it-significant-for-your-business>

[2] <https://casp-uk.net/news/what-is-a-pilot-study>

[3] <https://chatgpt.com/> (I can back up these facts with my own experience)

[4] <https://attack.mitre.org/techniques/T1218/005>

[5] <https://attack.mitre.org/techniques/T1053>

[6] Written after feedback on my pilot study

Qualitative and Quantitative Analysis

Tuesday, August 12, 2025 11:01 PM

Qualitative

Qualitative research or analysis is focused on stories, feelings, and other human complexities that cannot be properly captured by numerical analysis [1]. It is usually gathered through unstructured interviews or participant observation [2], which provides rich insights. These kinds of analysis compliment fields that are not well defined by numbers- psychology, social sciences, and education. Quantitative analysis may be able to spot a correlation, but qualitative analysis will stand a considerably better chance of explaining the connection behind it.

Quantitative

Quantitative analysis is the inverse of qualitative analysis, and uses statistics. While qualitative is usually focused on smaller groups with deeper results, quantitative analysis is performed on a much larger sample of people. Thus, it is easier to produce more reliable results using this method of analysis. Quantitative analysis is often done using software like Excel, PowerBI or Access- they help us visualise massive amounts of data in the form of graphs. This type of analysis tends to benefit fields like physics, computer science, and finance. The important part is how you can form a story using that data-

"Without analytics, data is just noise. Analyzing data helps you make decisions which are informed and free from bias." Shiv Sharma - Head of Content, Taskworld [3]

Quantitative data is excellent when you need a yes or no answer (exactly like a computer, funnily enough)- for example, does advertising method A or B result in more sales? Does script C or D make more efficient use of the machine's available resources?

[1]

S. Tenny, G. D. Brannan, J. M. Brannan, and N. C. Sharts-Hopko, "Qualitative Study," *PubMed*, 2021.

<https://www.ncbi.nlm.nih.gov/books/NBK470395>

[2] Unit 4 Topic 2: There's a theme here

[3] "Quantitative Data Analysis: A Complete Guide," *Hotjar.com*, 2023.

<https://www.hotjar.com/quantitative-data-analysis>

Pilot Study Feedback

Sunday, August 10, 2025 4:21 PM

The feedback from my pilot study was very helpful. There were some key factors I had not included in my pilot study plan. This included:

- Further explanation of parameters needed for controlled environment
- Advanced attack description
- Data ingestion rates, what volume of data is needed to try and replicate a real environment to a reasonable degree

Unit 1 Topic Questions/Answers

Wednesday, August 13, 2025 9:41 PM

You have come to university to study a particular subject, specific area of interest?

The specific area of interest that caused me to come to university to study is cyber security. I, at that point, had not been very interested in cyber for long, but it was a natural fit for me I feel.

What will make your research at the uni meaningful?

It's applicability in real scenarios to help bolster cyber security. I specifically aim to broaden my own understanding which will assist me in coming up with new hypotheses to discover new threats.

Think back to a purchase of something valuable. What did you do before parting with your money?

I think there are a couple of examples that come to mind. First are cosmetic purchases related to a hobby, and the second was the purchase of my car. Before purchasing my car, I considered:

- Whether I was getting a good price for the vehicle
- Whether I would need the money further down the line
- Whether if I waited longer, a better vehicle or price would come around
- If I really needed the car at all, is public transportation possible etc.
- The security of the establishment- secure transfer of money is assumed at almost all digital bank transfers nowadays, but when working with this much money, the concern is absolutely warranted. What if I lost it all and didn't even get the car?
- Can I trust this garage with a purchase of this volume?
- Do I have the correct amount of money in the correct account, ready to be withdrawn for the purchase?
- How quickly is the value of the car going to deteriorate?

When making the more cosmetic purchase, which was vastly cheaper but still a lot of money, I considered:

- Do I really need this?
- Have I earned the right to spend this amount of money?
- Will I be judged for spending this money?
- Will the value of the item increase or decrease over time? By how much?
- Is it available for cheaper anywhere else?
- Longevity? Is maintenance required?

Remembering and Recalling: The thank you tell me more approach

I am personally not great at receiving feedback. I have a bad habit of defending the work like my life is on the line, which is bad practice. I try my best to listen to and incorporate ideas in regardless of those feelings. On the whole, I should strive to be generally accepting of meaningful and constructive feedback. I think one thing that can help me do this is to try and contextualise feedback- sometimes you will be given feedback as simple as one sentence, which can be difficult to acknowledge. Trying to contextualise it can help you tie that feedback back into your project to

work on improving it.

Unit 2/3 Topic Questions/Answers

Wednesday, August 13, 2025 9:44 PM

Is your previous assignment just facts or is your knowledge being applied?

It kind of depends? Some of my assignments don't involve as much theory or research, they are either correct or incorrect. They are technical- they either work or they don't work. You can measure how correct they are by their efficiency, security etc. (there is a right answer and a wrong answer.) I would therefore argue that my knowledge is being applied constantly- and in the theoretical topics, my knowledge is applied by explaining what exactly the theory means, what it results in.

Mixed research is the best fit for my chosen subject as qualitative or quantitative research would not satisfy it's requirements. They are complex and fairly open ended.

How will you collect data for your next assignment?

Thanks to my workplace, I have a vast amount of high quality, premium tailored reports I have access to. I can pull from them to gather lots of information as needed. My project is likely to pertain to threat hunting- I work within CrowdStrike fairly frequently and they have good documentation on both the theoretical and practical implementation of threat hunting in CrowdStrike- both of these will help me build up my research. I also need to research threat hunting online, via videos, papers, reports, etc.

Torrance test of critical thinking:

As many uses for a brick as I can:

Paperweight

If you hollow it out, flowerbed

Construction

Hit someone over the head with

Projectile

Art, carve stuff into it

Use it as a doorstop

Other kind of weight, stop stuff rolling etc.

Woodlouse house

Home alone esque homemade traps

Use the 4 w's to critically evaluate a problem statement you will use

How do real world events influence the cyber threat landscape from a healthcare perspective?

Who- these influence literally everyone- everyone needs healthcare. In the short term, it mainly impacts criminals, state intelligence (GRU, MI5, FBI) and cyber security teams worldwide, particularly those working within critical infrastructure

What- Real world events, like war, politics, and terrorist attacks, heavily influence the cyber security

threat landscape. They can cause huge upsurges in attacks from a variety of sources for complex reasons.

Where- Everywhere. It affects me in the workplace, but for most the closest they would come to it is hospitals/healthcare. For example, the semi-recent attack on NHS Scotland that resulted in mass outages.

Why- How can we best combat this? Can we trust criminals and foreign intelligence not to put innocent lives in danger over war and politics?

What did the video tell you about the power of networks?

The video highlighted the concept of expanding your network into other fields or industries, often because they will have different methods of doing things and ways of thinking that you may not have originally thought of. I am too early in my career to really have been presented with opportunities from personal connections.

Unit 4 Topic Questions/Answers

Wednesday, August 13, 2025 9:48 PM

Write down what comes up when you look up quantitative analysis

What is quantitative analysis in cyber security? Quantitative analysis is a **critical component of cyber security risk management**. It involves the use of mathematical or computational methods to measure the potential impact or probability of a given security risk.

[What is Quantitative Analysis in Cyber Security ... - Cyber Insight](#)

Academic success to students reading age

I would argue there's probably a minor to medium positive trend here most likely- and would use a line graph with two lines. A bar chart would not work. This is because academic success is not measured wholistically with one single statistic (in my opinion), it is best measured by a longer series of results.

The aim of art is not to represent the outward appearance of things, but their inward significance.

An image may only show something in it's actual, real state. Art, on the other hand, is not necessarily real, and therefore, can have more meaning intrinsically built into it. Emphasis can be placed on aspects, the likes of which do not exist in real life. It can force us to consider things from angles that would not usually be considered.

Write a post in your **Module Blog** considering:

- Why is it important to begin to organise your qualitative data before data collection?
- How will this affect your planning for your qualitative research in future?

It depends how much data you're working with. If it's a small amount, you may not need to. With larger amounts of data though, I would argue it's impossible to work without first organising. I don't intend to record any interviews or anything, probably just transcribe answers and such.

Being able to sort your data and identify themes and patterns is the next stage in qualitative analysis and is often where your research starts to get exciting!

- Do you see patterns and themes in your everyday life?
- How do you recognise them and what do they look like?

Write a couple of sentences describing what a theme or pattern is.

Of course I do – we have pattern seeking brains and patterns are literally everywhere. Everything is part of a bigger pattern. For example, my sleep pattern. I have a pretty bad sleep pattern, which leads to a pattern of me waking up late, which leads to a pattern of me being late to things, which leads to a pattern of me missing important information, which leads to a pattern of me scrambling to catch up later.



Reflect

Reflect on your own experiences.

What does failure mean to you?

What mistakes or failures have you made in the past that really **set you back**?

During a training course, I failed to grasp at all the information I needed in a test/exercise sort of thing. I focused way too much on one aspect, and watching the host absolutely blast through it (obviously he knew beforehand but still) and seeing just how much I'd missed did a really big blow to my confidence.

Looking back though, it makes me glad just how far I still have to go. I'm nowhere near my potential, and I have plenty of time to reach towards it.

Can you think of someone famous or a personal acquaintance that seems to be successful from having been lucky?

By this we mean someone that has some claim to success from something they did rather than winning the lottery.

A great example is people made something ridiculous, and it captures the attention (and wallets of) many people. The person themselves does not know ahead of time that this is going to skyrocket them to success- there is no way they could have- but they are nonetheless. Culture and humour are nigh-impossible to predict, so absurd ideas like selling pet rocks can sometimes take off just because they are so comical. I would argue the fidget spinner, which exploded in popularity around 2014, was largely luck- the patent on the original design ran out and wasn't redeemed and someone took that idea and produced them at a massive scale for a wider audience and made millions.