ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

SecSAGE: NIST Cybersecurity Framework Visualization on SAGE2

Wudhichart Sawangphol[1], Assadarat Khurat[2] and Nasorn Niampradit[3]

ABSTRACT

Cybersecurity has been an area of great interest for an organization, given the significance of data and the increasing cybersecurity threats. The National Institute of Standards and Technology (NIST) has developed a cybersecurity framework intended for voluntary utilization by critical infrastructure owners and operators. Its primary purpose is to aid in the effective management of cybersecurity risks. This framework, similar to many other security standards, comprises a substantial volume of textual information that can be challenging to grasp comprehensively in a limited timeframe. In response to this challenge, we designed and developed an interactive visualization of the NIST Cybersecurity Framework using the SAGE2 platform. Our objective is to facilitate a better understanding of the framework. In addition, using SAGE2 enhances collaborative working. In our project, we analyze the content within the NIST document and map the framework's five core functions into a rich visualization workflow. Each function includes categories, sub-categories, and references that users can interactively explore. Our experiments show that our visualization can help participants correctly find the information about the NIST Cybersecurity Framework faster than manually finding the information in the document. For all tasks, participants can complete the tasks around 4.25 times faster than the manual method on average.

Article information:

Keywords: Data Visualization, SAGE2, High-resolution Display Wall, NIST Cybersecurity Framework

Article history:
Received: March 13, 2024
Revised: May 30, 2024
Accepted: July 25, 2024
Published: August 3, 2024
(Online)

DOI: 10.37936/ecti-cit.2024183.256034

1. INTRODUCTION

Cybersecurity is currently a critical concern for every organization, irrespective of its size, since being attacked successfully may bring significant losses to the organization, such as operational disruptions and data breaches. Security controls should be implemented to protect the organization. To choose security controls appropriately, some organizations start this task by follow *Security Standards* such as NIST Cybersecurity Framework [1], ISO/IEC 27001 [2], ISO/IEC 27002 [3], COBIT [4], NIST SP800-53 [5], and CIS Critical Security Control [6]. However, they usually contain a substantial volume of textual information, which can be challenging to grasp comprehensively within a constrained time frame. Especially for the newcomers who have not yet familiarized themselves with these standards, it will be time-consuming to acquire this knowledge.

Among these security standards, the *NIST Cybersecurity Framework* developed by the National Institute of Standards and Technology (NIST, United States) for voluntary use by critical infrastructure organizations to assist in managing risk [1] is our main focus. The version of this framework we are working on is 1.1. This framework introduces cybersecurity activities, outcomes, and relationships between several security standards. This relationship allows us to have more choices when selecting appropriate security controls. Like other security standards, the NIST Cybersecurity Framework also contains lots of textual information. Thus, we propose to represent this framework via visualization to reduce learning time.

The information visualization has been developed using traditional web-based applications support [7]. With the NIST Cybersecurity Framework, this tech-

[1,2]The authors are with the Faculty of Information and Communication Technology, Mahidol University, Nakhon Pathom, Thailand. E-mail: wudhichart.saw@mahidol.ac.th and assadarat.khu@mahidol.ac.th
[3]The author is with the Agoda Company Pte. Ltd., Thailand. E-mail:nasorn.nia@gmail.com
[2]Corresponding author: assadarat.khu@mahidol.ac.th

nique may not be proper since the information in the framework is too large and complex to be displayed on a single monitor using support from traditional web-based applications. As a result, the proposed visualization for the NIST Cybersecurity Framework will be developed using SAGE2 [8]. It is so-called *SecSAGE*. SAGE2 stands for Scalable Amplified Group Environment version 2. SAGE2 has been introduced as a collaborative learning environment, which utilizes a multiple high-resolution large-scale display wall for efficiently viewing and analyzing information on visualization. In addition, it can be exploited as a learning tool for content that needs a lot of space.

Our experiments show that our visualization can help participants find information about the NIST Cybersecurity Framework faster than manually finding information on the document. In some tasks, participants can complete those tasks around 18.51 and 13.15 times faster than the manual method. In addition, we found that participants can correctly complete all tasks in the experiments using our visualization system. The participants can perform each task even better as a group. From our interviews, the participants found that SecSAGE is very helpful in finding and learning the NIST Cybersecurity Framework. Moreover, their satisfaction with SecSAGE is very high.

The rest of the paper is organized as follows. Section 2. presents background knowledge about the NIST Cybersecurity Framework and SAGE2, as well as state-of-the-art visualization techniques. Our proposed approach, including design and implementation, is described in Section 3. The experiment setup and evaluation results are shown in Section 4. We discuss some development challenges in Section 5. Finally, Section 6. concludes this study work and states a future work.

2. BACKGROUND AND RELATED WORKS

2.1 NIST Cybersecurity Framework

According to the Cybersecurity Enhancement Act of 2014 (CEA) [1], the role of the National Institute of Standards and Technology (NIST) has to include "identifying and developing cybersecurity risk frameworks" for critical infrastructure organizations. The NIST created the first version (1.0) of the framework in 2014 and the latest version (1.1) in 2018. The framework focuses on guiding cybersecurity activities and introducing cybersecurity risks into the organization's risk management process.

There are three parts of the NIST Cybersecurity Framework, which are *Framework Core*, *Framework Implementation Tiers*, and *Framework Profile* [1]. The details are as follows:

1. The Framework Core presents a set of cybersecurity activities that lead to the desired cybersecurity outcomes and their references.

2. The Framework Implementation Tiers provide a definition of how well the organization handles the cybersecurity risks in three aspects i.e., risk management process, integrated risk management program, and external participation.

3. The Framework Profile represents the outcomes based on business needs according to the organization's decision on the Framework core.

Since our work focuses only on the Framework Core, which consists of four elements working together i.e., *Functions*, *Categories*, *Subcategories*, and *Informative references*, we only illustrate them as follows:

1. Functions - control the basic cybersecurity activities at the highest level. Each function supports the organization to manage its cybersecurity risks by addressing risk, organizing information, and so on. There are five core functions with their unique identifiers described as follows:

a Identify (ID) - intent to improve the organizational understanding to manage cybersecurity risk to systems, people, assets, data, and capabilities.

b Protect (PR) - implement appropriate security to assure the delivery of critical service.

c Detect (DE) - implement appropriate actions to recognize the occurrence of cybersecurity events.

d Respond (RS) - implement appropriate actions to respond to the detected cybersecurity circumstance.

e Recover (RC) - implement appropriate activities to maintain plans and to restore any competency or services that were damaged from cybersecurity incidents.

2. Categories - the subdivisions of each core function, which ties to programmatic needs and particular cybersecurity events. For example, Data Security and Protective Technology are two of the categories for Protect function.

3. Subcategories - the subdivisions of each category, which are even more technical and specific activities. The purpose of subcategories is to support the outcomes of each category.

4. Informative references - the references to security standards, guidelines, and practices indicating how to achieve the desired outcomes.

Table 1 shows the overview of the Framework Core, which includes *Functions*, *Categories*, and their unique identifiers.

2.2 SAGE2: Scalable Amplified Group Environment

With the security standards, which have large document structures, collaboration is essential in order to explore and exploit them. Therefore, SAGE2 [8] plays an important role here. Recently, there have been many web-based collaboration systems, such as Google Meet, WebEx, and MS Team. Most of them allow single users to work with remote collabo-

Table 1: *Function and Category Unique Identifiers.*

Function Unique Identifier	Function	Category Unique Identifier	Category
ID	Identify	ID.AM	Asset Management
		ID.BE	Business Environment
		ID.GV	Governance
		ID.RA	Risk Assessment
		ID.RM	Risk Management Strategy
		ID.SC	Supply Chain Risk Management
PR	Protect	PR.AC	Identity Management and Access Control
		PR.AT	Awareness and Training
		PR.DS	Data Security
		PR.IP	Information Protection Processes and Procedures
		PR.MA	Maintenance
		PR.PT	Protective Technology
DE	Detect	DE.AE	Anomalies and Events
		DE.CM	Security Continuous Monitoring
		DE.DP	Detection Processes
RS	Respond	RS.RP	Response Planning
		RS.CO	Communications
		RS.AN	Analysis
		RS.MI	Mitigation
		RS.IM	Improvements
RC	Recover	RC.RP	Recovery Planning
		RC.IM	Improvements
		RC.CO	Communications

rators through video conferencing and screen sharing. SAGE2 does not only support video conferencing and screen sharing but also allows groups of users to work in front of large multiple displays or Tiled Display Wall (TDW). The example of use of SAGE2 is shown in Fig. 1.

SAGE2[1] is a middleware for collaborative working with Scalable Resolution Shared Display [8]. SAGE2 is a redesign of old SAGE software to enhance its abilities in data-intensive co-located and remote collaboration. SAGE2 supports many types of applications, including custom applications, and also works with any resolution display, which can be accessed from any device as well. It is free software that supports large displays and can be run on a web server and Electron[2]. Some applications are developed on SAGE2, which enables groups to organize large conference-scheduling data. SAGE2 supports many operating systems, including Windows, macOS, and Linux. It can be accessed and controlled by any computer that has the server IP address through the remote-distributed system. This means that SAGE2 can be used in online conferences or meetings as well.

Numerous research studies employ SAGE2 as the principal platform for data visualization. This research mainly focuses on collaboration. Some of them use SAGE2 as a platform for visualizing medical research data that can help users explore the insight of those data [9,10]. In addition, SAGE2 can be used to visualize data in the disaster management system, such as monitoring flood [11].

2.3 Visualization techniques

Information Visualization (InfoVis) is the research field that helps users to explore and understand easily [12,13]. InfoVis has been widely used in many data visual analysis applications such as business data [14-16], scientific data [17-20], medical research data [9,21]. In addition, it has been used to visualize some medical classifications [22-24].

In this research, the main data that we focus on is the text with a large structure that links between many security standards. Visualization is a method of transformation and representation of content into knowledge. There are many visualization techniques, including Tree, Graph, and Histogram, which may be useful for our research [7].

The Graph visualization technique is a way of rep-

[1] https://sage2.sagecommons.org/
[2] https://www.electronjs.org/

Fig.1: *Using SAGE2 to enhance the collaborated working environment. It can be seen that with large displays, it is better to see the overview of large data.*

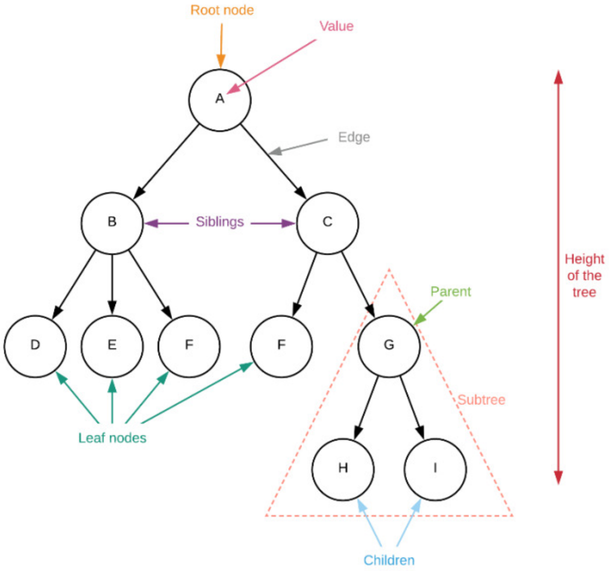

Fig.2: *An example of tree structure.*

resenting structural information as diagrams of abstract graphs data structure [25]. The graph structure consisted of 2 objects which are *nodes* and *edges*. Nodes are the objects, while edges represent their relationship. The Tree visualization is the visualization that used Tree data structure as a representation of the relationship between information [26]. The tree data structure is one type of Graph, but it does not contain any cycle in the structure. The first node of the Tree is called the root node, which branches to other nodes. The nodes that branch to other nodes are called parent nodes, and the branched nodes are called child nodes. The nodes without a child are called leaf nodes. An example of the tree can be seen in Fig. 2.

In addition, some research focuses on text and document visualization. The document visualization was proposed to display the information on small displays [27]. The researchers proposed a document segmentation and presentation system. This approach automatically segments documents regarding the screen size and the structure of content using a Tree structure. From a review article, many document visualizations have been reviewed [28]. Most of the techniques focused on visualizing words in the document regarding vocabulary, semantic structure, document version, and document relationship. The main visualizations used are tree and graph techniques. This research is not directly related to our work. Some parts, such as tree and graph techniques, are useful for this research. In addition, there is research working on an interactive visualization for topic models [29]. This research exploited Latent Dirichlet Allocation (LDA) to extract topics from text and use sunburst visualization to visualize the hierarchy of the topics. Another research focuses on visualizing the medical histories of patients [30]. This provides useful information for accurate treatment planning. They utilize a representation learning algorithm to create a semantic representation space of documents. Then they visualize the information using the tree technique. Again, it seems that the tree technique is interesting to be explored for visualizing the cybersecurity framework. To our knowledge, there is no visualization for the cybersecurity framework.

This research mainly focuses on visualizing the Framework Core, which consists of five functions: Identify, Protect, Detect, Respond, and Recover using the tree technique. Each core function represents the main cybersecurity activities and contains categories, which represent the sub-activities for each function. Moreover, each category also contains subcategories that contain informative references such as guidelines and references to the existing security standard [1]. Importantly, this research can help users explore the security standard collaboratively.

3. THE PROPOSED APPROACH

In this section, it articulates our research work with the ideas mentioned above. It presents the proposed approach that is exploited to create a visualization system for the NIST cybersecurity framework on SAGE2. This system is called *SecSAGE.*

3.1 Data description and processing

This section describes the data that is used for designing and developing SecSAGE, which is obtained from the NIST cybersecurity framework [1]. The structure of data in the NIST cybersecurity framework is described in Section 2.1 The input form of the data is in Portable Document Format (PDF).

We use *Optical Character Recognition* (OCR) [31] to extract the raw data from the NIST cybersecurity framework through *Python Tesseract* [3]. Then we transform the raw data into JSON [32] format.

3.2 Visualization technique

According to the NIST Cybersecurity Framework document, the framework core is categorized into four divisions. The visualization technique that we decided to use is the Tree visualization technique [26]. The reason that we decided to use this technique to visualize the framework is because this technique is appropriate for hierarchy data structure, and the framework structure can be seen as a hierarchy as well. In addition, tree visualization gives information to users at different levels. Therefore, users do not need to view the massive information of the framework.

This technique represents the first node as the root node of the Tree and the other node as its children. For this cybersecurity framework, there are 4 levels of the tree illustrated. The root node presents core functions, and the first level of child nodes represents a division, such as category division. The next level of child nodes represents the subcategory division that is related to the parent category. Finally, the last level of child nodes or leaf nodes represents the informative references division which contains the reference to each security standard. For example, a node represents five core functions, and its children represent categories that contain each function as shown in Fig. 3.

3.3 Color scheme

In the root node, the five core functions are given different colors, which are consistent with those in the NIST Cybersecurity Framework. The Identify function (ID) is assigned a blue color. Protect function (PR) is purple. Detect function (DE) is yellow. The Response function (RS) is assigned with red color, and the Recover function (RC) is green color [1]. This color scheme will help users get used to the cybersecurity framework easily.

3.4 SecSAGE Implementation

The SecSAGE system is mainly implemented based on the programming languages, i.e., HTML, CSS, and JavaScript with NodeJs and Electron. The information that we extracted from the NIST Cybersecurity Framework document using Python is stored as an object in a JSON file (`data.json`). JSON is a universal data structure for storing and transporting data [33]. In our work, we mainly implement SecSAGE using Electron. The structure of our code is having a JavaScript file that runs on Electron to call the HTML files and load their content to show on SAGE2 multiple-display

https://github.com/madmaze/pytesseract

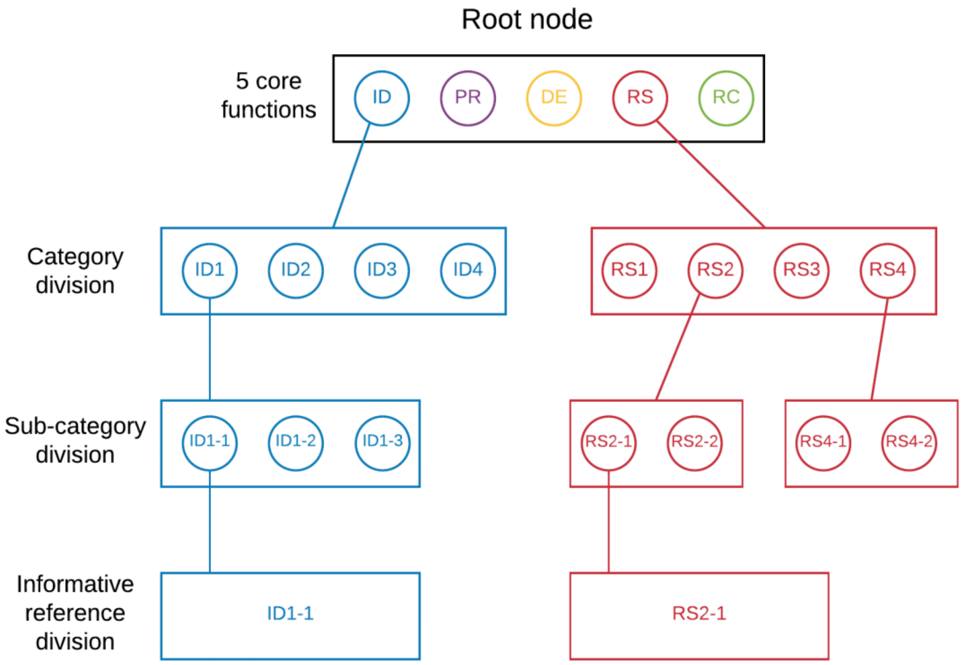

Fig.3: *An example map represents each node as a division.*

as shown in Fig. 1. In the HTML part, the program loads the information from the JSON file and displays it in the allocated section. There are 5 HTML files, each representing a division of the Framework core. They are `main.html`, `cat.html`, `sub-cat.html`, `control.html`, and `ref.html`. For the JavaScript file, we name it as `security.js`. The HTML files are the template files for filling content. The content of each HTML file will be dynamically gathered from `data.json` through `security.js`.

Fig. 4 shows the workflow of this visualization on the file level. The `main.html` file represents the five core functions of the framework core. It also works as a root node, where if any core function is clicked, it will navigate to that function's categories. From the root node (`main.html`), after the users click on any core function, the system will open a category node, which is implemented in `cat.html` file, which basically represents the category division. In order to get the correct category node, the `main.html` file has to send the information of the clicked core function to the `cat.html` file as well. The mechanism that we use to send this data is to add it as a parameter with the URL that links to the `cat.html` file. Therefore, it can retrieve the information of the core function through its URL. With the information of the core function, the `cat.html` file can access the right category in the JSON file and be able to retrieve the information inside it as well as display the retrieved information.

When each category of each core function is clicked, the `cat.html` will open a new HTML file, which is the `sub-cat.html` and send the information of the clicked category to that file through the URL with a parameter for opening the `sub-cat.html` file. The `sub-cat.html` works similarly to the `cat.html` file. It retrieves the information from its URL and uses that information to get the information of the subcategory inside the JSON file.

When each sub-category is clicked, the `sub-cat.html` file will open another HTML file, which is the `control.html` and send the information of the clicked sub-category inside the URL. The `control.html` file represents the informative references of the subcategory that expands to this node. The content of the `control.html` file is the description of the activity of the sub-category and the references to the cybersecurity standard that has this activity. The security standards that have been referenced in this part are NIST SP 800-53 [34], COBIT 5 [35], CIS CSC [36], ISO/IEC 27001:2013 [37], ISA 62443-2-1:2009, and ISA 62443-3-3:2013 [38]. For instance, the subcategory ID is PR.DS-1, which is about the activity that the data-at-rest is protected. The `control.html` file will show the description of activity PR.DS-1, and the reference to security standards like Appendix 8.2.3 of ISO/IEC 27001:2013. Furthermore, each cybersecurity standard can be interacted with, and it will lead to `ref.html` file, which retrieves the content in those cybersecurity standard documents. The `ref.html` file reads the content of the security standard document from PDF using *PDF.js* [4] to extract the content inside.

The next part is the JavaScript file, which is named as `security.js`. This file works as an intermediary,

https://mozilla.github.io/pdf.js/

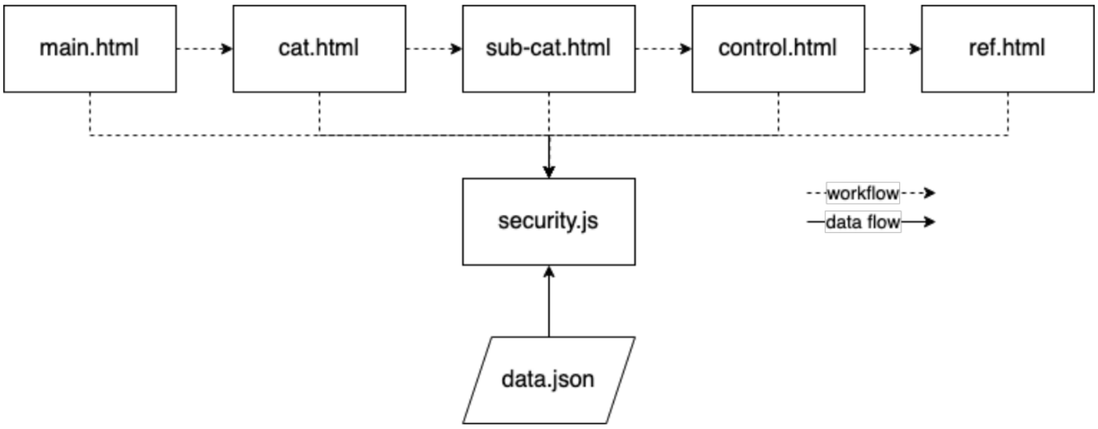

Fig.4: *The work flow of visualization program*

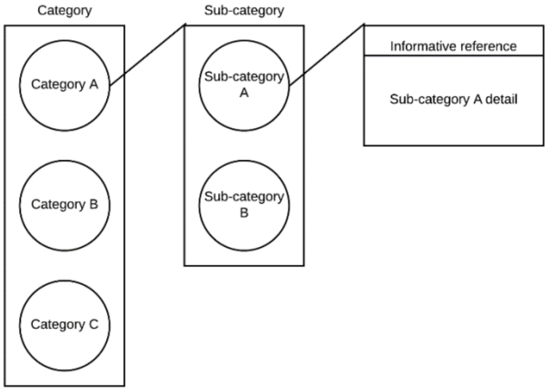

Fig.5: *The prototype of the security visualization application*

which pulls the content from HTML and shows it on the SAGE2 display. The method that we use is to open the HTML file as a node by using the property of *WebView*, which is a free application on SAGE2. The prototype of our application can be seen in Fig. 5. When an HTML file tries to open another node, this `security.js` file will open itself as another application that works as a child application. Then the parent node draws a line to connect itself with the opened child using the top left of the child window and the clicked topic's coordinates as coordinates of the line.

4. SYSTEM EVALUATION

This section presents the evaluation process of SecSAGE. The main goal is to determine how users will use the system in information and knowledge finding and how SecSAGE is useful for them. We are particularly interested in experts' reactions to and usage of the visualization features. This will help us understand the advantages of visualization in the cybersecurity framework field. and find potential difficulties that users may experience. To achieve this, the task-based test and semi-structured interviews were conducted, and we obtained feedback and comments from experts.

4.1 Study setup

4.1.1 Participants

We invited five security experts (three females and two males) from both the private and public sectors. They will be represented as P1 - P5. We asked them to evaluate their familiarity with the NIST Cybersecurity Framework (NIST CSF), COBIT, ISA 62443-2-1:2009, ISA 62443-3-3:2013, ISO/IEC 27001, ISO/IEC 27002, NIST SP 800-53 Rev.4, CIS CSC, Information Security Management System (ISMS), and SecSAGE on 5-point Likert scale (0-Never Heard of This and 4-Very Familiar). The result for each topic is shown in Table 2, indicating that the participants are not familiar with the SAGE2 application; not quite familiar with ISA 62443-2-1:2009 and ISA 62443-3-3:2013; and relatively familiar with NIST Cybersecurity Framework, ISO/IEC 27001 and ISO/IEC 27002.

4.1.2 Procedures

We deployed SecSAGE on our multiple display wall system equipped with SAGE2. We divided our system evaluation into 3 parts, individual task-based test, group task-based test, and interviews. We ob-

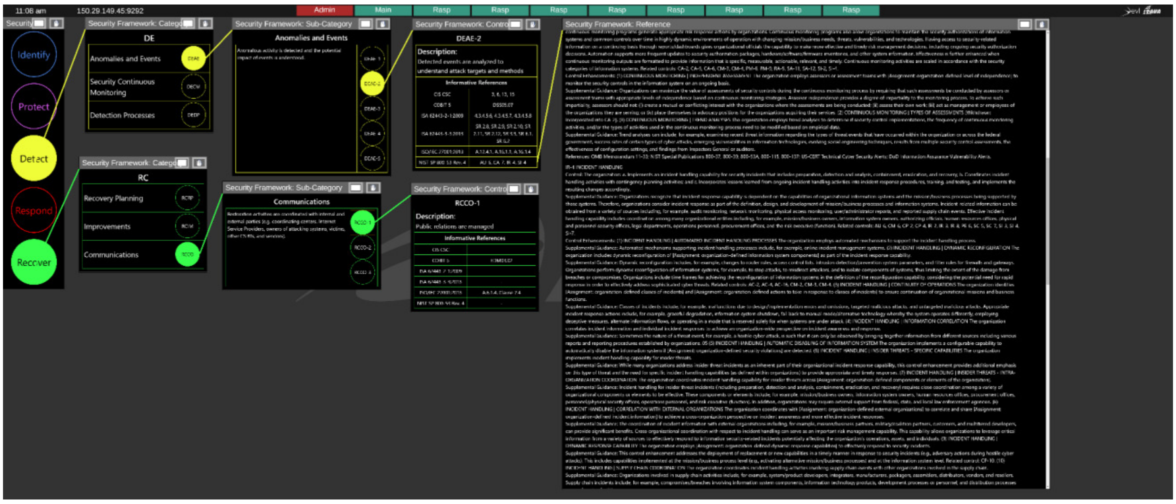

Fig.6: The visualization application on SAGE2

Table 2: Results of participants' pre-test

	P1	P2	P3	P4	P5	Avg.
NIST CSF	3	4	4	2	2	3
COBIT	2	4	3	1	0	2
ISA 62443-2-1:2009	2	2	0	0	0	0.8
ISA 62443-3-3:2013	2	2	0	0	0	0.8
ISO/IEC 27001	3	4	4	3	3	3.4
ISO/IEC 27002	2	4	4	3	3	3.2
NIST SP 800-53 Rev.4	2	3	4	0	1	2
CIS CSC	3	4	4	0	0	2.2
ISMS	3	4	4	1	2	2.8
SAGE2	0	0	0	0	0	0

tained the signatures of the participants on consent forms. To gain deeper insights, we provided the participants with detailed training and asked them to use the system before the task-based tests and interviews. In the individual and group task-based tests, the participants were asked to perform each task using different methods, which are manually using the NIST cybersecurity framework in PDF format, called *manual method*, and using SecSAGE. For each task, we read the task for them and ensured that they understood the task before they performed the task. Then we recorded the time spent for each task to compare between the manual method and SecSAGE. For the group task-based test, all participants were asked to perform each task together using the manual method and SecSAGE. After the individual and group task-based tests were completed, each participant was asked to fill in a questionnaire to rate the system and to have semi-structured interviews.

4.1.3 Tasks and Interview Questions

As mentioned above, our goal is to gather a deeper understanding of the reaction of users and usage of SecSAGE. Therefore, the tasks and interview questions center around the goal. Particularly, we are interested in the ease of learning and use of visualization to understand whether it is challenging for participants who are familiar with the NIST cybersecurity framework. The example tasks are listed as follows:

T1: How many categories of the function "Protect" (PR)?

T2: How many subcategories of the category "Awareness and Training" (PR.AT)?

T3: Please list the clause or annex of ISO/IEC 27001:2013 that is related to the subcategory "PR.AT-1".

T4: How many standards are related to the subcategory "PR.AT-3"?

T5: Assume that you want to explain the standards that are related to the subcategory "PR.AT-2". How would you describe the information shown in this visualization?

The tasks for the manual method and SecSAGE test are similar. There are only 2 differences, which are 1. **T5** was not performed in manual method test since this question is to explain the flow of the visualization and 2. the core function, category, and subcategory were different for each test. Similarly, the tasks for individual task-based tests and group task-based tests are similar. The only difference is the core function, category, and subcategory were different for each test. These tasks were evaluated by recording the time that the participants spent. Task **T5** is used to get insights and evaluate whether the participants understand the whole visualization.

In the interviews, there are 4 questions as follows:

Q1: Do you think the visualization helps give information about the NIST cybersecurity framework and its related standards?

Q2: Do you think the SAGE2 application is helpful for collaboration?

Q3: How much are you satisfied with the visualization?

Q4: How much are you satisfied with the SAGE2 application?

Note that the SAGE2 application in the questions refers to SecSAGE. **Q1** and **Q2** focus on the effectiveness of SecSAGE. **Q3** and **Q4** focus on the satisfaction of SecSAGE. All questions were rated by 5-point Likert scale (**Q1** and **Q2**: 0-Not Effective and 4-Very Effective and **Q3** and **Q4**: 0-Not Satisfied and 4-Satisfied). In the interviews, we also asked their opinions on the usefulness of SecSAGE for knowledge sharing and learning.

4.2 Results and analysis

With the proposed approach, the visualization was implemented as a SAGE2 application called SecSAGE. The result of our visualization program is shown in Fig. 6. This visualization can illustrate a lot of information on the Tiled Display Wall (TDW), which can help users work together and get an understanding of the NIST cybersecurity framework. Furthermore, the information visualized provides the mapping between different cybersecurity standards based on the NIST cybersecurity framework. In addition, it supports collaboration in designing cybersecurity standards that are suitable for them. Moreover, this visualization allows users to save the working state via the save state function. The save state function is a built-in function in SAGE2 that allows users to save and load the current screen.

Table 3 shows the time spent by each participant and group of **T1 - T4** in the task-based test. As can be seen from Table 3, it shows that most participants can complete correctly tasks **T1 - T4** using SecSAGE faster than the manual method, especially **T1** and **T2**. On average, all participants can perform **T1** 8.90, **T2** 5.51, **T3** 1.40, and **T4** 1.17 times faster than the manual method. One participant can complete **T1** and **T2** using SecSAGE around 18.51 and 13.15 times faster than the manual method, respectively. For Task **T1**, 2 participants incorrectly answered this task using the manual method. On the other hand, no one incorrectly answered all tasks using SecSAGE. Therefore, this shows that SecSAGE can help users find information correctly from the NIST cybersecurity framework compared to its document in PDF format. For the task **T5**, all participants can correctly explain the standards that are related to a given subcategory. In addition, they can use SecSAGE to navigate correctly to the related standards.

For the group experiment, it can be seen again that using SecSAGE outperforms the manual method. Moreover, the performance of the group is even better than the performance of all individual participants. Thus, this shows that SecSAGE, which is developed on SAGE2, allows users to collaboratively find information on the NIST cybersecurity framework and easily work together.

Table 4 shows the interview results of the participants. As can be seen, most participants gave 4 to all questions. Next, we discuss their feedback based on these questions.

Helpfulness of Visualization (Q1). All participants agreed that the overall visualization can help them effortlessly find information about the NIST cybersecurity framework. In addition, SecSAGE is easy to learn and use.

Helpfulness of SAGE2 (Q2). Most participants agreed that the SAGE2 system can help them easily find and navigate through the NIST cybersecurity framework as a team. However, some participants commented that the pointer on the SAGE2 system confused them at some points, resulting in different scores.

Visualization Satisfaction (Q3). All participants showed open and positive attitudes towards the tree visualization on SecSAGE. They stated that the visualization did not overwhelm them by the information. They can navigate through the NIST cybersecurity framework level by level. In addition, they can focus on the information on each level.

SAGE2 Satisfaction (Q4). All participants held positive attitudes towards the SAGE2 system. They agree that the SAGE2 system can be used not only to visualize the NIST cybersecurity framework but also other standards, which are more complex and need more collaboration.

In addition to the answers to the four questions above, P3 and P4 were confused about the function's unique identifier since they were not shown at the first window when the program started. P3 said, "I suggest giving the full name of each function's unique identifier at the first window." P4 gave a similar comment, "I think it should have full name of the function unique identifier at first" In terms of practices of this application, P1 explained, "I think it is not used often in practices. It can be used once a year during the security controls reviewing process. However, it can be very useful for training new employees to learn and remember the framework easier and faster. In addition, using different colors and separate windows is very helpful" P2 mentioned, "It is not quite useful in real jobs to find the number of categories or subcat-

Table 3: *Results of participants and group of **T1 - T4** in the task-based test (Time spent in second).*

	Manual method				SecSAGE			
	T1	**T2**	**T3**	**T4**	**T1**	**T2**	**T3**	**T4**
P1	8.00	24.68	17.23	11.90	6.25	8.56	12.11	11.00
P2	97.90	59.91	25.00	29.06	10.33	11.10	16.13	18.83
P3	70.37	13.00	20.00	11.25	9.48	11.00	15.73	11.00
P4	127.00	78.00	13.00	6.00	6.86	5.93	11.28	5.53
P5	42.98	23.13	10.23	7.75	5.51	4.68	6.30	7.00
Group	18.90	35.45	5.33	5.76	3.40	4.45	5.35	5.00

Table 4: *Results of participants' interviews*

	P1	**P2**	**P3**	**P4**	**P5**
Q1	4	4	4	4	4
Q2	3	3	4	4	2
Q3	4	4	4	4	4
Q4	4	3	3	4	4

egories as in the test questions. However, it is very useful for IT auditors to write reports referencing the standards, which are hard to remember. It especially helps train newcomers since there is a high turnover rate in this field.

5. DEVELOPMENT DISCUSSION

The development of SecSAGE involves several challenges across data gathering and formatting, user interface, and compatibility. This discussion presents these challenges and our proposed solutions.

As mentioned above, the data used in SecSAGE is primarily sourced from the NIST cybersecurity framework PDF file. In our development, we use Optical Character Recognition (OCR) to extract data from the PDF file using Python Tesseract. The main challenge is correctly extracting all the information from the PDF file. After using Python Tesseract, the results were not entirely accurate. Some data was missing, and there were inaccuracies, such as extra commas and special characters. To address this problem, we need to create a function to detect these issues and reformat the results from Python Tesseract. Next, we need to format the results in JSON format. In this case, a function to detect the topics of necessary data is needed since we need to extract the data and fill in the JSON structure that we defined. For example, the JSON structure requires a title and the relationship between categories and sub-categories.

The next challenge is developing a user interface on the SAGE2 platform. Even though applications developed on the SAGE2 platform are based on JavaScript and NodeJS, we need to consider the structure and components of the SAGE2 platform, especially SecSAGE. As can be seen in Figure 3, the tree structure of SecSAGE has four levels (divisions), each of which is connected. During development, we need to create different HTML files for different levels as templates for filling content. The challenge is how to connect these HTML files. Fortunately, the SAGE2 platform provides a function called `SAGE2_AddNewChild`, which allows SecSAGE to connect these files in a hierarchical manner. For example, in this case, `cat.html` is a child of `main.html`, as shown in Figure 4. Moreover, the SAGE2 platform provides a function called `SAGE2_SendDataToChild`, which allows SecSAGE to pass data from a parent node to a child node.

Finally, the compatibility of SecSAGE for the tiled display wall needs to be considered during development since the tiled display wall may have different resolutions. Therefore, we need to calculate the resolution of SecSAGE carefully. Again, for this case, the SAGE2 platform provides `manifest.json`, where the resolutions of each level can be defined. This allows us to set the resolution of each level appropriately.

6. CONCLUSION AND FUTURE WORK

In order to effectively visualize the NIST cybersecurity framework, we have determined the built-in function in SAGE2 and adapted it to create our visualization program called *SecSAGE*. The visualization on SAGE2 also allows users to interact and navigate through the information easily. In addition, the successful visualization of the NIST cybersecurity framework would help many organizations get a better understanding of the context of the security framework and the security standards. Using our application with a high-resolution Tiled Display Wall (TDW) may be the better approach to understanding the information.

However, there are some limitations. Some information is text-based, which we believe it can be improved by transforming that information into visual language with the help of an ontology. Another limitation is the leading lines, which are drawn from categories to sub-categories. They may overlap with the content if users navigate through many categories. Therefore, our visualization application can be further developed to cover more references of other security standards rather than NIST SP 800-53 Rev.4 with better formatting. Furthermore, it would cover all the built-in functions of SAGE2 and fix the limitations of the visualization.

ACKNOWLEDGEMENT

We acknowledge the support of the Faculty of Information and Communication Technology, Mahidol University, Thailand, and thank the experts for their generous support and participation in this project.

AUTHOR CONTRIBUTIONS

Conceptualization, W.S. and A.K.; methodology, W.S.; software, W.S. and N.N.; evaluation, W.S. and A.K.; validation, W.S. and A.K.; formal analysis, W.S. and A.K.; investigation, W.S. and A.K.; data curation, W.S. and A.K.; writing—original draft preparation, W.S. and A.K.; writing—review and editing, W.S. and A.K.; visualization, W.S., A.K. and N.N. All authors have read and agreed to the published version of the manuscript.

References

[1] National Institute of Standards and Technology, "Framework for Improving Critical Infrastructure Cybersecurity," 2018.

[2] Joint Technical Committee ISO/IEC JTC 1, "ISO/IEC 27001 Information security, cybersecurity and privacy protection - Information security management systems - Requirements," 2022.

[3] Joint Technical Committee ISO/IEC JTC 1, "ISO/IEC 27002 Information security, cybersecurity and privacy protection - Information security controls," 2022.

[4] ISACA, "COBIT2019 Framework," 2019.

[5] ISACA, "Security and Privacy Controls for Information Systems and Organizations," 2020.

[6] Center for Internet Security, "CIS Critical Security Controls," 2021.

[7] R. M. Rohrer and E. Swing, "Web-based information visualization," in IEEE Computer Graphics and Applications, vol. 17, no. 4, pp. 52-59, July-Aug. 1997

[8] T. Marrinan *et al.*, "SAGE2: A new approach for data intensive collaboration using Scalable Resolution Shared Displays," *10th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*, Miami, FL, USA, pp. 177-186, 2014.

[9] J. Mitrpanont, J. Roungsuriyaviboon, T. Sathapornwatanakul, W. Sawangphol, D. Kobayashi and J. H. Haga, "Extending MedThaiVis-Thai medical research visualization to SAGE2 display walls," *2017 2nd International Conference on Information Technology (INCIT)*, Nakhonpathom, Thailand, pp. 1-6, 2017.

[10] J. Mitrpanont, W. Sawangphol, S. Sillapathadapong, S. Suthinuntasook, W. Thongrattana and J. Haga, "MedThaiSAGE2: Enhancing the Decision Support System using Rich Visualization on SAGE 2," *2020 - 5th International Conference on Information Technology (InCIT)*, Chonburi, Thailand, pp. 128-133, 2020.

[11] D. Kobayashi, M. Ready, A. Gonzalez Martinez, N. Kirshenbaum, T. Seto-Mook, J. Leigh and J. Haga, "Sage river disaster information (sagerdi): Demonstrating application data sharing in sage2," in *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, ISS '18, (New York, NY, USA), pp. 33–42, Association for Computing Machinery, 2018.

[12] H. Shiravi, A. Shiravi and A. A. Ghorbani, "A survey of visualization systems for network security," *IEEE Transactions on visualization and computer graphics*, vol. 18, no. 8, pp. 1313–1329, 2011.

[13] S. Liu, W. Cui, Y. Wu, and M. Liu, "A survey on information visualization: recent advances and challenges," *The Visual Computer*, vol. 30, no. 12, pp. 1373–1393, 2014.

[14] S. Liu, N. Cao and H. Lv, "Interactive Visual Analysis of the NSF Funding Information," *2008 IEEE Pacific Visualization Symposium*, Kyoto, japan, pp. 183-190, 2008.

[15] B. Alper, T. Hollerer, J. Kuchera-Morin and A. Forbes, "Stereoscopic highlighting: 2d graph visualization on stereo displays," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2325–2333, 2011.

[16] S. Ko, R. Maciejewski, Y. Jang and D. S. Ebert, "Marketanalyzer: an interactive visual analytics system for analyzing competitive advantage using point of sale data," in *Computer Graphics Forum*, vol. 31, pp. 1245–1254, Wiley Online Library, 2012.

[17] O. Kersting and J. Döllner, "Interactive 3d visualization of vector data in gis," in *Proceedings of the 10th ACM International Symposium on Advances in Geographic Information Systems*, GIS '02, (New York, NY, USA), pp. 107–112, Association for Computing Machinery, 2002.

[18] R. ElHakim and M. ElHelw, "Interactive 3d visualization for wireless sensor networks," *The Visual Computer*, vol. 26, no. 6-8, pp. 1071–1077, 2010.

[19] R. D. Müller, X. Qin, D. T. Sandwell, A. Dutkiewicz, S. E. Williams, N. Flament, S. Maus and M. Seton, "The gplates portal: cloud-based interactive 3d visualization of global geophysical and geological data in a web browser," *PloS one*, vol. 11, no. 3, p.e0150883, 2016.

[20] D. Bouyssié, J. Lesne, M. Locard-Paulet, R. Albigot, O. Burlet-Schiltz and J. Marcoux, "Hdx-viewer: interactive 3d visualization of hydrogen–deuterium exchange data," *Bioinfor-matics*, vol. 35, no. 24, pp. 5331–5333, 2019.

[21] J. Mitrpanont, N. Janekitiworapong, S. Ongsri-

trakul and S. Varasai, “MedThaiVis: An approach for thai biomedical data visualization,” *2017 6th ICT International Student Project Conference (ICT-ISPC)*, Johor, Malaysia, pp. 1-4, 2017.

[22] J. Geller, C. Ochs, Y. Perl and J. Xu, “New abstraction networks and a new visualization tool in support of auditing the snomed ct con- tent,” in *AMIA Annual Symposium Proceedings*, vol. 2012, p. 237, American Medical Informatics Association, 2012.

[23] V. Della Mea *et al.*, “A Web-Based Tool for Development of a Common Ontology between ICD11 and SNOMED-CT,” *2014 IEEE International Conference on Healthcare Informatics*, Verona, Italy, pp. 144-148, 2014.

[24] D. Chen, R. Zhang, H. Zhao and J. Feng, “A bibliometric analysis of the development of icd-11 in medical informatics,” *Journal of Healthcare Engineering*, vol. 2019, 2019.

[25] I. Herman, G. Melan¸con and M. S. Marshall, “Graph visualization and navigation in information visualization: A survey,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 6, no. 1, pp. 24–43, 2000.

[26] H.-J. Schulz, “Treevis.net: A tree visualization reference,” *IEEE Computer Graphics and Applications*, vol. 31, no. 6, pp. 11–15, 2011.

[27] K. Kit Hoi, D. Lun Lee and J. Xu, “Document visualization on small displays,” in *Mobile Data Management* (M.-S. Chen, P. K. Chrysanthis, M. Sloman, and A. Zaslavsky, eds.), (Berlin, Heidelberg), pp. 262–278, Springer Berlin Heidelberg, 2003.

[28] Q. Gan, M. Zhu, M. Li, T. Liang, Y. Cao and B. Zhou, “Document visualization: an overview of current research,” *WIREs Computational Statistics*, vol. 6, no. 1, pp. 19–36, 2014.

[29] A. Smith, T. Hawes and M. Myers, “Hiearchie: Visualization for hierarchical topic models,” in *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, pp. 71–78, 2014.

[30] H. Z. Yerebakan, Y. Shinagawa, P. Bhatia and Y. Zhan, “Document representation learning for patient history visualization,” in *Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations*, (Santa Fe, New Mexico), pp. 30–33, Association for Computational Linguistics, Aug. 2018.

[31] N. Islam, Z. Islam and N. Noor, “A survey on optical character recognition system,” *ArXiv*, vol. abs/1710.05703, 2017.

[32] F. Pezoa, J. L. Reutter, F. Suarez, M. Ugarte and D. Vrgoč, “Foundations of json schema,” in *Proceedings of the 25th International Conference on World Wide Web*, WWW’16, (Republic and Canton of Geneva, CHE), pp. 263–273, International World Wide Web Conferences Steering Committee, 2016.

[33] Refsnes Data, “What is JSON?,” 2019.

[34] J. T. Force and T. Initiative, “Security and privacy controls for federal information systems and organizations,” *NIST Special Publication*, vol. 800, no. 53, pp. 8–13, 2013.

[35] P. Bernard, COBIT®5-A management guide. Van Haren, 2012.

[36] Center for Internet Security, “CIS Center for Internet Security,” 2017.

[37] G. Disterer, “ISO/IEC 27000, 27001 and 27002 for information security management,” *Journal of Information Security*, vol. 4, no. 2, pp. 92–100, 2013.

[38] International Society of Automation, “International Society of Automation,” 2005.

Wudhichart Sawangphol is a lecturer at the Faculty of Information and Communication Technology, Mahidol University, Thailand. In addition, he is a program chair of the Master of Science in Game Technology and Gamification. He holds a Master of Information Technology (MIT) with Honours in Data Management, Software Engineering, and Knowledge Engineering and a Ph.D. in Ontology Reasoning and Optimization from Monash University. His interesting research areas are Automated Reasoning, Ontology, Ontology Reasoning, Optimization, Machine Learning, Deep Learning, and Data Visualisation.

Assadarat Khurat is a lecturer at the Faculty of Information and Communication Technology, Mahidol University, Thailand. She obtained a Master’s in Information and Communication Systems and a Ph.D. in Computer Security. Her research interests are Application Security, Network Security, Privacy, Intrusion Detection Systems, IT Audit, Information Security Management, and Risk Management.

Nasorn Niampradit is a software engineer at Agoda. He graduated from the Faculty of Information and Communication Technology, Mahidol University, Thailand, with first-class honors in Computer Science. His areas of interest are Artificial Intelligence, Deep Learning, Data Visualization, and Algorithms.