Visualization for Cyber Complex Systems: Application, Issues and Future Work

Georgi Nikolov
Research Unit for Cyber Defence
Royal Military Academy
Brussels, Belgium
g.nikolov@cylab.be

Margaret Varga
Department of Biology
University of Oxford
Oxford, United Kingdom
margaret.varga@biology.ox.ac.uk

Susan Träber-Burdin
Human Machine Systems
Fraunhofer FKIE
Bonn, Germany
traeber-burdin@posteo.de

Carsten Winkelholz
Human Machine Systems
Fraunhofer FKIE
Bonn, Germany
carsten.winkelholz@fkie.fraunhofer.de

***Abstract*—Technology advances at a rapid pace; new hardware being developed offers the opportunity to create even more intricate networks of devices, interconnected not only locally, but also on a global scale, offering increasingly fast processing and generating immeasurable amounts of data. New software takes advantages of the leaps in hardware to perform even more intricate computations and facilitating our daily lives, in the work and home environment. Unfortunately, when an issue arises in the network, it is difficult to immediately pinpoint the origin and find an appropriate solution readily. The infrastructures that we have built have become so complicated and so interconnected, that they pass the threshold from complicated into the realm of complex systems. Feedback loops, non-linearity, and emergent behavior are underlying characteristics of complex systems. These characteristics contribute to the unpredictable and often intricate behaviors that arise from the interactions between different elements in the system. Complex systems are thus difficult to manage without an in-depth knowledge of the underlying components and their interactions - where the whole is greater than the sum of its parts. To aid in this task, new ways to visualize the system need to be developed to represent them accordingly and aid cyber operators and analysts to quickly and correctly detect and identify problems and find solutions. In this paper, we offer a detailed explanation on what Cyber complex systems are, the difficulties encountered and how Visual Analytics and visualization can resolve some of the issues. Examples of visual representations of cyber complex systems will be discussed, together with techniques used for their evaluations in terms of their usefulness and usability. Finally, a brief overview of possible future advancements that can support better understanding and management of complex systems will be discussed.**

***Index Terms*—Visual Analytics, Visualization, Complex Systems, Cyber, Situation Awareness, Evaluation**

I. Introduction

Today's society is overwhelmingly dependent on cyber systems ranging from individual daily activities to the functioning of governments and industry. This dependency exhibits not only the ubiquitous use of digital technology but also its deep-rooted integration into practically all aspects of life. Indeed, our daily lives are governed by many different systems, be that the environment, our community or work organization, such as social interaction (email, social media), business communication (video conferencing), e-commerce (digital retail), financial services (online banking), critical infrastructure (energy grid, transportation network), public services (tax filing), healthcare (digital health records, telemedicine), etc. We benefit tremendously from these systems, but we also face challenges that require careful management to maintain a stable and safe environment.

These systems are not always easy to understand, consisting of a structure composed of any number of autonomous parts, which interact constantly with each other and produce unexpected or unpredictable results. This inherent complexity can lead to mismanagement, sometimes producing catastrophic mistakes. There are three types of systems, namely simple, complicated and complex systems [1]. In this paper we will focus on complicated and complex systems. Often issues arise when a system, which exhibits all characteristics of a Complex System is treated as though it is a Complicated one, or vice versa. A clear picture of a complicated system can be gleamed-each piece of hardware is attributed a specific role, from the hard disk which stores data, the different cables transporting electricity, to the motherboard functioning as the backbone of the machine, aiding the communication between the different components. However, applying the same methodology to the understanding of a complex system such as the Internet falls quickly short. We can separate the obvious distinct components, such as the computers, routers and humans, but their interaction is far more nebulous. The social aspect of communication between users, the vastly different types of data and even the geographical aspect have a significant influence on the system and there is no one correct way of representing these factors to better comprehend them. An in-depth description and explanation of complex systems will be presented in section II.

As the Cyber domain can be equated to a complex system, much effort has been spent on understanding and representing how different network components interact to improve the ability to monitor and detect anomalous or malicious behavior. With the rapid advancement of technology, adversaries evolve and present new ever greater threats. Cyber attacks have multiplied and their sophistication has increased many folds. To combat the ever present and increasing threat of these attacks, new detection technologies and methodologies have been developed, but it is an ever evolving arms race between defenders and attackers.

One way to counter cyber attacks is through the use of technologies such as Intrusion Detection Systems (IDS), Honeypots or Behavior-based analysis. More often than not, all the information is gathered and then relayed to a Security Information and Event Management (SIEM) system for analysis. The large quantities of data generated daily and fed into a SIEM are often impossible to handle without the use of visualization techniques to transform these inherently non-visual high dimensional data into an intuitive visual form to provide an effective means to explore, analyze and process its meaning. The clear and appropriate visual representation of the information can greatly aid analysts and enhance their Cyber Situation Awareness (CSA). The application of CSA to complex systems will be discussed in section III. Following that, the Visual Analytics techniques applied for the enhancement of the various stages of CSA will be presented in section IV. Finally, we will discuss what are the methodologies and techniques used for the evaluation of such specialized visualizations in section V.

II. Complex Systems

In order to manage a system it is important to establish the differences between a complicated system from a complex system as the approaches applied to manage them are radically different. Complicated systems are deterministic. Their formal and functional structure, i.e. their units and relationships, can in principle be fully analyzed - even if only with certain expertise. Based on the knowledge of the complete structure and functioning of the systems, their behavior can be fully predicted and regulated or optimized accordingly. This applies, for example, to a database management system with many different functions and features, consisting of many interacting modules with well-defined interfaces.

In contrast, complex systems are non-deterministic. They have an open functional structure, i.e. they interact with subsystems in their environment, which in turn interact with subsystems in their environment. This results in a network of feedback loops that lead to highly dynamic, non-linear behavior and the emergence of new system properties or functions that cannot be understood if entities are analyzed in isolation, i.e. the whole is more than the sum of its parts. Furthermore, the formal and functional structure can change on its own due to changes within and outside the system (self-organization). Due to this dynamic, open and highly interdependent nature, the structure of complex systems cannot be fully analyzed and therefore the behavior cannot be fully predicted. This makes the management of complex systems much more difficult and requires a shift from reductionist to systems thinking [2] and the application of management and problem-solving methods based on this [3]. The internet is an example of a complex system. It is composed of innumerable individual networks, devices, and protocols interacting in ways that can produce unpredictable outcomes, such as network traffic patterns. or the spread of misinformation. Furthermore, any cyber security system that involves substantial human interaction becomes inherently complex. Human behavior introduces variability and unpredictability, such as when responding to alerts, or unknowingly creating vulnerabilities.

Often in the cyber domain, complicated and complex systems interact in manners that are integral to the functionality and security of the modern digital infrastructures. These interactions can have deep implications and effects on the resilience, performance as well as vulnerability of the systems. Complicated systems such as databases, networking hardware, and software applications are designed to perform specific tasks with a high degree of predictability and reliability. These systems often are the infrastructure and backbone upon which complex systems such as the Internet operate. Therefore the performance and behavior of complex systems is influenced by the functioning of the individual complicated systems / components. For example, the overall performance of the Internet relies on the correct functioning of various routers, servers, and protocols, where each is a complicated system. Interactions between complicated and complex systems often create feedback loops where the outcome / behavior of one influence the input / behavior of another. A vulnerability in a software application (a complicated system) can lead to extensive security breaches or failures in the network (a complex system), which in turn requires changes or updates in the software (feedback loops). As more complicated systems are added within complex systems, new behaviors or properties (emergence) may emerge that were not predictable by analysing the behavior of the individual components in isolation. This is particularly relevant in cyber systems when incorporating more devices and connections can result in unexpected network dynamics or introduction of new vulnerabilities. In a complex cyber environment, resilience in operation is often governed by how well the complicated systems can recover from failures / compromise and remain in operation under adverse conditions. The interactions between these systems, such as redundancy protocols (complicated systems) within network architecture (complex system), is vital in maintaining service continuity. To monitor and manage these interactions, a holistic approach that takes into account both the complicated systems (micro-level details) and the dynamics of the complex systems (macro-level) is vital [2] [3].

III. Cyber Situation Awareness

Cyber Situation Awareness (CSA) is concerned with the perception and understanding of the stability and safety status within a cyber environment, together with the ability to predict, detect, and respond to existing vulnerabilities and potential new threats effectively. This is derived from the situation awareness in other domains, such as military operations and aviation domains, where it is vital to maintain up to date understanding of an environment to support timely and effective decision making.

The CSA is governed by three distinct phases [7], as shown in Fig. 1, each play a significant role in supporting the correct management and threat mitigation.

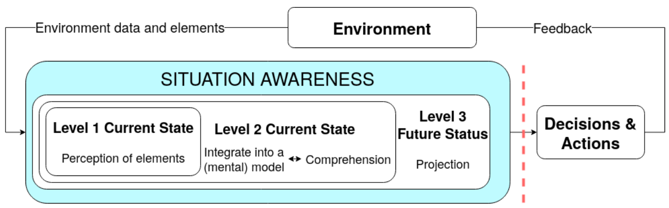

Fig. 1. Situation Awareness as defined by [7]

- **Perception of elements in the environment** : This phase refers to the capability of humans to monitor, detect cues in the environment and basic recognition. These capabilities are vital in the Cyber domain for the surveying of the different network elements (events, people, systems, etc.) and their current state. Displays such as dashboards, network graphs and heatmaps can be used to represent data in relation to network traffic, user activities, and alerts.
- **Comprehension of the current situation** : Following the initial perception phase, the observed information needs to be interpreted and possible patterns recognized and analyzed. This step will lead to better understanding of the meaning of the observed information, how the elements are interlinked and how it relates to the state and security of the network. In essence, this phase translates raw data into a contextual form that can support informed decision making. Interactive exploration and analysis are required so that users can interact with data, e.g. zooming, panning, drilling down, or filtering information of interests to explore, analyse and understand the data to gain understanding and insight into events.
- **Projection of future status** : By correctly understanding the data and the observed patterns and behavior, a clear model of the evolution of the system can be constructed, future impacts can be assessed and mitigation techniques can be put in place. The projection phase thus encompasses predicting future states of the cyber environment based on the information gathered and comprehended in the previous two stages. Visualization tools can incorporate predictive analytics to exploit historical data / past events to predict potential future threats - graphs and trend lines depicting previous attack patterns to support prediction of what might/could happen.

Indeed, effective CSA is essential to support informed decision making in maintaining the stability and safeguarding the functioning of systems and networks. It relies on advanced tools and technologies such as intrusion detection systems, SIEM systems, and threat intelligence platforms to gather, analyze, and interpret data. The aim is thus to maintain a high level of awareness that facilitates proactive responses to cyber threats, mitigating damage and reinforcing overall security.

IV. Visual Analytics and Visualization

Visual Analytics and visualization techniques can greatly aid in the different stages of CSA. Sadly, many problems arise when we try to apply the theory into practice. The nature of the cyber environment is so vast and can even be considered border-less, it is often difficult to determine which parts of the environment the analyst should focus their attention on. This stems from the position the analyst is in, i.e.- they are passive observers of the events in the system, using only the information gathered by the different sensors.

Indeed, visualization and visual analytics play a vital role in enhancing Cyber Situation Awareness (CSA) by providing users with intuitive and interactive techniques / systems to perceive, understand, and react to the ever-growing complex cyber data and events [6]. These tools help users and decision-makers to interpret vast amounts of different types of data and identify potential security threats and anomalies in a timely manner.

A. Human vs System centric approach

To better understand the information gathered about the system, there are different approaches to human machine interface design can be developed and applied to address the different operational and users' needs, for example:

- human centred design (HCD) approaches, and
- system based approaches, such as the Ecological Interface Design (EID).

These two approaches provide SA in a different manner. HCD approach focuses on the users' and their tasks' needs, the users' skills, work environment and limitations, as well as their mental models [12] [13]. While, the EID focuses on the system [13] with the objective to show the complex relationships in the system to the users in a readily informative and intuitive manner. It is a user interface design particularly suitable for real-time dynamic and complex socio-technical systems [12].

The first step is defining who is the intended audience of the representation, and the objectives and tasks, which will govern the decision on human or a system centric approach [5]. The human centric approach takes into account the knowledge and expertise of the user, their work environment, their tasks, available data sources as well as their mental model and appropriate visualization approaches, designing a representation using which they can apply their knowledge in the analysis of the system. This means that the visualization needs to be highly customizable to suit the needs of the operator to conduct their tasks and offer different ways to represent the data, be that using bar, line or pie charts or more sophisticated visualizations such as Sankey. To enhance the operator's CSA, it is of great importance that they can compare, filter, correlate and dive deeper into the data [15]. To enable a fully interactive and interactive visual analysis using the human centric design approach, one must follow the taxonomy of interactive dynamics for visual analysis [4], Fig. 2 shows an example human centric dashboard. The user-centric visualization approach provides an effective means of analysing, detecting, discovering and identifying patterns, anomalies, violations and threats; as well as correlating events. The resulting intuitive visualizations are suitable for the provision of detailed information on the performance of network components, such as IPs, ports, protocols, packages, CPU load, disk and memory usages, etc.

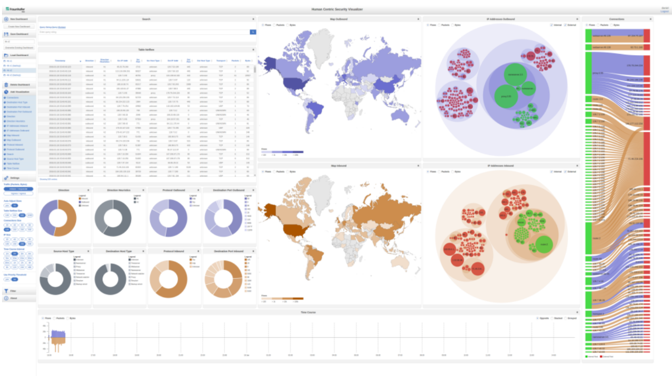

Fig. 2. Example of a Human Centric Design through a multi-visualization dashboard

Ecological Interface Design (EID) is an approach based on the idea that by understanding how a system works, users can diagnose problems and manage a system more effectively. The objective of EID is to utilize the knowledge of the system to develop interfaces that support the natural human ability to understand complex systems. It is particularly suitable in domains where the systems are complex and dynamic, such as nuclear power control and cyber situation awareness [14].

EID manages complexity by creating interfaces that reveal the underlying structure of the system to the users. It facilitates a deeper understanding of how different components of the system interact and how they contribute to the overall functioning of the system. In EID, the visualization covers both the operational states and the constraints that govern system behavior. EID supports three different types of behaviour, namely Skills (routine actions), Rules (act on pre-defined rules), and Knowledge-based (deeper understanding for more complex or new situations).

By designing interfaces that cover the three levels, EID supports operators to perform their tasks effectively across routine and unusual scenarios.

Fig. 3 depicts a logical network topology, the functionalities of the network, showing the relationships and dependencies between servers, firewalls etc. It provides a visualization that guides the users to understand the functioning of the network. Once users are familiar with the patterns of the 'normal situation', they can readily detect any changes from the normal patterns. Thus, in the EID concept, analysts can easily see the operational aspects of the network, i.e. the big picture. In this approach users can derive great insight into how the system works; but it lacks the capabilities in providing information about the reason of the issue or the events that might have led to it [16].

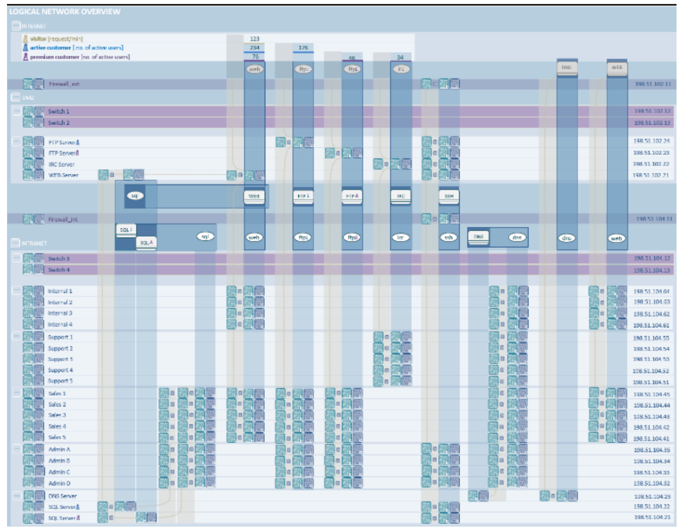

Fig. 3. Example of a System Based Approach depicting a logical network topology

The two approaches complement each other in providing awareness and information of different aspects of the network situation. Thus, integrating the HCD and EID approaches can produce a CSA interface that addresses the user needs as well as providing information on the functionality of the system, thus exploiting the benefits of both approaches.

B. Levels of visualization

A different approach to representing data is to focus on specific level of information. This can vary from a very low level, or in other words "bottom-up", to a very high level, "top-down" view. Each level has its own advantages and issues that must be considered when developing a visual representation. On the lower level, the important information is the flow of data through the network, such as Netflow or PCAPs. Here, the analyst can gain insight into specific events, their origin and their effect on specific machines. This is great for the identification of specific patterns that might be of interest for future analysis. The problem with the "bottom-up" representation, is that while focusing on specific low level of information, the operator may lose sight of the system as a whole and how one event may orchestrate a change in behavior and impact various parts of the system and thus the functioning of the whole system. The "top-down" representation of information provides the user with a good overview of the system, the flow of information and possible issues that may arise in various elements of the network. Such visualizations are great for speedy identification of points of failure, which can quickly be remedied for the continuous functionality of the system. Because of the high level nature of the overview of the system, it is also easier to discern emergent behavior in the interactions between the different elements in the system and possibly adapt to the dynamic change. As with

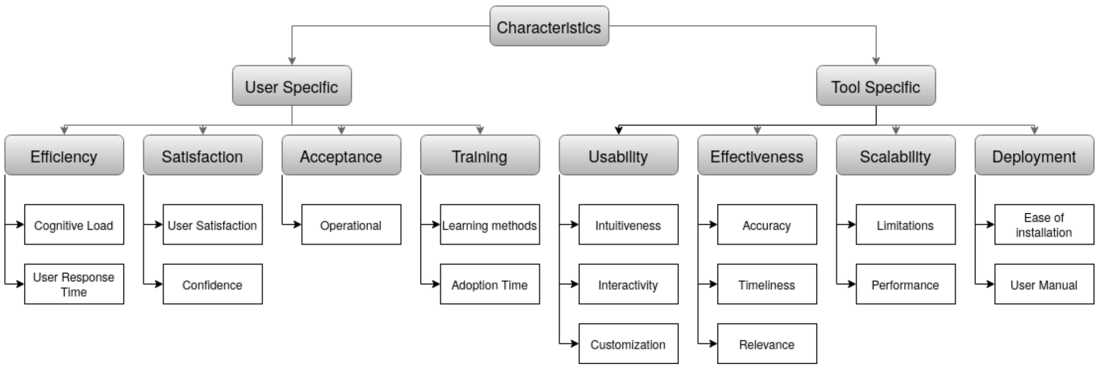

Fig. 4. Characteristics used for the evaluation of the design and implementation of visualizations

the system centric approach, an issue that arises is that it is difficult to infer what are the causes of the change in behavior and even worse, what are the specific effects on the specific elements of the system.

V. Evaluation of cyber visualizations

The dynamic nature of complex systems does not easily lend itself to one specific visualization. Indeed, much work has been done in designing and implementing various visualization tools to aid users in better identifying and understanding their environment. A systematic literature review of CSA visualizations [8] showcases different approaches for the enhancement of the CSA level of operators. One issue that arises is the difficulty of evaluating the usefulness and usability of the various visualizations and how they apply to specific situations.

Evaluating Cyber Situation Awareness (CSA) visualization involves assessing how intuitively and effectively the visual tools and interfaces facilitate the users to perceive, comprehend, and project information in a cyber environment. Evaluation is vital to ensure that the visualizations do indeed help to enhance the decision-making capabilities of those monitoring and responding to cyber threats. A selection of some of the elements and methodologies that can be considered in evaluating cyber visualizations as shown in Fig. 4. As presented, the characteristics can be broadly separated into two groups- elements pertaining to the user experience and those defining the visualization tool used. There are various ways of evaluating the aforementioned characteristics, usually pertaining to the use of diverse methods aimed at a formal assessment of a visualization design [11]. It is important to keep in mind, that the two groups are interdependent on each other. The user specific characteristics encompass how the operator interacts with the visualization tools and how the experience of the user is enhanced by them. To accomplish that, the visualizations need to be able to decrease the cognitive load of the user by presenting information in an easily comprehensible manner and lower the response time, leading to quick and appropriate reactions to what is shown on the screen. A well designed tool will lead to higher user satisfaction and confidence about their understanding of the cyber situation to make informed decisions. Furthermore, a higher confidence and satisfaction will cause a quicker operational acceptance of the tool and shorter adoption time, depending on the learning mechanism applied. All these characteristics are influenced by the visualization tool and its specificities. The user will more quickly adopt the visualization if it offers high usability regarding the intuitiveness of use, the degree of interactivity during data exploration and the level of customization available. Furthermore, the tool can be evaluated on the degree of accuracy it exhibits when displaying information, the ability of the visualization to update in real-time or near real-time to show the current status, and if it is able to provide information that is of use and relevance to the users to make informed decisions. On the technical side, it is important to also evaluate the scalability of the visualization tools and how easy it is for them to be deployed and used for operational use as well as affordability.

Alongside the evaluation of the characteristics dealing with the human-machine interaction, a need for the assessment of the level of CSA is needed. There are various methodologies that can be applied for this, each presenting their own benefits and drawbacks. Techniques such as Situation Awareness Global Assessment Technique (SAGAT) [9], Situation Awareness Rating Technique (SART) [9] and Situation Present Assessment Technique (SPAM) [10] have vastly different approaches to how the operator is questioned on their experience with the visual tools provided. More often than not, it is better to focus on a specific low level of analysis, a specific dataset or subsystem. This is done by presenting an operator, who has intimate knowledge of the network and system with the visualization to be evaluated, focusing on specific aspects that are well known and easily defined. Through an iterative process, the evaluation then can be scaled up to expand further and go to successively higher levels for a more general evaluation and estimate of the capability to detect emergent behavior in the complex system. Other techniques include: cognitive walkthroughs, heuristic evaluation, scenario based assessment [17], Key performance indicators (KPIs) [18], Controlled user experiments, analytics based evaluation and ecological validity.

VI. Future Work

The dynamic nature of complex systems has opened the door to new domains of research in the field of Visual Analytics. Designing visualizations to represent the intricate nature of complex systems is a difficult task. As presented earlier in Section IV-A, there are two major approaches for the visual representation of such systems. Our belief is that through the combination of the two, a more holistic approach can be achieved to visualize everything, without ignoring any parts of the system. The integration of both the user and system centric approaches will benefit both types of visualizations, enhancing the aspects where each of them presents deficiencies. To achieve that, specific capabilities need to be present to transition between the micro- and macro-levels seamlessly, offering the capability to not only investigate the root of problems in the system, but also the impact those can have globally. Furthermore, such visualizations can greatly help quickly identify emergent behaviors, be those benign or malicious, and estimate their short and long term effects.

To correctly evaluate the benefit of combining the two approaches, a rigorous testing needs to be performed not only on the design choices made for the visualization, but also the benefit they bring in enhancing the CSA level of the operator. To do so, we propose the use of specialized scenarios in a controlled environment, simulating various types of attacks / malfunctions in the network, positioning the origin outside and inside the system. The evaluation can be split in two parts, first evaluating the user experience and the tool capabilities, and second preparing a questionnaire to evaluate the three stages of CSA, as described in Section V. In both cases a choice needs to be made between the use of questionnaires filled in by the operator and the use of expert opinion. Indeed, relying on the subjective opinion of the users can be beneficial for the evaluation of the user experience, but does not produce reliable results for the evaluation of the CSA level. A better approach for the assessment of CSA is to use an objective methodology such as SAGAT [9] in combination with user feedback on the usefulness and usability of the visualization tool. By combining the two evaluation methods, a better overview of the advantages and disadvantages of a given visualization can be highlighted and any shortcomings can be amended.

VII. Conclusion

The rapid growth of the cyber environment provides possibilities for the development and application of new visualization techniques for management and defense against ever present threats. The inherent complexity of the cyber domain and its effect on the Situation Awareness of the people working within it can not easily be resolved. Designing visualizations with the user in mind can help take advantage of their knowledge and aid in the identification of the cause of specific issues, but lacks the capabilities to determine the impact on the system. Contrary to that, designing visualizations with the system based approach, the impact on the system can be observed, managed and mitigated, but deeper understanding of the root issues is lacking. A future solution would be to combine the two approaches to offer a multi-level view of the system, combining the micro- and macro-levels into a complete system overview, greatly benefiting the operators to apply their expertise and better detect, understand and determine future implications of emergent behaviors and patterns within the network.

Acknowledgment

This is the work of the NATO Visual Analytics for Complex Systems Research Task Group.

References

[1] R. Van der Merwe, "Collapsing the complicated/complex distinction: It's complexity all the way down." Interdisciplinary Description of Complex Systems: INDECS 21.1 (2023): 1-17.

[2] S. Träber -Burdin, M. Varga, How does Systems Thinking support the Understanding of Complex Situations? 24 October 2022 IEEE International Symposium on Systems Engineering (ISSE), DOI:10.1109/ISSE54508.2022.10005449Corpus ID: 255598986

[3] S. Träber-Burdin and M. Varga, Dealing with complex situations: towards a framework of understanding problems. 1431-1436. 10.1109/SMC53654.2022.9945102.

[4] J. Heer and B. Shneiderman, "Interactive dynamics for visual analysis,"Communications of the ACM, vol. 55, no. 4, pp. 45–54, 2012.

[5] M. Varga, C. Winkelholz, and S. Träber-Burdin, "Cyber situation awareness," NATO/OTAN (STO-MP-IST-148), 2016.

[6] M Varga, C. Winkelholz, Carsten, S. Träber-Burdin, P. Bivall and K. Kullman, Chapter 7 Cyber Situation Awareness, Exploratory Visual Analytics. 10.14339/STO-TR-IST-141, February 2023.

[7] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," in Situational awareness. Routledge, 2017, pp. 9–42.

[8] L. Jiang, et al. "Systematic literature review on cyber situational awareness visualizations." IEEE Access 10 (2022): 57525-57554.

[9] M. R. Endsley, et al. "A comparative analysis of SAGAT and SART for evaluations of situation awareness" Proceedings of the human factors and ergonomics society annual meeting. Vol. 42. No. 1. Sage CA: Los Angeles, CA: Sage Publications, 1998.

[10] M. R. Endsley. "A systematic review and meta-analysis of direct objective measures of situation awareness: a comparison of SAGAT and SPAM." Human factors 63.1 (2021): 124-150.

[11] J. Nielsen, Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Eds.), Usability Inspection Methods. John Wiley & Sons, New York, NY. 1994

[12] M. R. Endsley, B. Bolté, and D. G. Jones. Designing for situation awareness: An approach to user-centered design. CRC press, 2003.

[13] M. Varga, C. Winkelholz, and S. Träber-Burdin, S., An Exploration of User Centered and System Based Approaches to Cyber Situation Awareness, 15th IEEE Symposium on Visualization for Cyber Security (VizSec), 22nd October 2018, Berlin, Germany

[14] J. Rasmussen and K. J. Vicente. Coping with human errors through system design: Implications for ecological interface design. International Journal of Man-Machine Studies, 31, 517-534, 1989.

[15] K. Liggett and K. Kullman, Chapter 2 – Human factors considerations for visual analytics, Exploratory Visual Analytics (NATO). 10.14339/STO-TR-IST-141

[16] C. M., Burns, J. Kuo and S. Ng. Ecological interface design: a new approach for visualizing network management, Computer Networks 43, pp 369-388, Elsevier B. V., 2003.

[17] N. Looker, D. Webster, D. Russell, and J. Xu, Scenario Based Evaluation. Proceedings - 11th IEEE Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing, ISORC 2008. 148-154. 10.1109/ISORC.2008.56.

[18] D. Parmente, Key Performance Indicators Developing, Implementing, and Using Winning KPIs Third Edition, 2015, John Wiley & Sons, Inc., Hoboken, New Jersey.