# Visualizing Cybersecurity Data for Detection and Assistance in Cyber Operations

| Jason Ingalls | Kaur Kullman, PhD | Colonel Judson Dressler, PhD |
|---|---|---|
| Ingalls Information Security LLC | University of Maryland, Baltimore County | United States Air Force Academy |
| Woodworth, LA | Baltimore, MD | Colorado Springs, CO |
| jringalls@iinfosec.com | iitsec@coda.ee | judson.dressler@afacademy.af.edu |

## ABSTRACT

Modeling and analysis of cybersecurity data is necessary to provide education, timely analysis and reporting on threat and risk events, and leadership decision support. Current solutions suffer due to bottlenecks in how human analysts ingest, analyze, and react to the ever-growing amount of data. This is due in part to how data is presented to humans using legacy 2-Dimensional (2D) display systems and spreadsheet-formatted data output.

The advent of consumer-grade head mounted displays, that can immerse the user in stereoscopically perceivable multi-dimensional environments, including Mixed Reality (MR), Virtual Reality (VR) and some Augmented Reality (AR) headsets (xR), present an opportunity to advance cybersecurity data visualization into multiple new dimensions of capability. There are many benefits of this approach, however in this paper we will focus on three:

Immersive visualizations that use an intuitive User Interface (UI) to present multi-dimensional data visualizations that lower the technical barrier to entry for human analysts. These visualizations can be used to instruct and educate analysts in understanding complex and complicated systems that they need to work with.

Multi-Dimensional cybersecurity data visualizations allow trained cyber analysts to perform threat hunting and detection of suspicious or malicious activity that otherwise may bypass Machine Learning, heuristic, or signature-based detection capabilities that are currently used to identify this activity.

Finally, multi-dimensional cybersecurity data visualizations allow leaders who are not cyber domain experts to understand the context of a given event, using intuitive models of the event and a timeline of activity that can be used to "step-through" any precursor or other activity.

## ABOUT THE AUTHORS

**Jason Ingalls** is the Founder & CEO of Ingalls Information Security LLC, and inventor of Viewpoint, a patented cybersecurity data visualization technology.

**Kaur Kullman, PhD** is a cybersecurity researcher at the University of Maryland, Baltimore County, and the inventor of the Virtual Data Explorer cybersecurity data visualization tool.

**Colonel Judson Dressler, PhD** is the Permanent Professor and Head of Computer and Cyber Sciences at the U.S. Air Force Academy.

# Visualizing Cybersecurity Data for Detection and Assistance in Cyber Operations

**Jason Ingalls**
**Ingalls Information Security LLC**
**Woodworth, LA**
**jringalls@iinfosec.com**

**Kaur Kullman, PhD**
**University of Maryland, Baltimore County**
**Baltimore, MD**
**iitsec@coda.ee**

**Colonel Judson Dressler, PhD**
**United States Air Force Academy**
**Colorado Springs, CO**
**judson.dressler@afacademy.af.edu**

## INTRODUCTION

To maintain "availability, integrity, authentication, confidentiality, and nonrepudiation" of an information system, the personnel responsible for this task must maintain actionable situational awareness and situational understanding of the system. The larger and more complicated such a system becomes, the more vulnerabilities such a system will have. The longer that system remains in use, the higher the likelihood for those vulnerabilities to be discovered and (ab)used.

Therefore, the success of the task to monitor and protect such a system, and respond to incidents involving its components will (amongst other conditions) rely on

- that system's maintainers, operators and protectors' ability to have actionable mental models of that system, that are up to date with its growth and change over time, and;
- timely access to telemetry, logs, configurations, topology changes, and other relevant data collected from the to-be-protected system.

The amount of telemetry and logs generated by the networked devices, and applications running in these systems is usually overwhelming. However, analyzing this data to maintain actionable situational understanding is critical for protecting organizations and their data in such systems. Traditional visualizations and spreadsheet-based textual tools may work for seasoned analysts who are familiar with, and have actionable mental models of that system, but could have too steep of a learning curve for new team-members, outsourced help, and others.

To address this challenge, several organizations are exploring the potential of immersive data visualizations, using virtual and augmented reality headsets. These tools offer the potential to present complex and complicated (Traeber-Burdin & Varga, 2022) cybersecurity datasets in layouts that would be intuitive for its users, allowing analysts to identify patterns, relationships, and allow them to detect potential or investigate realized threats more easily.

In this whitepaper, we will explore the benefits of immersive data visualization for cybersecurity analysts and their education, discuss the technical requirements of such tools, potential limitations, and implications for future research, development, and deployment.

## CHALLENGES WITH CONTEMPORARY METHODS OF DATA ANALYSIS

The quantity of logs, telemetry, and various other data collected from networked devices and devices that are running these networks is increasing steadily, if not exponentially due to our society's growing dependence on interconnected information technology (Kaisler, Armour, Espinosa, & Money, 2014). This collected data is in turn instrumental for the protection of those interconnected devices that our society relies on. Cyber Defense Analysts, Cyber Defense Incident Responders and Network Operations Specialists (designated as PR-CDA-001, PR-CIR-001, OM-NET-001 respectively and bearing responsibilities for tasks identified in (NIST, Applied Cybersecurity Division, National Initiative for Cybersecurity Education (NICE), 2018)), referred to as Subject Matter Experts (SME) from here onwards, are working in Security and Network Operations Centers (SOC/NOC) to maintain Cyber Defense Situational Awareness (CDSA) and Situational Understanding (CDSU) to protect the interconnected systems, and our society.

More data does not in itself yield better information. To make sense of the collected data in a timely manner to maintain actionable CDSA & CDSU, SMEs need to be knowledgeable of the system they need to protect and be able to augment the expected but ever-changing baseline of their system with relevant and fresh data. The combination of baseline and fresh data transform into actionable CDSA, which will then feed into timely CDSU.

To fulfill that task, SMEs usually combine the output from command line tools with dashboard visualizations that allow them to interact (e.g., filter, drill, etc. (Shneiderman, 1996)) with the data depicted as charts and graphs. The dashboards that are deployed in a NOC/SOC usually provide an array of two-dimensional (2D) graphs and charts in addition to alphanumeric formats that summarize different types of data, such as system logs, network traffic, threat feeds, and so on. However, because that data is often multidimensional, the entity responsible for designing a dashboard must solve the problem of how to translate multidimensional data (for example: the topology of an interconnected system) into textual and/or 2D visualizations on flat screens (Healey, Hao, & Hutchinson, 2014).

Given that the datasets are usually large and inherently multidimensional, and that a considerable part of the human population is naturally better at interpreting visual than textual information, what visualization tools, in addition to contemporary flat-screen data visualizations would be useful in training new team members, and ease in reporting of the CDSA and CDSU?

## ADVANTAGES OF IMMERSIVE DATA VISUALIZATION FOR DATA ANALYSIS

Mixed and Virtual Reality headsets with high-quality resolution and interactive usability for providing SMEs with customized, interactive, stereoscopically perceivable, multidimensional data visualizations (ISPMDV, or immersive visualizations), may enhance SMEs capability to understand the state of their systems in ways that flat displays with either text, 2D, or 3D visualizations cannot afford; provided that the visualizations of their data is aligned with SMEs internalized representations (mental models) of their data (Kullman, Buchanan, Komlodi, & Engel, 2020).

Such visualizations could be deployed in multi-user settings (localized or remote) to facilitate inter- or intra-group collaboration, knowledge transfer, training, but also in providing CDSU to the next command tier. In the latter case, the ISPMDV layouts used for reporting upstream may need to differ from the ones used for maintaining CDSA: visualizations of complex and complicated systems (computer network topology) must be useful for SMEs, but do not have to make sense to others who do not regularly work with the dataset that those ISPMDV layouts were created for.

To compensate for the sheer amount of data that is collected with or created by monitoring tools, heuristic and other algorithmic methods are often used to filter and prioritize data to present only the information that these tools find most relevant. While this approach can help manage the volume of data, it also increases the risk of missing suspicious or malicious activity. Essentially, analysts are only seeing data that the best detection capabilities available for them can detect, which may not be enough to identify all potential threats. Instead of relying only on one or the other, the input from algorithmic methods should be merged into ISPMDV that align with SME's mental models of the system they need to protect.

The lack of a standardized model to depict relationships between actors, systems, and activity means that there is no common analytic process that can be quickly and efficiently learned without having to first learn underlying technologies and information technology concepts. This can make it difficult for new analysts to get up to speed quickly and is also challenging for different teams or organizations to share information and collaborate effectively. Multi-user, collaborative ISPMDV could be helpful in such scenarios for transferring knowledge from seasoned SMEs to (new) team members or outsiders, using an immersive visualization of the functional topology of the complicated (or complex) system that they will be working with.

## IMPLEMENTING MULTI-DIMENSIONAL DATA VISUALIZATION

The implementation of a multi-dimensional cybersecurity data visualization tool can be accomplished using a client-server architecture where

a) the server stores, processes, and communicates data between storage (or connected data source), pre-processes the data before sending it to
b) the client, which immerses the user in that data as ISPMDV.

### Application Server

At the heart of the system is the application server. The server is responsible for retrieving network traffic metadata and relevant discrete data from individual computer hosts and connections in the monitored network. The server processes this data by constructing a graph data structure, a common data organization method that represents networks of interconnected nodes and edges.

To provide context, imagine the network traffic metadata as data about all the interactions that are happening in the network, and the discrete data as separate pieces of information that provide additional details about these interactions. The graph data structure is like a detailed map of a city, with each node representing a building (a computer host) and each edge representing a road connecting these buildings (the network traffic).

The server then embeds additional layers of information within the graph data structure. These layers are derived from the discrete data and provide more in-depth insights about the individual computer hosts and connections in the network. The server's operation can be broken down into three main modules:

- A *Retrieval Engine Module* that retrieves the network traffic metadata and relevant discrete data.
- A *Graph Generator Module* that processes the retrieved metadata and builds the graph data structure.
- An *Overlay Generator Module* that works with the Graph Generator to embed the additional information into the graph data structure.

### Application Client

An immersive visualization client generally will require significant use of Software Development Kits (SDK) and libraries that would support the various hardware display systems available. Several game engines exist that allow development of applications using a common framework or platform. Unity 3D is an example of a game engine that can be leveraged to produce such an application.

Regardless of the engine or framework a client is built in, a client application would involve a few key components.

### Communicator Module

The Communicator Module is responsible for handling communication between the server and the client. In Unity, you could accomplish this by implementing a networking library, like Unity's built-in Networking API, or a third-party library like Photon. The module would handle sending and receiving data, possibly in the form of serialized JSON or XML, representing the graph data structure from the server.

### Data Visualization Layout Modules

These modules dictate the arrangement and movement of nodes and connections in the immersive visualization environment. Unity's built-in physics engine could be used to model natural movement and interaction between nodes. For layout, you could use a force-directed graph drawing algorithm, which simulates attraction between connected nodes and repulsion between all nodes, leading to an aesthetically pleasing and intuitive layout. Alternatively, you could use a hierarchical or grid-based layout for more structured visualizations. Further, Unity's Transform component would be used to move and position nodes in 3D space, while the LineRenderer or a custom shader could be used to draw connections between nodes.

### Mixed/Virtual/Augmented Reality Graphical User Interface (xRGUI)

An xRGUI displays the ISPMDV immersive visualization and provides user interaction capabilities. Unity has extensive support for xR development, and depending on the hardware used (e.g., Oculus/Meta, Microsoft HoloLens, Magic Leap, and various mobile devices with AR capabilities), it provides unified integration with that hardware’s SDK (e.g., OpenXR, Windows Mixed Reality, ARCore/ARKit). Users can interact with the visualization using their gaze, gestures, controllers, or other input devices provided or supported by the xR platform. Unity's Input System can be used to handle user interactions by defining what happens when a user looks at, clicks on, or hovers over a node.

For the visual aspect, Unity's Shader Graph or Visual Effect Graph could be used to create custom shaders for nodes and connections, allowing you to create a wide variety of visual effects. For example, a SME might want to color nodes based on their overlay data, or change their size based on their traffic volume.

## EXEMPLARY DATA VISUALIZATION TOOLS

Authors are aware of several tools that are in development for visualizing functional or logical topologies of computer networks that rely on ISPMDV or immersive visualizations to provide its users with useful environments for analyzing cybersecurity related datasets. In this paper we focus on two of such systems: Viewpoint and Virtual Data Explorer.

### Viewpoint

The Viewpoint cybersecurity data visualization tool has been in development since 2015 (United States Patent No. US10965561B2, 2021). Viewpoint provides users with a way to view all available network, host, and application log data in an immersive visualization environment that is configured for display using standard computer monitors and Mixed Reality display devices such as the HTC Vive and Oculus Rift. Network flow data, commonly called NetFlow, is visualized as connections with animated and colorized graphics, so that users can see the direction of network traffic from a client to a server as well as the protocol used. In addition to NetFlow visualization, Intrusion Detection System (IDS) data, Endpoint Detection & Response (EDR) data, system log data, and forensic agent data is available for review by selecting a specific host.

Viewpoint presents two specific types of data to users: metadata and discrete data. Metadata is defined as data about data (e.g., NetFlow data, counts of IDS alerts, EDR alerts, system, and application logs, etc.), and discrete data is defined as specific information regarding a specific asset or actor (e.g., EDR data about a specific file, user account, etc.). Viewpoint uses metadata to provide a background visualization that communicates context (e.g., how many logs are on a given system, how many connections the system makes to other systems, etc.). This allows users to see the general activity of a computer network in a way that provides context and situational awareness, while at the same time providing the ability to review specific information as part of a threat hunting exercise.

A screenshot of Viewpoint can be seen in Figure 1, which shows multiple systems communicating, along with other metadata describing the amount of logs and other information for each host.

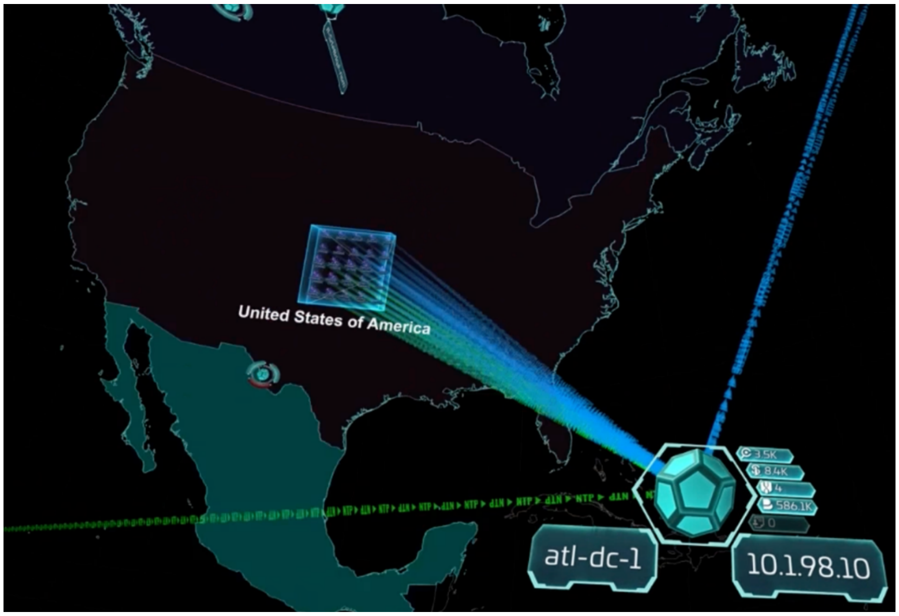

**Figure 1. Viewpoint screenshot**

Viewpoint remains in development and testing and has been used with live Security Operations Center (SOC) data to perform threat hunting and IDS alert adjudication.

For more information, please see: iinfosec.com/viewpoint

### Virtual Data Explorer

Virtual Data Explorer (VDE) has been in development since 2014. VDE allows its user to customize visualization layouts created with Mental Model Mapping Method for Cybersecurity (M4C, (Kullman, Buchanan, Komlodi, & Engel, 2020)) using two textual configuration files, which will be parsed by VDE Server (VDES) and visualized by VDE Client (VDEC).

VDE functionality is decoupled to Server and Client components to pre-process the incoming data in a more powerful environment (than a standalone xR headset) before its visualized in an xR headset. VDES also acts as a multi-user relay to synchronize the visualizations (e.g., grabbed objects position in connected users' views) between connected user sessions.

Thread-safe messaging is used extensively for asynchronous data processing, browser-based User Interface actions etc., but most importantly for keeping the Client component (VDEC) visualization in sync with (changes in) incoming data (example: response to a Moloch/Arkime query).

### Architecture

VDES has 7 main components:

1. Core is responsible for starting up the various VDES components, according to command line parameters and configuration file(s).
2. Webserver serving User Interface for monitoring and controlling VDES and connected VDE Clients' behavior.
3. WebSocket server facilitating communication with UI served by VDES Webserver and VDE Client sessions.
4. Browser Extension for enabling inter-tab communication in Chromium-based-browsers to monitor for query results in supported tools (ex. Moloch/Arkime) and transfer these results to VDES via the SignalR WebSocket.
5. Entity Templates created from topology configurations that are used while processing incoming data.
6. Data Processor that parses the incoming query response according to the currently active configuration.
7. Messenger component, relaying communications between VDES components and threads.

### Virtual Data Explorer Client

VDEC was created using Unity 3D, to facilitate deployment to various xR HMD's. Unity 3D was chosen for it's (by then) aggressive development cycle (facilitating various xR integration frameworks, that were being released by hardware manufacturers), it's use of C# as (one of) the scripting language (as opposed to C++ in Unreal Engine), and for the fact that the team I was going to work with at the US Army Research Lab was already using Unity 3D for their other VR projects. Over time this has proven to be a correct decision, as all relevant HMDs that have been released since 2015 have had Unity 3D support since early releases of their SDKs.

### Virtual or Mixed Reality

Although VDE was initially developed with Virtual Reality headsets (Oculus Rift DK2 and later CV1 with Oculus Touch), its interaction components were always kept modular so that once mixed reality headsets such as the Meta 2, Magic Leap, and HoloLens became available, their support could be integrated into the same codebase. The underlying expectation for preferring Mixed Reality to Virtual Reality is the user's ability to combine stereoscopically perceivable data visualizations rendered by an MR headset with relevant textual information represented by other sources in the user's physical environment (SIEM, dashboard, or another tool), most likely from flat screens. This requirement was identified from early user feedback: trying to input text or define / refine data queries while in VR would be vastly inferior to the textual interfaces that users are already accustomed to operating while using conventional applications on a flat screen for data analysis. Hence, rather than spend time on inventing a three-dimensional data-entry solutions for xR, it was decided to focus on creating and improving stereoscopically perceivable data layouts and letting users use their existing tools to control the selection of data that is then fed to the visualization.

A major advantage provided by the VR environment, relative to MR, is that VR allows users to move (fly) around in a larger scale (overview) visualization of a dataset while becoming familiar with its layout(s) and/or while collaborating with others. However, once the user is familiar with the structure of their dataset, changing their position (by teleporting or flying in VR space) becomes less beneficial over time. Accordingly, as commodity MR devices became sufficiently performant, they were prioritized for development - first, the Meta 2, later followed by support for the Magic Leap and HoloLens.

**User Interaction**

As an example, in VDE user can:

1. point to select a visual representation of a data-object - a node (ex. a cube or sphere) or an edge - with a "laser" or dominant hand's index finger of either the virtual rendering of the hand or users real hand tracking results (in case of Oculus Quest and MR headsets). Once selected, detailed information about the selected object (node or edge) is shown on a line of text rendered next to user's hand, (Shneiderman Task Level 4).
2. grab (or pinch) nodes and move (or throw) these around to better perceive its relations by observing the edges that are originating or terminating in that node: humans perceive the terminal locations of moving lines better than that of static ones, (Shneiderman Task Levels 3, 5).
3. control data visualization layout's properties (shapes, curvature, etc.) with controller's analog sensors, (Shneiderman Task Levels 1, 5).
4. gesture with non-dominant hand to trigger various functionalities. For example: starfish – toggle the HUD; pinch both hands – scale the visualization; fist – toggle edges; etc.
5. Once user moves closer to a part of the visualization that might be of interest, textual labels are shown for upper tier groups first, while the rectangular representations of these groups are disappeared as the user gets closer, to enable focusing on the subgroups inside, and then the nodes with their IP addresses as labels. To convey the changes in visualization as the user moves, screenshots are provided sequentially, from upper left to right. Please see VDEC behavior in MR on Figure 2 and in VR on Figure 3.

In addition to active gestures and hand recognition, the user position and gaze (instead of just their head direction, if available) are used to decide which visualization sub-groups to focus on, to enable textual labels, to hide enclosures, to enable update routines, colliders, etc. (Shneiderman Task Levels 2, 3, 4, 5, 7). Therefore, depending on user's direction and location amongst the visualization components and on the user's gaze (if eye-tracking is available), a visualization's details are either visible or hidden, and if visible, then either interactive or not.

The reasons for such a behavior are threefold:

1. Exposing the user to too many visual representations of the data objects will overwhelm them, even if occlusion is not a concern.
2. Having too many active objects may overwhelm the GPU/CPU of a standalone MR/VR headset - or even a computer rendering into a VR headset - due to the computational costs of colliders, joints, or other physics.
3. By adjusting their location (and gaze), the user can:
    a. Gain Overview of the entire dataset (Shneiderman Task Level 1),
    b. Zoom on an item or subsets of items (Shneiderman Task Level 2),
    c. Filter irrelevant items (Shneiderman Task Level 3),
    d. Get details-on-demand for an item or subset of items (Shneiderman Task Level 4),
    e. Relate between items or subsets of items. (Shneiderman Task Level 5).

Please see (Kullman & Engel, User Interactions in Virtual Data Explorer, 2022) for more details. For descriptions of Shneiderman Task Levels referred above, please see (Shneiderman, 1996).

For more information on VDE (Kullman, Interactive Stereoscopically Perceivable Multidimensional Data Visualizations for Cybersecurity, 2023) please visit coda.ee.

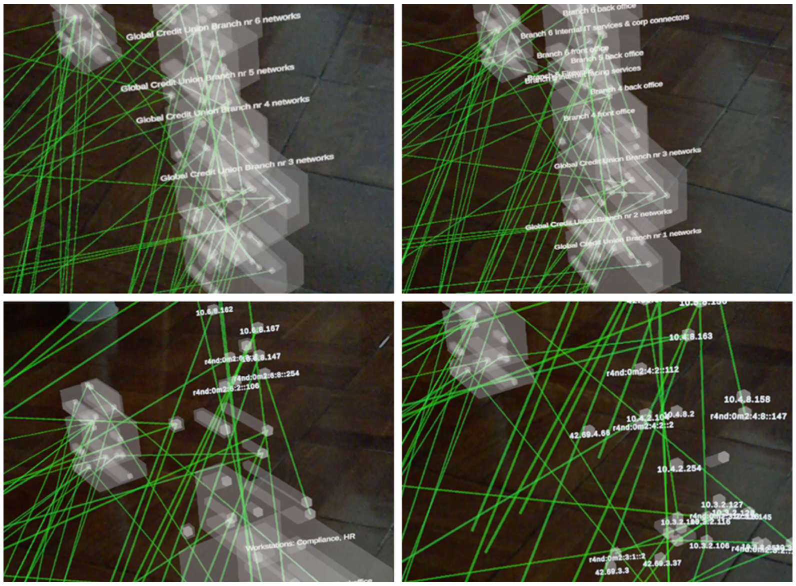
Global Credit Union Branch nr 6 networks
Global Credit Union Branch nr 5 networks
Global Credit Union Branch nr 4 networks
Global Credit Union Branch nr 3 networks
Branch 6 back office
Branch 6 Internal IT services & corp connectors
Branch 6 front office
Branch 5 back office
Branch 4 back office
Branch 4 front office
Global Credit Union Branch nr 2 networks
Global Credit Union Branch nr 1 networks
Workstations: Compliance, HR
10.6.8.162
10.4.8.163
10.4.8.158
10.4.2.254
10.3.2.127
42.69.4.68
42.69.3.37
42.69.3.3

**Figure 2. VDEC in MR**

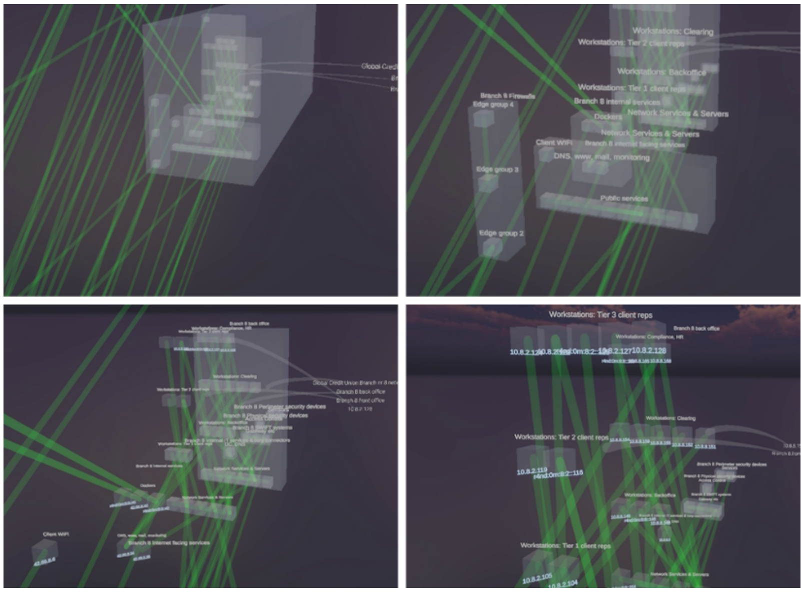
Workstations: Clearing
Workstations: Tier 2 client reps
Workstations: Backoffice
Workstations: Tier 1 client reps
Branch 8 Firewalls
Edge group 4
Branch 8 internal services
Dockers
Network Services & Servers
Client WiFi
DNS, www, mail, monitoring
Edge group 3
Public services
Edge group 2
Workstations: Tier 3 client reps
Branch 8 back office
Workstations: Compliance, HR

**Figure 3. VDEC in VR**

## LESSONS LEARNED FROM IMPLEMENTATION

The technical requirements for displaying cybersecurity data in an immersive visualization environment using xR headsets can be complex and vary depending on the specific use case. However, some general requirements include:

### Data Collection and Ingestion

To display cybersecurity data in an immersive visualization environment, the data must first be collected and ingested into a compatible format. This may involve integrating data from multiple sources into a single database or process at runtime. Dressler, Bowen, Moody and Koepke proposed six classes of information should be taken into account in order to establish operationally relevant situational awareness: the threat environment, anomalous activity, vulnerabilities, key cyber terrain, operational readiness, and ongoing operations in and out of the cyber domain

(Dressler, Bowen, Moody, & Koepke, 2014). The data must also be formatted and optimized for use in immersive visualization environments, such as by converting it into a 3D model or mesh.

### Immersive Visualization Environment Generation

Once the data has been collected and ingested, an immersive visualization environment must be generated in which to display the data. This may involve using a 3D modeling tool to prepare a virtual representation of the network or system that is being analyzed before running the software or implement a capability to generate such visualizations on runtime. The environment must also be optimized for use with standalone xR headsets, which may require specific technical requirements such as rendering distance, frame rate, and resolution.

### xR Headset Integration

Creating an immersive visualization environment may involve developing a custom xR application or using an existing software platform that is compatible with the chosen headset. The software must be able to display the immersive visualization environment and data in a way that is intuitive and easy to interact with, using features like hand gestures and voice commands.

Overall, the technical requirements for displaying cybersecurity data in an immersive visualization environment using xR headsets can be complex and may require a high degree of technical expertise. However, the potential benefits of this approach, including improved visualization, data analysis, and decision-making capabilities, make it a promising area for further research and development.

### Simulator Sickness

A crucial property for an immersive visualization environment is the capability to avoid users experiencing simulator sickness. Various experiments have shown that applying certain limitations to a user's ability to move in the virtual environment – limit their view and other forms of constrained navigation – will limit confusion and help prevent simulator sickness while immersed. For example, if an immersed user can only move their viewpoint either forwards or backwards in the direction of user's gaze (or head-direction), the effects of simulator sickness can be minimized or avoided altogether (Pruett, 2017).

### User Interactions

The ability to interact with data visualization while immersed, namely, to query detailed information about a visual representation of a data point using input devices is imperative. While gathering feedback from SMEs (Kullman, Ben-Asher, & Sample, Operator Impressions of 3D Visualizations for Cybersecurity Analysts, 2019) it was confirmed to be also crucial for users' immersion in the data visualization to allow them to explore and build understanding of the visualized data while querying it intuitively (Kullman & Engel, User Interactions in Virtual Data Explorer, 2022).

## TECHNICAL LIMITATIONS

While immersive visualization offers many potential benefits for cybersecurity analysis and education, there are also some potential limitations to consider. These limitations include:

### Technical Requirements

One of the primary limitations of immersive visualization is the high degree of technical expertise required to develop and implement these tools. This may include expertise in areas like 3D modeling, data processing, and software development, as well as specialized hardware and software requirements. These technical requirements can make it difficult for organizations to adopt and use immersive visualization tools, particularly smaller organizations with limited resources.

### User Adoption

Another potential limitation of immersive visualization is user adoption. While these tools can provide more intuitive and immersive ways to analyze data, they may also require users to learn new skills and adapt to new interfaces. This can be particularly challenging for organizations with large and diverse user bases, where some users may be resistant to change or may struggle to learn new tools.

### Data Overload

Immersive visualization can also lead to data overload, particularly if the visualization is not designed in a way that effectively filters and highlights the most relevant data. This can make it difficult for analysts to identify important patterns and trends and may even lead to information overload and cognitive overload.

While occlusion in VR can be addressed by measures such as transparency, transparency adds significant overhead to the rendering process. To optimize occlusion-related issues, VDE strikes a balance between the necessity of transparency of visualized objects, while adjusting the number of components currently visible (textual labels, reducing the complexity of objects that are farther from the user's viewpoint, etc.) based on the current load (measured FPS); on objects' relative positions in user's gaze (in-view, not-in-view, behind the user); and on the user's virtual distance from these objects. Such an approach to semantic zooming proves a natural user experience, visually akin to the semantic zooming techniques used in online maps which smoothly but dramatically change the extent of detail as a function of zoom level (showing only major highways or the smallest of roads, toggling the visibility of street names and point of interest markers).

### Limited Interactivity

While immersive visualization can provide a more immersive experience, it may also have limitations in terms of interactivity. For example, it may be more difficult to manipulate and explore data in an immersive visualization environment compared to a 2D environment, particularly if the interface is not designed in a way that is intuitive and easy to use. This may limit the ability of analysts to interact with and explore data in real time.

In contrast to fully immersed visualization (in VR), controlling VDE's server and client behavior, including data selection and transfer, turned out to be more convenient when done in combination with the VDE Server's web-based interface and with existing conventional tools on a flat screen. For example, in the case of cybersecurity related datasets, the data source could be a SIEM, log-correlation, net flow, PCAP analyzing environment or something else (see Data Collection and Ingestion).

## CONCLUSION

In conclusion, immersive visualization represents a powerful new tool for cybersecurity analysis and education. By providing an immersive and intuitive way to visualize data, these tools have the potential to transform the way that analysts, decision-makers, and students approach cybersecurity.

Throughout this paper, we have explored the benefits of immersive visualization for cybersecurity analysis and education, as well as the technical requirements, potential limitations, and implications for future research and development. We have seen how immersive visualization can enable more effective threat hunting and detection, lower the technical barrier to entry for cybersecurity education and training, and provide decision-makers with a more comprehensive and intuitive view of their organization's cybersecurity posture.

While there are potential limitations to consider, including the technical expertise required to develop and implement these tools, the potential for data overload, and the potential cost of implementing these tools, the benefits of immersive visualization make it a promising area for further research and development. By continuing to explore the potential of immersive visualization, we can work to create a more secure and resilient cyberspace for everyone.

As we move forward, it will be important to continue to research and develop new tools and techniques that leverage the power of immersive visualization. This will require collaboration across disciplines, including cybersecurity, data science, and human-computer interaction, as well as a commitment to ongoing education and training. By working together, we can unlock the full potential of immersive visualization for cybersecurity analysis and education.

## REFERENCES

Dressler, J. C., Bowen, C. L., Moody, W., & Koepke, J. (2014). Operational data classes for establishing situational awareness in cyberspace. *6th International Conference On Cyber Conflict.* Tallinn.

Healey, C. G., Hao, L., & Hutchinson, S. E. (2014). Visualizations and Analysts. In A. Kott, C. Wang, & R. F. Erbacher, *Cyber Defense and Situational Awareness* (pp. 145-165). Springer.

Ingalls, J., Richards, A., Perinelli, E., Piccinelli, N., & Arena, R. (2021, 03 30). *United States Patent No. US10965561B2.* Retrieved from https://patents.google.com/patent/US10965561B2

Kaisler, S., Armour, F., Espinosa, A. J., & Money, W. (2014). Big Data: Issues and Challenges Moving Forward. *2014 47th Hawaii International Conference on System Sciences.* Wailea.

Kullman, K. (2023). *Interactive Stereoscopically Perceivable Multidimensional Data Visualizations for Cybersecurity.* Tallinn: Tallinn University of Technology.

Kullman, K., & Engel, D. (2022). User Interactions in Virtual Data Explorer. In D. F. Schmorrow (Ed.), *International Conference on Human-Computer Interaction. 13310*, pp. 333–347. Springer. doi:10.1007/978-3-031-05457-0_26

Kullman, K., Ben-Asher, N., & Sample, C. (2019). Operator Impressions of 3D Visualizations for Cybersecurity Analysts. *18th European Conference on Cyber Warfare and Security.* Coimbra, Portugal.

Kullman, K., Buchanan, L., Komlodi, A., & Engel, D. (2020). Mental Model Mapping Method for Cybersecurity. *HCI for Cybersecurity, Privacy and Trust* (pp. 458-470). Tallinn: Springer International Publishing. doi:10.1007/978-3-030-50309-3_30

NIST, Applied Cybersecurity Division, National Initiative for Cybersecurity Education (NICE). (2018, 01 18). *Reference Spreadsheet for the NICE Framework, NIST SP 800-181.* Retrieved 01 2020, from https://www.nist.gov/itl/applied-cybersecurity/nice/nice-cybersecurity-workforce-framework-resource-center/current

Pruett, C. (2017, 06 09). *Lessons from the frontlines modern VR design patterns.* Retrieved from Oculus for Developers: https://developer.oculus.com/blog/lessons-from-the-frontlines-modern-vr-design-patterns/

Shneiderman, B. (1996). The eyes have it: a task by data type taxonomy for information visualizations. *Proceedings 1996 IEEE Symposium on Visual Languages.* Boulder, CO, USA, USA: IEEE. doi:10.1109/VL.1996.545307

Traeber-Burdin, S., & Varga, M. (2022). How does Systems Thinking support the Understanding of Complex Situations? *2022 IEEE International Symposium on Systems Engineering (ISSE)* (pp. 1-7). Vienna: IEEE. doi:10.1109/ISSE54508.2022.10005449