Training on social media cybersecurity skills in the healthcare context

Mario Fernandez-Tarraga, Alejandro-David Cayuela-Tudela, Pantaleone Nespoli, Joaquin Garcia-Alfaro, Félix Gómez Mármol

Training on Social Media Cybersecurity Skills in the Healthcare Context

Mario Fernández Tárraga[1 [0009−0009−6403−3243]], Alejandro David Cayuela Tudela[1 [0009−0008−6627−6399]], Pantaleone Nespoli ✉[1,2 [0000−0002−4041−1205]], Joaquin Garcia-Alfaro[2 [0000−0002−7453−4393]], and Félix Gómez Mármol[1 [0000−0002−6424−3322]]

[1] Departamento de Ingeniería de la Información y las Comunicaciones, Universidad de Murcia, 30100, Murcia, Spain
{mario.fernandezt,alejandrodavid.cayuelat, pantaleone.nespoli, felixgm}@um.es

[2] SAMOVAR,Télécom SudParis, Institut Polytechnique de Paris, 19 place Marguerite Perey, 91120 Palaiseau, France
{pantaleone.nespoli, joaquin.garcia_alfaro}@telecom-sudparis.eu

Abstract. In the last decades, social media has experienced exponential growth due to its popularity and advantages in communication, connection, and broadcasting, etc. Nevertheless, social media also presents disadvantages and threats that can be exploited for evil ends. Indeed, it has become a vector of attack for cybercriminals and scammers. Broadcasting and public access have allowed the public to post fake news and disinformation. The wrong use of social media by users due to negligence or unawareness allows these threats to succeed. In this scenario, the healthcare sector is also affected by social media threats, as can be seen over the years in the frequent daily phishing attacks, vulnerable devices, data leaks of personal information, etc. This article proposes an automated tool for training social media cybersecurity competencies in the social and professional sectors of the healthcare environment, built on the Cyber Range context. It allows automating the generation and configuration of simulated social media exercises with several levels of content, difficulty, and realism, creating multiple hyperrealistic situations across a wide range of possibilities. Training phishing attacks, crisis management, social media attacks, and pandemic or disease crisis disinformation are some of the possibilities, both theoretical and practical.

Keywords: Cyber Range · Cybersecurity · Cyberdefense · Social media simulation · Social media cybersecurity skills · Digital education · E-health · Digital health · Healthcare cybersecurity

1 Introduction

Since the second half of the 20th century, technology has evolved exponentially, culminating in the present technological and digital state. One of the main exponents of such an evolution is social media, whose hyperconnectivity has facilitated the globalisation of information and communication with thousands of people, among many other advantages [5]. Even with the daily use of social media, not all users are aware of the

disadvantages and threats posed by the use of these platforms, where different actors with malicious objectives participate directly or indirectly.

Technological evolution affects society in all areas. The transformation of sectors such as healthcare towards a more digitalised environment has led to the widespread use of the internet, making social media and other online services indispensable. Consequently, healthcare is also vulnerable to various threats associated with internet and social media usage. The study conducted by the European Network and Information Security Agency (ENISA) [6] outlines cybersecurity threats within healthcare environments, many of which are connected to or stem from social media.

On the one hand, there are problems related to communication and interaction between patients and professionals on social media, as well as threats related to privacy and security [4], the influence of social bots and their dangers [15], threats related to disinformation or misinformation in the medical field [10]. Not to mention the use of social media in managing potential social and health crises such as the COVID-19 pandemic [8]. On the other hand, social media enables both gathering information to prepare more elaborate attacks and attacking vulnerable users with social engineering, phishing, OSINT techniques [14] or uploading malware. Phishing is a common attack in this area, with high success rates, including within hospital environments [18]. The majority of successful attempts can be attributed to user distractions and negligence, since many attacks are recognisable with awareness. The gravest consequences, such as data breaches, can result in catastrophic losses, both in terms of patients' personal information and the loss of essential patient diagnostics and records necessary for providing continuous, essential services [19]. Ultimately, many of these challenges converge on the vulnerability of the human factor in the cybersecurity chain, as underscored by numerous studies [3, 7, 9, 12].

This article proposes an automated tool for social and professional training to mitigate the aforementioned challenges, building upon the structure of well-known Cyber Ranges [16], through an automated social media training system. This proposal holds immense importance, especially in the healthcare context, where training and knowledge can make good habits and reduce cybersecurity incidents [21]. This training proposal toggles with the architecture required for its implementation and its automation. It allows for the easy setup and control of cyberexercises by automating their creation, execution, and adaptation to extensible, multidisciplinary, and hyperrealistic challenges.

The article is structured as follows: Section 2 shows the lack of academic work. Then, Section 3 explains the social media simulator proposed and its architecture. Next, Section 4 shows the general implementation of the proposal. Then, Section 5 displays an example of provisioning a basic exercise within the social media platform. Finally, Section 6 outlines the drawn conclusions and future work for this proposal.

2 State of the Art

Social simulation does not represent a novel idea in the literature and has been extensively explored in various studies focusing on behaviour, polarization, topic-centered communities, relationships between simulated agents, and more [1]. However, the specific simulation of social media platforms, especially concerning competency training

in cybersecurity, remains a relatively unexplored domain. In other words, there is a frightening lack of tools that simulate social media scenarios to train competencies of cybersecurity. Furthermore, the few solutions are limited or specific approaches, and most of the solutions are developed by private companies, leaving little research in academia. The following paragraphs analyse the main solutions found in the literature, showing their main characteristics and comparisons.

The master's thesis [2] (CYRAN Cyber Range Extension) uses the first versions of open source social media platforms, with major limitations in functionality. These are the impossibility of using multimedia content and the automation of user creation, as well as the configuration of user profiles. The publication of content is done through pregenerated posts and dynamically generated (autogenerated) posts, which are published automatically according to the parameters entered in the templates for exercise generation. In this sense, exercises can be configured manually using these templates, but there are no mechanisms to automate the exercise creation process. In addition, the tool does not support hot configuration and settings, and does not allow multiple runs to be controlled simultaneously. However, the tool is designed to support several real users or students simultaneously and concurrently.

Another project [13] (Somulator) is developed by the Norwegian Defence Research Establishment (FFI) and by the Norwegian University of Science and Technology (NTNU), supports multimedia content, and allows the creation and configuration of users, automating only the creation processes. The publications used are pregenerated and loaded with templates. It does not allow the configuration of statistics, but it does allow the configuration of exercises (and also during the execution). The tool uses several open source platforms and is designed to be used simultaneously by many real users or students. In addition, it is complex to control several instances in parallel.

Besides, the third software service [17] (Prevency) also supports multimedia content and allows users to be created both individually and collectively by automating this process. The configuration of users and profiles can also be done both individually and automatically when creating such collective profiles. Publications are pregenerated and must be published deliberately. Additionally, it allows the configuration of statistics for both users and publications and permits hot settings. The tool is intended for use with several real users simultaneously; however, like its counterparts, it is not aimed at deploying an exercise across various individual instances (parallel control).

Table 1 shows the main differences according to the features of each proposal or tool, as well as their comparison with the proposal in this article. In particular, the last lines of the table show the new concepts explained in Section 3.1, which have not been explored by any of the previous proposals. In the following list, the meaning of each of the entries in the table previously mentioned is described:

- **Multimedia content**: The solution is not limited to text only, but also allows the use of multimedia content.
- **User creation**: The tool enables user manual creation and generation for the exercise within the platform.
- **User creation automation**: The tool automates the user creation process, both individual and collective creation.

Table 1: State of the art - Feature comparison.

Features	CYRAN	Somulator	Prevency	Proposal
Multimedia content	✗	✓	✓	✓
User creation	✓	✓	✓	✓
User creation automation	✗	✓	✓	✓
User configuration	✗	✓	✓	✓
User auto-configuration	✗	✗	✓	✓
Pregenerated content	✓	✓	✓	✓
Self-published	✓	✓	✗	✓
Exercise configuration	✓	✓	✓	✓
Exercise automation	✗	✗	✗	✓
Hot configuration	✗	✓	✓	✓
Real platforms	✓	✓	✗	✓
Parallel control	✗	✗	✗	✓
Multiple simultaneous users	✓	✓	✓	✓
Psychological profiles	✗	✗	✗	✓
Behavior profiles	✗	✗	✗	✓
User relationships	✗	✗	✗	✓

– **User configuration**: The configuration of the profiles, advanced options, and privacy of the created users is contemplated.
– **User auto-configuration**: The created users can be auto-configured automatically by the tool.
– **Pregenerated content**: The solution allows the creation and configuration of publications and contents, which can be used during the execution of the cyberexercise.
– **Self-published**: Assigned content can be published automatically, either by prescheduling or dynamic publishing, without the need for direct human intervention.
– **Exercise configuration**: Exercises can be created and configured manually, adding users, publications, and other elements according to the requirements.
– **Exercise automation**: The process of creating social media exercises is fully or partially automated.
– **Hot configuration**: The exercise configuration can be modified while the exercise is running, adding or modifying elements such as users and publications.
– **Real platforms**: The solution uses existing social media platforms as the basis, not those specifically designed from scratch that are a software product rather than a social media platform.
– **Parallel control**: The tool is designed to monitor and manage multiple instances of the same social media exercise simultaneously.
– **Multiple simultaneous users**: Several users can connect to the same instance and operate at the same time.
– **Psychological profiles**: Psychological descriptions can be set up on users, allowing them to interpret incoming content and generate content based on their preferences.

- **Behaviour profiles**: Specific actions and behaviours can be configured for specific situations and events, such as publishing content, replying to publications and mentions, updating relationships, etc.
- **User relationships**: Basic relationships and connections between users created in social media can be configured, such as follow, block, mute, etc.

3 Proposal

This section outlines the article's proposal within the context of social media simulation exercises for cybercompetencies training. Concretely, a cybercompetency is a skill related to the secure use of social media platforms and the internet. Section 3.1 presents the social media training simulator and Section 3.2 its corresponding architecture.

3.1 Social Media Training Simulator

The proposal is centred around a tool that automates the process of creating and provisioning simulated cyberexercises in social media within the Cyber Range context and their advanced configuration. According to NIST a Cyber Range is an interactive platform to simulate all types of cybersecurity scenarios with their components with simulations of internet traffic or services. This virtualisation platform provides the opportunity to present and use a secure, controlled and legal environment to develop cybersecurity skills [16]. Typically, the structure of a Cyber Range is a composition of a user interface based on front-end technologies, the features that the Cyber Range can deploy or that are available (virtualisation, monitoring, simulation of services and internet, etc.) and finally, the Cyber Range infrastructure itself managed by the orchestrator (proper virtualisation, containerisation, emulation and simulation) [20]. Thus, the Cyber Range environment enables the generation of hyperrealistic simulations for cybercompetencies training in any scenario or situation, qualifying various social stratum and professional sectors against imminent threats and risks caused by, related to, or developed in social media. The training is carried out by three roles: students who perform the social media cyberexercise, instructors who design, prepare, and build the cyberexercise, and administrators who manage the Cyber Range. It is worth noting that automation is a crucial and fundamental new feature of the proposal. It facilitates the effortless, swift, and efficient generation of training, thereby fostering the development of competencies with minimal effort.

The proposed solution utilises open source social media platforms, concretely Mastodon[3], which serves as an alternative to Twitter. It is a social media platform based on microblogging similar to Twitter (now known as *X*), and the decision to use Mastodon for social media simulation is based on the advantages and features it offers, as well as the current social context, adapting perfectly to the needs and objectives of the proposal. The main advantages are its open source code, the number of libraries and functionalities it offers to build the necessary modules and functions, as well as its easy installation, configuration, and deployment. Also, the most valuable feature is

[3] `https://docs.joinmastodon.org/`

the automation capability for creating and configuring users, including the ability to skip authentication steps for new users by using the platform's own internal administration commands. Furthermore, using a Twitter-like platform allows students to learn in a close-to-reality environment, enhancing the training.

The following new innovative concepts presented in this proposal, combined with the improved foundational concepts, facilitate advanced simulations with hyperrealistic scenarios for various contexts.

- **Personality**: It attributes psychological descriptions and preferences. Such a feature aids in the interpretation and classification of incoming content, and generating automatic responses and posts that are aligned with the simulated user personality.
- **Behaviour**: It provides simulated users with actions and interactions, enabling them to perform tasks such as posting, and responding to notifications. It also allows the configuration and generation of concrete reactions to specific events, differentiating between target users by user type or current relationship.
- **Relationships**: It establishes connections between users, allowing them to form and accept friendship requests, follow, block, or mute other users, and accept or decline follow requests.
- **Profile configuration**: It provides simulated users with profile and privacy settings. It is used to customise the simulated users' profiles with biographies, aliases, profile pictures, other privacy options, and other features.

Besides, several enhancements are proposed through process automation. These automations are achieved using superparameters and base templates. Templates used to generate exercises and the required configurations for provisioning the simulation. The templates or base templates are guides or JSON-format models containing elements with the configuration and provisioning of a simulation (see Section 4, Table 3), and superparameters are parameters used to determine the automated construction and configuration of a base template (see Table 2). Concretely, proposed automations include:

1. **Creation and configuration of social media cyberexercises**: The creation of cyberexercises is automated, allowing the instructor to obtain a usable, configured base template with a few easy-to-use superparameters, such as the type of exercise (disinformation, phishing, social engineering attacks, and more), the topic (COVID-19 pandemic, health crisis, and others.) or the number of simulated interactions.
2. **Advanced configuration of cyberexercises**: The basic templates generated can be configured at different levels of detail, with infinite possibilities thanks to the new concepts proposed. Thus, the configuration of specific user profiles, new relationships between users, specific complex actions, and much more are possible.
3. **Creation and customisation of users**: User creation is automated, distinguishing between relevant and random users, for example, the configuration of the novel concepts previously introduced (i.e. profiles, relationships, personalities, and behaviours) is automated.
4. **Cyberexercise content**: The generation of content and key publications necessary for the development of the cyberexercise is automated, and classified into types according to the topic, exercise type, or other superparameters. When creating base templates for a concrete exercise, compatible template elements are selected and

filtered by the specified superparameters. This way, choosing a disinformation exercise will return publications, users, personalities and behaviours needed to develop a disinformation exercise.

5. **Simulated Non Playable Character (NPC) content**: Interactions, content generation, and posts used as background traffic are automated to animate the simulation. An NPC could follow another user, post a photo, and more. In this way, a real user the social media is completely authentic with NPC users talking or posting.

In addition, it is relevant to note that the tool is designed to be integrated into a Cyber Range and therefore needs to be applied in multiple cyberexercises instances. Also, in particular, deployments of the same cyberexercise should be consistent for all students, with the base configuration being identical, differing only in the dynamic background traffic. This way, the training of multiple students can be easily controlled by varying unimportant background traffic while maintaining the key features of the cyberexercise. Similarly, the tool should allow the instructor to interact with the simulations during the cyberexercise, whether for individual or group deployments. The Hot Configuration will be explained in the section 3.

3.2 Architecture

The proposed architecture is based on the typical structure of a Cyber Range, expanding upon the architecture proposed in the COBRA Cyber Range [11] to focus on the development of features for social media training modules. The tool's design, based on independent systems and modules from a specific Cyber Range, renders it exceedingly extensible and autonomous, despite reusing the architecture from the Cyber Range COBRA proposal. Fig. 1 illustrates the proposed architecture, and the following sections describe and detail its components.

Front-end Architecture It is an architecture based on Docker containers reused from COBRA. The Fig. 1 shows the systems to which the front-end should be connected. As a result of these connections with the managers, it is abstracted from the inclusion of new modules into the systems to raise the functionality, and it is very easy to change the front-end with other options.

Scenarios, Challenges and Cyberexercises Creation System The objective of the Creation System is to provide the necessary configurations and capabilities to Virtual Machines (VMs) that will be used to carry out the training. Subsequently, the System will deploy the VMs, and configure the challenges and cyberexercises. A challenge is a task that students need to overcome. Moreover, the choice of features and parameters of social media, where the training is conducted, is determined here. A Cyberexercise is a series of one or more challenges that may or may not share the same training objective. Each instance of a Cyberexercise is referred to as a deployment. The configurations of virtual machines and social media platforms are carried out here.

Firstly, the front-end communicates with the Creation Manager, which handles the requests for the creation of scenarios, challenges, and cyberexercises. Secondly, the

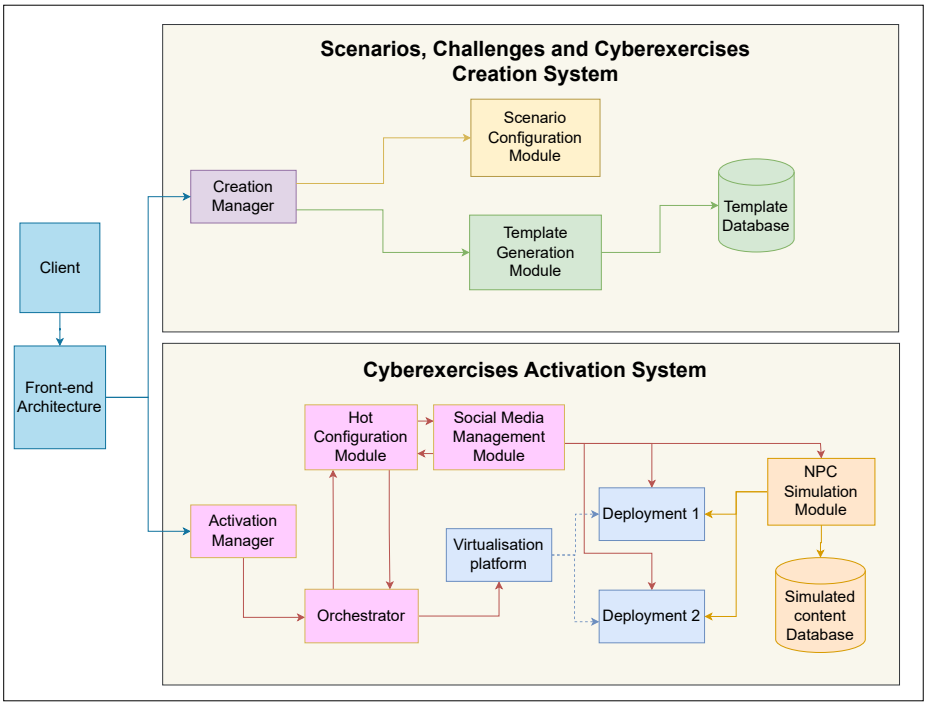

Fig. 1: Architecture of the social media simulator

manager forwards all information in the requests to the Scenario Configuration Module (SCM). Finally, it communicates with the Template Generation Module (TGM) to obtain the final challenge template.

The SCM is responsible for creating scenarios with one or multiple VMs. It is possible to configure parameters such as the operating system, services or applications, files that should be present on the machine, and the characteristics of the VM itself. To ensure the configuration is modular and extensible, services are implemented as *Docker* containers, and any type of file can be sent to the VMs. For social media, training is mandatory for at least one of the VMs to run the Mastodon service.

Following the scenario creation, the instructor proceeds to create and configure the challenge using superparameters explained in Section 4. The TGM will provide a JSON template containing the complete configuration of the challenge, including instructions for configuring the instance, which will be loaded into Mastodon after activation.

For generating templates, the TGM possesses a database with previous templates with social media content generated based on topics, kinds of training, objectives, and so forth. It also has configurations for personalities and behaviours, automation for user parameterisation and deployment, number of posts and much more.

Cyberexercises Activation System The Activation System is designed to guarantee suitable deployment, activation, and execution of cyberexercises. Either an instructor

or student of the training user can access this system. Similar to the Creation Manager, the front-end communicates with the Activation Manager, which handles requests and allows the activation/deactivation of cyberexercises. Furthermore, it enables the connection between the student and the virtual environment where the training will take place.

On the one hand, if access is granted with the instructor role, it is possible to start the cyberexercise. The orchestrator automates processes related to virtual machine creation and management, and performs communication and management tasks for different modules within the cyber range. It requests the virtualization platform to deploy the VMs and provisions them with the needed operating system and services, such as Mastodon. Once the deployment is successful, the orchestrator contacts the Hot Configuration Module to send the cyberexercise base template to prepare the initial state of the challenge.

It is worth noting that each deployment has its Social Media Management Module (SMMM) and NPC Simulation Module, which run locally in the VM where Mastodon is located. The SMMM module is responsible for ensuring that the training can be carried out. The SMMM module guarantees the execution of cyberexercises, making it the foremost module with extensive functionality. Specifically, it updates and manages exercise information dynamically. The SMMM receives and analyses the template, generates the configuration, and loads it into the local Mastodon instance. Once loaded, it continues to monitor the health of the social media platform.

On the other hand, if a student accesses, it could connect to the VMs via remote desktop applications. Once connected to the VM, the student can start the cyberexercise to train, and, thus develop the competencies specified by the instructor. This cyberexercise can be conducted for the duration specified in the configuration file.

Additionally, the NPC simulation module is employed to bring to life the social media simulation. This module generates simulated traffic, meaning that it creates and publishes content related to the personalities of the NPCs, and creates interactions among simulation users. Similarly to the database in the Template Generation module, the NPC simulation module also possesses a database with previously generated content. This database is accessed by the NPC simulation module to generate automatically responses or publications made by the NPC users using the pregenerated content based on their personality, behaviour and desired topic.

4 Implementation

This section explores the general implementation of the proposal, following the processes previously described from the creation of scenarios and challenges, the activation of a cyberexercise and its execution. These processes include the modules discussed in Section 3, whose functionalities allow automating and simplifying the proceedings of generation and configuration of the social media platform instances to generate more efficient, effective and realistic cyberexercises.

The following subsections explain the modules used in the proposal. Subsection 4.1 details the template generation module, which automates the creation of social media exercises, while Subsection 4.2 shows the Social Media Management Module. Then,

Table 2: Superparameter description.

Superparameter	Definition	Objective	Required	Dependencies	Values
Topic	Main theme of the exercise used for relevant content	Generate a themed template for the exercise.	✗	✗	Text
Type	Type of exercise or use case to be generated	Generate a base template for a specific use case	✓	✗	Use case
Subtype	Subtype of exercise or specific use case to be generated	Generate a base template for the implementation of a specific use case	✗	Type	Specific use case
Users amount	Simulated users amount	Generate a random users amount	✗	✗	Positive integer
Simulated traffic	Simulated traffic frequency and amount	Configure the amount and frequency of interactions generated by NPC	✗	✗	Percentage
Topic divergence	Divergence between simulated traffic topic and main topic	Configure the emergence of topics separate from the main topic	✗	✗	Percentage
User divergence	Divergence between user behaviour and user personality	Generate users with different behaviour and personality	✗	✗	Percentage
Configuration level	Automatic template configuration	Generate templates at various configuration levels	✗	✗	Percentage
Bot amount	Percentage of automated bots	Generate a percentage of bot accounts	✗	Automation level	Percentage
Automation level	Humanity of bots	Configure bot humanity level	✗	Bot amount	Percentage

in Subsection 4.3, the Hot Configuration Module is explained, and finally, in Subsection 4.4, the NPC Simulation Module is described.

4.1 Template Generation Module

The first step is to generate and configure a base template (as described in Section 3.1), which will be used to provision the social media instances. These templates are typically generated manually, but such a procedure requires extensive knowledge of the tool due to the extensive configuration possibilities. To simplify and automate this creation process, the Template Generation Module is provided, enabling the template generation by using a few superparameters, as explained in Section 3.1. This module consists of a database that maintains fragments of templates and permits the generation of a base template according to these superparameters (used as filters). Table 2 shows the superparameters, their definition, their use, whether they are required or mandatory, the dependencies between other superparameters, and the value that must be assigned (e.g. the superparameter *type* refers to filter of the type of exercise required, such as a disinformation exercise, a phishing scenario, etc., used to classify and retrieve compatible template elements. The superparameter is always required, it has no dependencies on other superparameters, and valid values are the defined exercise types in the filter).

Based on the superparameters chosen by the instructor, the cyberexercise TGM retrieves automatically all compatible fragments from the database, and then it builds the base template by adapting it to the specifications, automating the process. The recurring elements of a base template are displayed in Table 3, together with their description and usage (e.g. the template element *publications* is a JSON entry list of publications, where each publication has its content, used to create and schedule a user's pregenerated posts on the social media platform). These received templates must be reviewed by the instructor, who will decide if it is valid or if they must be manually modified for the cyberexercise (Section 3.2).

4.2 Social Media Management Module (SMMM)

Once the cyberexercise is activated, the templates configured by the instructor are sent and loaded into the selected deployments within the virtual environment. The SMMM receives and maps the templates to objects that store and work with the information provided. It also executes the necessary actions and procedures to provide the social media instances with the template configurations and update the information for its internal management.

The procedure followed when loading a template is divided into two phases. The first phase consists of receiving and processing the base template, traversing it and its elements to convert it into objects. Behaviour and personality lists are loaded, and relevant and random users are assigned. In the second phase, all User objects are traversed, and the simulation is provisioned with the assigned configuration. The provisioning process in the second phase is divided into several stages:

1. **User creation and profile configuration**: Users are created and individual sessions with the social media platform are generated. Profiles are then configured.

Table 3: Template elements.

Template element	Definition	Use
Predefined templates	Previously created base templates	Loads pre-existing configured exercise base templates
Random users	Configurations for user creation	Set up and create different numbers of NPC users with common characteristics
Relevant users	Configurations for relevant user creation	Set up and create users configured with the given characteristics
Profile configuration	Specific profile configuration for users	Set up users profile configuration
Relationships	Configurations for user relationships	Set up base relationships as add, block, mute requests
Publications	List of user publications with assigned content	Create pregenerated publications and post them
Behaviour	Permitted action configuration	Set up user actions and behaviours for publishing content and interact
Behaviour functions	NPC authorised actions	Set up NPC user actions
Personality	User psychological description	Set up user personality for interpretation and content generation

2. **Relationship generation**: For each user follow, mute, block and other requests are made. Then, for each user again, received requests are accepted or rejected, updating the current relationships.
3. **Content publication**: For each user assigned pregenerated publications are sent.
4. **User simulation**: For each user, a thread is created to run the simulated traffic, based on the user's personality and behaviour. The execution of these threads is included in the NPC Simulation Module.

4.3 Hot Configuration Module

While the cyberexercise is running and deployments are operational, it may be necessary to make changes to these instances, for which we need a hot configuration module. This module facilitates modifying the running social media simulations. To achieve this, several mechanisms must be developed to enable the addition of new elements through the use of templates (new personalities, behaviours, users, and publications or content). These additions can be applied to specific or grouped deployments. Furthermore, the module should provide control actions that can be sent to influence these simulations. Some control actions can go directly to social network functioning, background traffic, NPC or relevant users, etc.

4.4 NPC Simulation Module

This module permits simulating users their personalities and their assigned behaviours. Specifically, personalities dictate the user's *personality*, determine how the simulated

users will interpret a particular message, comment or topic, and respond based on their tastes and interests (whether they agree or disagree, if they insult, support, etc.). Then, behaviours dictate the actions they can take to generate background traffic, such as replies to public content, private notifications, etc.

The module consists of running threads for each user, which perform a looping process, selecting actions by categories according to the functionality required in the behaviour. Firstly, the publication actions are performed, both public and private. Then the actions related to the notifications received are performed. Next, the public publications are processed, and finally, the control and internal update actions are executed.

5 Demo

This section shows the result of provisioning a template to demonstrate that the proposal can automate the deployment of a social media exercise, as well as the interactions and actions that can be performed by simulated users within the platform.

In this template, three relevant manually configured users (*DrSigmunFraud, DefinitelyNotAHacker, and Fake News Channel*) and six other random autoconfigured users (they will be autoconfigured as *Samuel29, emorris, michaelmiller, zacharyperry, wcole, jessicaalexander*) are declared. Particularly, several behaviours and personalities are also configured to enable the autoconfigured users to generate automatic actions within the platform, such as replying to new messages, boosting posts, adding them to favourites, following users etc. Moreover, some users have declared predefined publications and relationships with other users.

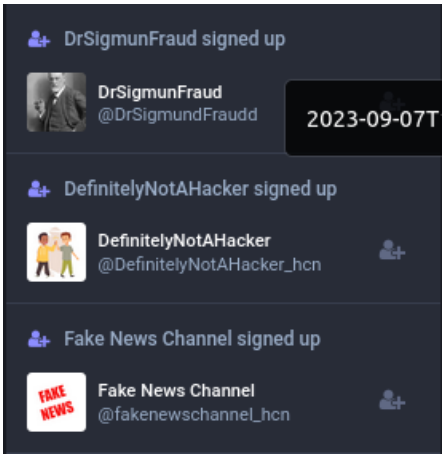

Fig. 2: Relevant users generation.

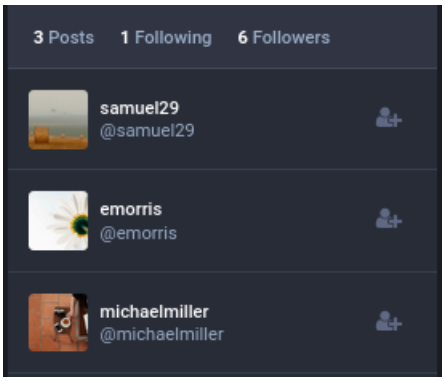

Fig. 3: Creating relationships

Firstly, Fig. 2 presents the creation of the three manually configured users (*DrSigmunFraud, DefinitelyNotAHacker, and Fake News Channel*), who have been provided with profile pictures, descriptions, publications, and relationships manually.

Secondly, Fig. 3 reveals some generated relationships, where *samuel29, emorris, and michaelmiller* and some users follow the shown account (*Fake News Channel*). In addition, the behaviour declared in the template may also generate follow, block, or mute relationships.

Then, Fig. 4 displays the self-configured random creation of six users, resulting in realistic profiles of a social media platform, while Fig. 5 reveals an autoconfigured user profile (*samuel29*) showing various user statistics. These statistics are related to the assigned behaviour, as it allows replying to posts, marking them as favourites, following other users, etc., automatically. The figure highlights its profile images, biography, boosted posts, replies, and followed users.

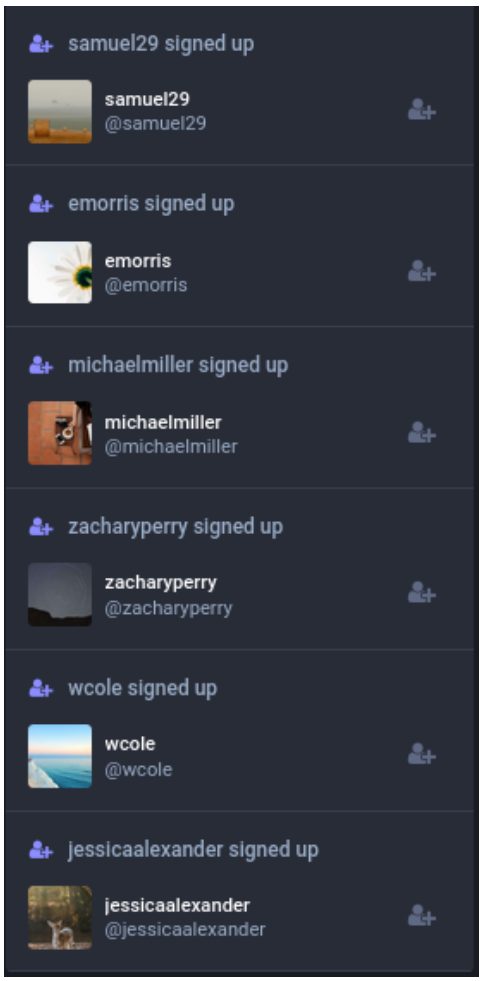

Fig. 4: Random user generation.

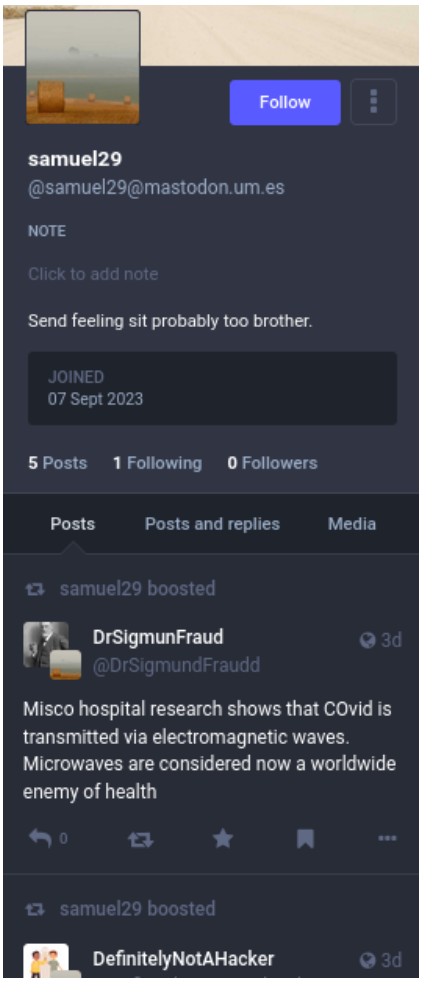

Fig. 5: User interactions.

Finally, Fig. 6 illustrates some responses and interactions of a disinformation fake new about COVID-19, which could be part of a disinformation campaign. Here it is possible to observe those user behaviours that have generated several responses to a posting event of another user they follow. Moreover, if we look at the responses of users jessicaalexander and zacharyperry, a more complex response action can be observed, which corresponds to a mention to the original user of the post.

Fig. 6: Behaviour actions and replies.

To sum up, with this demo, it is shortly demonstrated that the proposal of this article can generate both concrete users manually and random users, individually or collectively, in an automatic way. Then, the setup of posts enables the design and advancement of particular cyberexercises. The utilization of specific behaviors or actions

enhances the realism and functionality of the social media platform, enabling the creation of highly realistic exercises and scenarios. This, in turn, facilitates the training of cybersecurity competencies within social media platforms.

6 Conclusions and Future Work

This article introduces a proposal designed for training cybersecurity competencies on social media platforms. The proposed tool is prepared for integration into a Cyber Range. Specifically, this tool aims to make the training and education of professionals more efficient, achieving this by presenting the necessary architecture for proper deployment and automation in generating and managing cyberexercises. Also, the Cyber Range framework allows for simulations of adaptable difficulty levels and learning objectives, catering to competencies, topics, and final goals. The proposal is flexible, providing the capability to generate highly configurable, extensible, and multidisciplinary challenges. This involves an automated simulation generation feature utilizing the foundational templates outlined in Section 4 as a basic or customisable element. Furthermore, the proposal introduces the option to incorporate NPC users with specific characteristics, such as the innovative concepts of behavior, personality, and relationships, adding functionality, depth and realism to the training environment.

To demonstrate the feasibility of the proposal, a demonstration showcasing its functionality is provided in Section 5. It illustrates the creation of two types of users, the publication of social media content, the generation and building of relationships between users, and the configurations and behaviors that contribute to making a social media platform realistic. Additionally, it highlights the automatic execution of actions essential for creating diverse cyberexercises.

Looking ahead to future work, there are several key areas of focus. These include exploring the integration of other open-source social media platforms to simulate highly known platforms like Facebook, YouTube, Telegram, etc., thereby making the proposal more versatile. Furthermore, it is proposed to design and create an automated module for dynamic content generation. This module would enable the obtaining of existing content and the generation of new content, aligned with the requirements of behavior and personality. Finally, there is a plan to develop a module for the collection and evaluation of student data, aiming to enhance the efficacy and adaptability of the training platform.

In conclusion, the ongoing development and expansion of the proposed training tool for cybersecurity competencies on social media platforms promise to bring about a more versatile, adaptive, and effective solution for professionals, organisations in the field of health or other sectors seeking comprehensive training experiences related to the social part of the internet.

Acknowledgements

This work has been partially funded by the Spanish Ministry of Universities linked to the European Union through the NextGenerationEU programme, from the postdoctoral

grant Margarita Salas (172/MSJD/22). Authors acknowledge as well support from the CybAlliance project (Grant no. 337316).

References

1. Amblard, F., Bouadjio-Boulic, A., Gutiérrez, C.S., Gaudou, B.: Which models are used in social simulation to generate social networks? a review of 17 years of publications in jasss. In: 2015 Winter Simulation Conference (WSC). pp. 4021–4032. IEEE (2015)
2. Braidley, S.: Extending our cyber-range cyran with social engineering capabilities. Master's thesis. De Montfort University, Leicester, England (2016)
3. Colwill, C.: Human factors in information security: The insider threat–who can you trust these days? Information security technical report **14**(4), 186–196 (2009)
4. Denecke, K., Bamidis, P., Bond, C., Gabarron, E., Househ, M., Lau, A., Mayer, M.A., Merolli, M., Hansen, M.: Ethical issues of social media usage in healthcare. Yearbook of medical informatics **24**(01), 137–147 (2015)
5. Drahošová, M., Balco, P.: The analysis of advantages and disadvantages of use of social media in european union. Procedia Computer Science **109**, 1005–1009 (2017)
6. ENISA: Health Threat Landscape. Technical report, ENISA (July 2023). https://doi.org/10.2824/163953, https://www.enisa.europa.eu/publications/health-threat-landscape
7. Geeng, C., Yee, S., Roesner, F.: Fake news on facebook and twitter: Investigating how people (don't) investigate. In: Proceedings of the 2020 CHI conference on human factors in computing systems. pp. 1–14 (2020)
8. González-Padilla, D.A., Tortolero-Blanco, L.: Social media influence in the covid-19 pandemic. International braz j urol **46**, 120–124 (2020)
9. Hadlington, L.: Human factors in cybersecurity; examining the link between internet addiction, impulsivity, attitudes towards cybersecurity, and risky cybersecurity behaviours. Heliyon **3**(7), 18 (2017)
10. Naeem, S.B., Bhatti, R., Khan, A.: An exploration of how fake news is taking over social media and putting public health at risk. Health Information & Libraries Journal **38**(2), 143–149 (2021)
11. Nespoli, P., Albaladejo-González, M., Pastor Valera, J.A., Ruipérez-Valiente, J.A., Gómez Mármol, F.: Capacidades avanzadas de simulación y evaluación con elementos de gamificación. In: VII Jornadas Nacionales de Investigación en Ciberseguridad (JNIC '22). pp. 55–62 (2022)
12. Nifakos, S., Chandramouli, K., Nikolaou, C.K., Papachristou, P., Koch, S., Panaousis, E., Bonacina, S.: Influence of human factors on cyber security within healthcare organisations: A systematic review. Sensors **21**(15), 5119 (2021)
13. Norwegian University of Science and Technology, Norwegian Defense Research Establishment: Somulator. https://www.ntnu.no/ncr/somulator and https://www.ffi.no/forskning/prosjekter/somulator, accessed: September 15, 2023
14. Pastor-Galindo, J., Nespoli, P., Gómez Mármol, F., Martínez Pérez, G.: The not yet exploited goldmine of osint: Opportunities, open challenges and future trends. IEEE Access **8**, 10282–10304 (2020). https://doi.org/10.1109/ACCESS.2020.2965257
15. Pastor-Galindo, J., Zago, M., Nespoli, P., Bernal, S.L., Celdrán, A.H., Pérez, M.G., Ruipérez-Valiente, J.A., Pérez, G.M., Mármol, F.G.: Spotting political social bots in twitter: A use case of the 2019 spanish general election. IEEE Transactions on Network and Service Management **17**(4), 2156–2170 (2020). https://doi.org/10.1109/TNSM.2020.3031573

16. Petersen, R., Santos, D., Smith, M.C., Wetzel, K.A., Witte, G.: Workforce Framework for Cybersecurity (NICE Framework). Tech. rep., NIST (2020), https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-181r1.pdf
17. Prevency: Project NATO Stratcom COE. https://prevency.com/en/projects/project-nato-stratcom-coe/, accessed: September 15, 2023
18. Rizzoni, F., Magalini, S., Casaroli, A., Mari, P., Dixon, M., Coventry, L.: Phishing simulation exercise in a large hospital: A case study. DIGITAL HEALTH **8**, 20552076221081716 (2022). https://doi.org/10.1177/20552076221081716, pMID: 35321019
19. Seh, A.H., Zarour, M., Alenezi, M., Sarkar, A.K., Agrawal, A., Kumar, R., Ahmad Khan, R.: Healthcare data breaches: insights and implications. In: Healthcare. vol. 8, p. 133. MDPI (2020)
20. Ukwandu, E., Farah, M.A.B., Hindy, H., Brosset, D., Kavallieros, D., Atkinson, R., Tachtatzis, C., Bures, M., Andonovic, I., Bellekens, X.: A review of cyber-ranges and testbeds: Current and future trends. Sensors **20**(24) (2020). https://doi.org/10.3390/s20247148, https://www.mdpi.com/1424-8220/20/24/7148
21. Zafar, H.: Cybersecurity: Role of behavioral training in healthcare. In: Americas Conference on Information Systems (2016)