Interaction Humain Machine avec un Agent Cognitif automatisant une plateforme militaire de Tests d'Intrusion

John Pyrgies, Justine Pyrgies, Théo Pyrgies

▶ **To cite this version:**

John Pyrgies, Justine Pyrgies, Théo Pyrgies. Interaction Humain Machine avec un Agent Cognitif automatisant une plateforme militaire de Tests d'Intrusion. IHM'24 - 35e Conférence Internationale Francophone sur l'Interaction Humain-Machine, AFIHM; Sorbonne Université, Mar 2024, Paris, France. hal-04493665

HAL Id: hal-04493665
https://hal.science/hal-04493665v1

Submitted on 13 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Interaction Humain Machine avec un Agent Cognitif automatisant une plateforme militaire de Tests d'Intrusion

Human Centered Interaction with a Cognitive Agent automating a military Penetration Testing platform

JOHN PYRGIES
SkyAngels Lab, SkyAngels SRL, Belgique, john.pyrgies@skyangels.eu

JUSTINE PYRGIES
SkyAngels Lab, SkyAngels SRL, Belgique, justine.pyrgies@skyangels.eu

THÉO PYRGIES
SkyAngels Lab, SkyAngels SRL, Belgique, theo.pyrgies@skyangels.eu

Sur base d'un des thèmes de l'appel à projets Fonds Européen de la Défense 2023 qui couvrait l'automatisation, avec de l'Intelligence Artificielle, des Tests d'Intrusion militaires, nous avons réalisé un état de l'art sur les associations [IHM-Cybersécurité] et [IHM-Agent Cognitif]. Nous avons étudié HCISec qui améliore l'utilisabilité des dispositifs de Cybersécurité. Nous avons effectué l'Analyse Socio-Technique d'un Agent Cognitif destiné à automatiser ces Tests d'Intrusion militaires . Nous avons énoncé un processus de Conception Centré Utilisateur adéquat et l'avons inséré dans un Cycle de Vie de Développement Logiciel AgileUX. Nous avons ensuite étudié les Interactions Humain-Machine entre l'acteur humain et cet Agent Cognitif sur base d'un Cas d'Utilisation. Nous avons conçu l'architecture logicielle basée sur l'Architecture Cognitive Soar et un Module de Contrôle de tests d'intrusion exposant des Interfaces Humain-Machine. Nous avons enfin décrit ces interfaces consistant en 3 écrans principaux et la navigation entre ces écrans.

CCS CONCEPTS • Security and privacy→Intrusion/anomaly detection and malware mitigation • Human-centered computing → Human computer interaction (HCI)• Computing methodologies→Artificial intelligence→Distributed artificial intelligence

Based on a theme of the 2023 European Defence Fund (EDF) addressing the automation, with Artificial Intelligence, of military Penetration Testing, we have carried a state-of-the-art upon the [HMI-Cybersecurity] et [HMI-Cognitive Agent] associations. We have studied HCISec which aims at improving the usability of Cybersecurity mechanisms and tools. We have performed the Socio-Technical Analysis of a Cognitive Agent aimed at automating military Pen Tests. We have defined a User Centered Design process and embedded it within an AgileUx Software Development Life Cycle. We then studied the Human-Machine Interactions between the Human Actor and this Cognitive Agent based on a Use Case. We have then designed the software architecture based on the Soar Cognitive Architecture and a Pen Tests control module exposing Human-Machine Interfaces. At last, we have described those interfaces as 3 screen layouts and the navigation flow between those screens.

Mots-clés additionnels : Interaction Humain-Machine (IHM), Cybersécurité, HCISec, Agent Cognitif, Intelligence Artificielle (IA), Analyse Socio-Technique, Conception Centrée Utilisateur, Cyberdéfense, Centre Opérationnel de Sécurité.

Reference : Pyrgies, J., Pyrgies, J., Pyrgies, T. 2024. Interaction Humain-Machine avec un agent cognitif automatisant une plateforme militaire de tests d'intrusion. Dans IHM'24: Articles Industriels de la 35ème Conférence Internationale Francophone sur l'Interaction Humain–Machine, Mars 25–29, 2024, Paris, France.

1 INTRODUCTION

Les Cyberattaques menées avec succès contre les Systèmes d'Information des entreprises privées et des organisations publiques font régulièrement la une de nos journaux. Dans cet environnement où la menace est omniprésente, les Centres des Opérations de Sécurité (SOC) ont recours aux Tests d'Intrusion afin de détecter et d'exploiter 'éthiquement' les vulnérabilités de sécurité de leurs réseaux, systèmes d'exploitation, intergiciels et logiciels applicatifs. Ces Tests d'Intrusion effectués tant bien avant la mise en service de ressources informatiques que régulièrement au cours de leur exploitation, se déroulent suivant des processus standardisés, dont le 'Penetration Testing Execution Standard' [1], qui comporte les étapes suivantes : *Le pré-engagement*, la *collecte de renseignements*, la *modélisation de la menace*, l'analyse de *vulnérabilité*, l'*exploitation*, *la post exploitation* et le *rapport*. Ces Tests d'Intrusion nécessitent du temps et mobilisent des compétences

humaines rares et des moyens techniques non négligeables. L'automatisation des Tests d'Intrusion est dès lors opportune et a fait l'objet d'articles scientifiques [2] ou de thèses doctorales [3] qui mettent en œuvre des techniques d'Intelligence Artificielle à cette fin.

Les ressources informatiques des Forces Armées sont particulièrement complexes : Elles combinent des Technologies de l'Information (IT), comme un Système d'Information donnant la vue complète d'un théâtre d'opération, avec des Technologies Opérationnelles (OT), comme un Système de Contrôle des radars d'une base aérienne, ainsi que des Systèmes Embarqués Internet des Objets Défense (DIoT), comme des capteurs à bord d'un drone de combat. Les Forces Armées sont également la cible de Cyberattaques avec une menace aggravée par la puissance, en termes de compétences, de moyens financiers et de technologies, des Cyber agresseurs internes/externes qui, dans un contexte de Cyberdéfense, peuvent élaborer des 'menaces persistantes avancées' (APT) particulièrement difficiles à détecter car extrêmement sophistiquées et exploitant typiquement des vulnérabilités 'du jour zéro' (i.e. non publiées dans la communauté informatique et/ou sans correctif disponible).

Conscientes de ces enjeux, l'Agence Européenne de Défense (EDA) a constitué le groupe de travail CapTech Cyber dont un des projets de recherche MASFAD 2 a abouti sur un prototype de système multi-agents pour la détection d'APT [4] et la Commission Européenne lance régulièrement, à travers son programme du Fonds Européen de la Défense (EDF) 2021-2027 des appels à projets en Cybersécurité militaire (CYBER). Lors de l'appel EDF 2023, un des thèmes couvrait notamment l'automatisation, avec de l'Intelligence Artificielle, des Tests d'Intrusion effectués par les administrateurs système eux-mêmes pour mettre à l'épreuve leurs réseaux d'ordinateurs militaires [5].

Le cahier des exigences succinct de cet appel à projets lié aux Tests d'Intrusion militaires automatisés sert de base au présent article élaboré par le SkyAngels Lab, laboratoire de recherche privé, qui a lancé le programme de recherche CollosusGuardian visant à exploiter l'Intelligence Artificielle, et particulièrement les Architectures Cognitives, pour automatiser les tâches liées à la Cyberdéfense au sein des forces armées Européennes. L'objectif n'est pas de remplacer l'Humain dans ces tâches Cyber mais bien de créer des équipes d'Agents Cognitifs et de Cyberanalystes Humains pour augmenter l'efficience et l'efficacité de ceux-ci dans leur mission.

Dans cet article, nous nous focalisons donc sur les Interactions Humain-Machine d'une plateforme militaire de tests d'intrusion automatisés avec un Agent Cognitif. Le projet de recherche est au stade de l'analyse et de la conception de l'architecture logicielle et ses premiers livrables, décrits dans cet article, serviront de base à la recherche de partenaires industriels et académiques pour élaborer des dossiers de candidatures à des appels à projets futurs du Fonds Européen de la Défense (EDF) couvrant particulièrement les applications innovantes de l'Intelligence Artificielle pour les SOCs militaires.

2 INTERACTION HUMAIN-MACHINE ET CYBERSECURITÉ : HCISEC

Une de ces exigences spécifiait que la 'solution logicielle devait être conviviale' (*user-friendly software solution*). Si la grande majorité des cahiers des exigences de logiciels spécifient des exigences d'utilisabilité en ligne avec les standards idoines [6] et les bonnes pratiques de l'Ingénierie des Exigences [7] qui, selon le standard ISO 9241-11:2018 [8], *permettent aux utilisateurs d'atteindre leurs objectifs de manière efficace, efficiente et satisfaisante, en tenant compte du contexte d'utilisation,* il y a un autre enjeu lié à l'utilisabilité des interfaces impliquées dans les Interactions Humain-Machine des fonctions (identification, authentification, autorisation, chiffrement…) et outils (pare-feu, anti-virus, infrastructure à clé publique, tests d'intrusion…) de Cybersécurité : La Cybersécurité elle-même !

En effet, jusqu'à présent la réponse des cyberdéfenseurs à la recrudescence des Cyberattaques a été essentiellement technologique. Cela a une double conséquence : Tout d'abord, les Cyberattaquants se sont adaptés en ciblant le nouveau maillon faible de la chaine de Cybersécurité : l'Humain, à travers des Cyberattaques d''Ingénierie Sociale' basées sur la manipulation psychologique des utilisateurs finaux, des administrateurs et même des cyberdéfenseurs du système informatique. Ensuite, l'apport de nouvelles technologies de Cybersécurité plus complexes rend leur usage sécurisé plus difficile à cause de l'incompréhension et du manque d'adhésion de leurs utilisateurs finaux, de leurs administrateurs voire de leurs cyberdéfenseurs et il est donc essentiel d'analyser avec soin les Interactions Humain-Machine des dispositifs (fonctions, outils…) de Cybersécurité.

Ainsi, déjà en 1883, Auguste Kerckhoffs, dans son ouvrage dédié à *La cryptographie militaire* [9] spécifiait que *le Système* (de cryptographie militaire) *soit d'un usage facile, ne demande ni tension d'esprit, ni la connaissance d'une longue série de règles à observer*. [10] nous apprend que lorsque les professionnel(le)s de la Cybersécurité sont confronté(e)s à des interfaces utilisateur de logiciels de Cybersécurité peu intuitives et réactives, leur capacité à terminer leurs tâches et leur motivation diminuent au point d'influencer négativement le niveau de Cybersécurité de leur organisation tout entière. Dans la littérature, le concept d'HCISec s'est développé et a pour origine le besoin identifié par les experts en IHM d'améliorer l'utilisabilité des dispositifs de Cybersécurité. [11] fournit un modèle d'analyse de la menace sécurité-utilisabilité, dans ce contexte HCISec, qui se focalise sur les erreurs non-malveillantes des utilisateurs légitimes du Système pouvant compromettre sa sécurité. Le modèle compare la difficulté de réaliser les scénarios d'usage désirés avec la facilité de réaliser les scénarios de menaces non désirés. Cette méthode prend comme postulat que les utilisateurs vont suivre le chemin de la facilité pour atteindre leurs objectifs, même si cela induit des risques de Cybersécurité.

3 INTERACTION HUMAIN-MACHINE ET AGENT COGNITIF

[12] nous définit le concept d'agent selon : '*Tout ce qui peut être vu comme percevant son environnement et agissant sur cet environnement*' et 'd'*agent rationnel qui pour chaque séquence de perception sélectionne l'action qui maximise sa mesure de performance compte tenu de ses connaissances*'.

[13] définit la '*confiance en l'automatisation*' comme '*la perspective qu'un agent technologique contribuera aux objectifs d'un acteur humain dans une situation caractérisée par l'incertitude et la vulnérabilité*' et décrit comment celle-ci peut être prise en compte dans le processus itératif de Conception Orientée Utilisateur d'un outil de Cybersécurité destiné à l'usage de professionnels(les) de la Cybersécurité pour élucider et spécifier ses exigences d'utilisabilité. L'agent technologique interagit avec l'acteur humain dans le cadre d'une Interaction Humain-Machine qui évolue, par anthropomorphisme, en fonction de la capacité de l'agent technologique d'interagir avec l'acteur humain de manière naturelle et interactive ainsi que de participer aux processus de prise de décision en équipe. La '*confiance en l'automatisation*' doit avoir un niveau approprié, pour ce faire, il faut identifier son périmètre, définir les conditions associées aux différents niveaux de confiance qu'on lui accorde, décrire les interactions que doit avoir l'acteur humain avec l'agent technologique en fonction de ces niveaux de confiance et former l'acteur humain en conséquence. La '*confiance en l'automatisation*' dépend de la nature de la tâche automatisée, des caractéristiques émotionnelles (satisfaction de l'agent technologique), cognitives (niveau d'expertise en Cybersécurité et compréhension de l'agent technologique) et de trait de caractère (propension à faire confiance) de l'acteur humain et de l'utilisabilité et de la fiabilité de l'agent technologique.

Dans [14], on lit qu''*Un agent cognitif est un agent dont la conception se fonde sur des propriétés que l'on attribue habituellement aux êtres humains*' comme la prise de décision, l'apprentissage, le raisonnement… et qu'un *Système de Confiance* permet à un utilisateur d'évaluer le niveau de confiance vis-à-vis d'un système multi-agents sur base des interactions antérieures de l'utilisateur avec ces agents et des recommandations que certains agents fournissent à propos d'autres agents. Il existe 2 types de *Système de Confiance* : Les modèles logiques basés sur les logiques modales et les modèles numériques basés essentiellement sur les probabilités. L''*agent cognitif*', tel que défini dans [14], est bien le concept dont nous avons besoin pour jouer le rôle d'agent technologique dans les dispositifs (fonctions, outils…) de Cybersécurité.

4 QUESTIONS DE RECHERCHE

Ayant introduit le contexte de notre démarche et effectué un état de l'art sur les associations [IHM-Cybersécurité] d'une part et [IHM-Agent Cognitif] d'autre part, nous sommes maintenant en mesure d'énoncer les questions de recherche que va tenter d'élucider le présent article :

1. Quel est le contexte Socio-Technique de l'Agent Cognitif destiné à automatiser les tests d'intrusions militaires ?
2. Quel Processus de Conception Centrée Utilisateur (UCD) pour développer un tel Agent Cognitif ?
   2.1. Comment s'inscrit ce Processus UCD dans son Cycle de Vie du Développement Logiciel (SDLC) ?
3. Quelles sont les Interactions Humain-Machine entre l'Acteur Humain et l'Agent Cognitif ?
   3.1. Quels est l'objectif de l'Acteur Humain ?

3.2. Quelles sont les Interfaces exposées à l'Acteur Humain pour interagir avec l'Agent Cognitif ?

5 CONTEXTE SOCIO-TECHNIQUE D'UN AGENT COGNITIF POUR AUTOMATISER LES TESTS D'INTRUSION MILITAIRES

Nous allons analyser, du point de vue d'une startup innovante conceptrice, le contexte socio-technique entourant l'Agent Cognitif, objet technique en question, destiné à automatiser les Tests d'Intrusion dans un environnement militaire pour maximiser les chances que cette innovation soit acceptée non seulement par ses utilisateurs cible dans son environnement mais également par toutes les parties prenantes intervenant dans ses phases de conception et d'exploitation.

Cette étude se matérialise sous la forme d'une analyse des risques que l'innovation ne soit pas acceptée et qui met en perspective les risques (sociaux, techniques, légaux et éthiques), les parties prenantes et la remédiation proposée pour chaque risque identifié et est présentée dans la Table 1 suivante:

Table 1: Analyse des risques menaçant l'innovation introduite par l'Agent Cognitif destiné à automatiser les Tests d'Intrusion militaires

Risque	Catégorie	Partie prenante	Remédiation
Pas d'accès aux émetteurs d'exigences	Social	QGs Militaires	Développer des contacts, habilitations
Développer une solution déjà existante	Social	Entreprises Défense	Développer des contacts, habilitations
Rejet IA pour applications militaires	Ethique	Public, Politiques	Sensibilisation
Infraction au cadre législatif IA	Légal	Législateurs	S'informer sur les récentes législations IA
Ne répond pas aux besoins Utilisateurs	Technique	Concepteur	Conception Centrée Utilisateur dans SDLC
Mauvaise conception Agent Cognitif	Technique	Concepteur	Analyse approfondie des modèles cognitifs
Failles Cybersécurité Agent Cognitif	Technique	Concepteur	Conception Centrée Cyber dans SDLC
Mauvaise utilisation Agent Cognitif	Technique	Concepteur	Conception Centrée Utilisateur dans SDLC
Comportement déviant Agent Cognitif	Technique	Concepteur	Spécification/Vérification formelles
Manque Confiance Agent Cognitif	Social	Utilisateurs finaux	Conception Centrée Utilisateur dans SDLC
Mauvaise utilisation Agent Cognitif	Social	Utilisateurs finaux	Formation

6 PROCESSUS DE CONCEPTION CENTREE UTILISATEUR

L'Analyse du contexte Socio-Technique nous apprend que pour remédier aux risques :

1. De concevoir et de développer un Agent Cognitif qui ne répond pas aux besoins de ses utilisateurs finaux dans leur environnement opérationnel.
2. D'une mauvaise utilisation de l'Agent Cognitif par ses utilisateurs finaux par défaut d'utilisabilité.
3. D'un manque de confiance des utilisateurs finaux en l'Agent Cognitif.

il faut appliquer un processus de Conception Centrée Utilisateur.

[15] nous propose le modèle suivant qui nous semble particulièrement bien adapté au développement de logiciels innovants tels que notre Agent Cognitif:

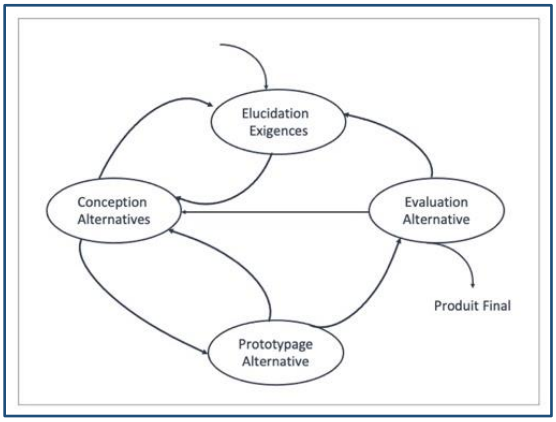


Figure 1: Processus de Conception Centrée Utilisateur pour logiciels innovants selon [15]

[15] et la norme ISO 9241-210 :2019 [16] nous énoncent les principes d'un processus de Conception Centrée Utilisateur pour les systèmes interactifs à travers leur Cycle de Vie du Développement Logiciel (SDLC) qui sont présentés dans la Table 2 suivante :

Table 2: Principes d'un processus de Conception Centrée Utilisateur et étape du SDLC correspondante :

Principes d'un processus de Conception Centrée Utilisateur	Etape du SDLC
Prise en compte des utilisateurs, de leurs objectifs et tâches dès le début du SDLC	Elucidation Exigences
Prise en compte des caractéristiques des utilisateurs	Elucidation Exigences
Répartition idoine des fonctions entre utilisateurs et technologie	Conception
Prise en compte du comportement des utilisateurs pour réaliser leurs tâches et contexte d'utilisation	Evaluation
Mesure et observation des réactions et de la performance des utilisateurs cible	Evaluation
Processus par itérations	Toutes
Participation active des utilisateurs à travers chaque étape du SDLC	Toutes

7 CYCLE DE VIE DU DEVELOPPEMENT LOGICIEL (SDLC) : AGILEUX

Les principes d'un processus de Conception Centrée Utilisateur énoncés précédemment nous permettent de nous lancer dans la sélection d'un Cycle de Vie du Développement Logiciel (SDLC) idoine pour élucider les exigences, concevoir le logiciel, réaliser l'analyse fonctionnelle et technique, coder, intégrer et évaluer notre Agent Cognitif.

AgileUX est une tentative de convergence entre les processus de Conception Centrée Utilisateur et les méthodes Agile telles qu'initiées par le *Manifeste Agile* [17]. De prime abord, un SDLC agile est attirant car itératif et nous avons vu qu'une des bases de la Conception Centrée Utilisateur est de suivre un processus itératif.

Selon [15], AgileUX présente encore des défis comme le fait que les équipes agiles doivent assimiler que la Conception Centrée Utilisateur n'est pas l'apanage d'un seul membre de l'équipe, typiquement le concepteur UX le cas échéant, mais une discipline et un état d'esprit que doit adopter tous les membres de l'équipe de développement. Un autre grand défi posé par l'AgileUX est le rôle clé central du 'Propriétaire du Produit' (Product Owner) au sein de l'équipe agile (*Scrum*). Homme/femme orchestre, il a pour rôle de maximiser la valeur du logiciel produit, d'établir les priorités et il est censé connaître les besoins de toutes les parties prenantes, ce compris ceux des utilisateurs finaux. Le 'Propriétaire du Produit' est rarement un utilisateur final et au vu de la rapidité de l'enchainement des itérations, il a rarement l'occasion de consentir du temps à élucider les besoins d'utilisabilité auprès des utilisateurs finaux, qui sont rarement conviés aux démonstrations du logiciel en fin d'itération pour donner leur avis. Il y a encore du chemin à faire…

Il y a une pléthore de méthodes agiles qui se combinent typiquement dans le cadre d'un projet de génie logiciel et nous en dénotons 3 principales : *l'Extreme Programming* (XP), *Kanban* et *Scrum*. Nous avons choisi *Scrum* pour notre projet car *Scrum* se prête mieux au prototypage itératif de nouveaux logiciels innovants mis en œuvre par une équipe de projet incluant plusieurs rôles différents, en plus des rôles typiques de 'Représentant du Client' et de programmeur, comme le rôle de Concepteur UX ou d'Expert en Cybersécurité. *Kanban* est plus orienté sur l'efficacité des équipes de développement pour la maintenance de logiciels existants de grande taille, avec des itérations plus longues le cas échéant ou selon un processus continu sans itérations. L'*Extreme Programming* est plus adapté à des équipes de développement constituées uniquement de 'Représentant du Client' et de programmeurs.

Le SDLC AgileUX que nous proposons dans la figure 2 ci-dessous est basé sur *Scrum* mais a comme particularité de traiter en amont d'une itération (càd lors de l'itération 0 et ensuite lors d'une, voire 2, itérations antérieures à l'itération en cours) l'élucidation des exigences et la conception des alternatives afin de pouvoir traiter en toute sérénité dans ces 2 phases, comme nous l'avons vu plus haut dans la table 2:

1. La prise en compte des utilisateurs, de leurs objectifs et tâches dès le début du SDLC.
2. La Prise en compte des caractéristiques des utilisateurs.
3. La Répartition idoine des fonctions entre utilisateurs et technologie.

En effet, comme également indiqué dans [15], réaliser ces 3 activités dans l'itération 1 et ensuite lors de l'itération en cours est un véritable défi car elles peuvent être difficilement réalisées en quelques jours.

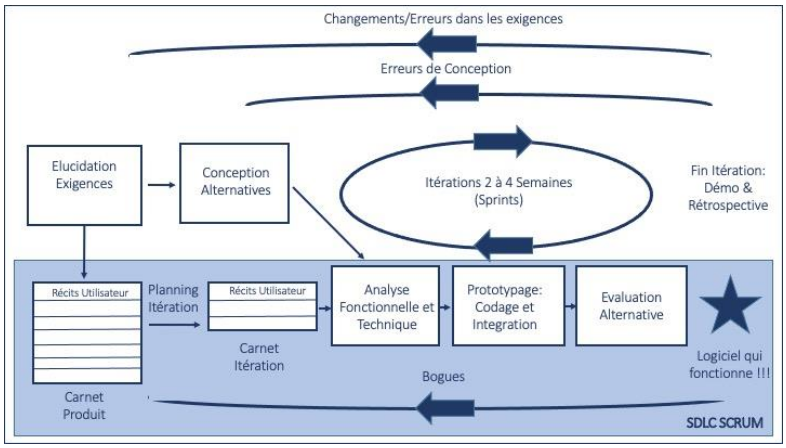


Figure 2: Cycle de Vie de Développement de Logiciel AgileUX basé sur SCRUM

8 INTERACTIONS HUMAIN-MACHINE ENTRE L'ACTEUR HUMAIN ET L'AGENT COGNITIF

8.1 Objectif de l'Acteur Humain

Lors de la phase d'exploitation des vulnérabilités selon le Penetration Testing Execution Standard [1], l'objectif de l'Acteur Humain est de déléguer à l'Agent Cognitif l'exécution de tests d'intrusion pour mettre à l'épreuve des réseaux d'ordinateurs militaires dont l'Acteur Humain est administrateur système.

8.2 Cas d'Utilisation

Le cas d'utilisation, spécifié en langage naturel, permettent de décrire les interactions entre l'Acteur Humain et l'Agent Cognitif et de spécifier les exigences fonctionnelles à satisfaire par ce dernier dans le cadre de ces interactions. De [5], on dérive le Cas d'Utilisation suivant :

Table 3: Cas d'Utilisation:

Cas d'Utilisation : Exploiter les vulnérabilités		
Acteur	Cyberanalyste appartenant au Centre des Opérations de Sécurité militaire.	
Pré condition	L'analyse de vulnérabilité est clôturée avec succès.	
Événement Déclencheur	Le Cyberanalyste reçoit l'ordre de démarrer l'étape d'exploitation des vulnérabilités	
Scénario principal	**Acteur**	**Agent Cognitif**
	1 - Crée une nouvelle attaque.	
		2 - Affiche l'écran de configuration de l'attaque.
	3 - Configure le profil d'attaque, les actions (non)autorisées, les contraintes et lance l'attaque.	
		4 - Sélectionne la procédure d'attaque adéquate, l'exécute et affiche l'écran de suivi de l'attaque.
		5 - Envoie des commandes vers les ressources informatiques cibles tenant compte de leurs réponses.
		6 - Affiche à l'acteur ses actions passées, leur évaluation et ses actions futures.
Alternative	6.b L'acteur arrête définitivement l'attaque.	
Post conditions	Le Système affiche à l'acteur le résultat de l'attaque terminée (6) ou stoppée (6.b).	

8.3 Architecture Logicielle

L'architecture Cognitive Soar (state operator and result) [18] [19] est le modèle cognitif idoine pour concevoir un Agent Cognitif qui répond aux exigences fonctionnelles spécifiées dans le cas d'utilisation. La présente section est destinée à justifier cette assertion : Issu de la Psychologie Cognitive et transposé à l'Intelligence Artificielle, un modèle cognitif permet de modéliser des processus cognitifs humains comme la perception, l'apprentissage, la mémoire, l'attention, le raisonnement, la catégorisation, l'émotion, le langage, la prise de décision, la planification et l'action et leurs interactions pour permettre à un Agent Cognitif d'interagir avec son environnement complexe et muable pour atteindre des objectifs fixés. Dans notre contexte, les deux principales caractéristiques de l'intelligence sont l'apprentissage permettant d'acquérir et de conserver dans sa mémoire de la connaissance et l'adaptativité permettant de faire face à de nouvelles situations.

La Figure 4 met en perspective le module de contrôle de tests d'intrusion qui va envoyer des commandes, en format Soar Markup Language, vers le module de Perception de Soar et recevoir, également sous forme de Soar Markup Language les réponses envoyées par Soar après le traitement de ces commandes dans la mémoire de travail. Ces commandes sont celles correspondants à l'étape 3 du cas d'utilisation 'exploiter les vulnérabilités' : Configurer et lancer l'attaque ainsi qu'à l'exception 6.b : Stopper l'attaque. De plus, ce module de contrôle va également permettre de mettre en œuvre la surveillance de chaque étape du déroulement de l'attaque tel que spécifié dans l'étape 6 du cas d'utilisation décrit dans la table 3:

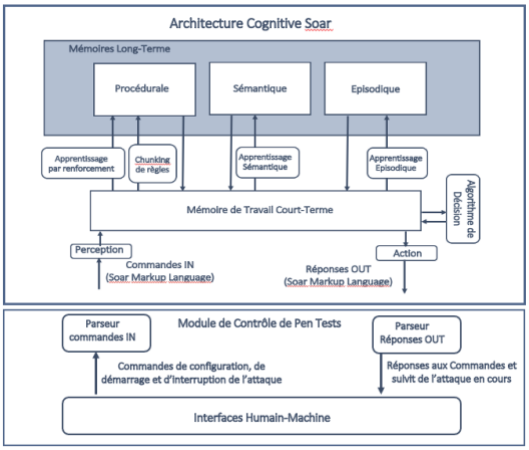


Figure 4: Diagramme de bloc de l'Architecture Cognitive Soar couplée avec son module de contrôle de tests d'intrusion

8.4 Interfaces Humain-Machine

Des interfaces Humain-Machine permettent donc à un Cyberanalyste humain d'exercer un contrôle de type 'Humain-dans-la-Boucle' suivant la définition spécifiée dans [21] : *La capacité d'une intervention humaine dans chaque cycle de décision du Système*. Dès lors, il est important que ces Interfaces Humain-Machine soient particulièrement conviviales.

La Figure 5 fournit un diagramme d'état-transition qui illustre la navigation à travers les 3 écrans supportant la gestion des attaques :

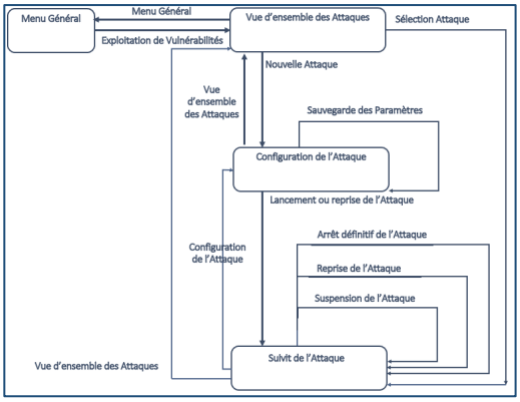
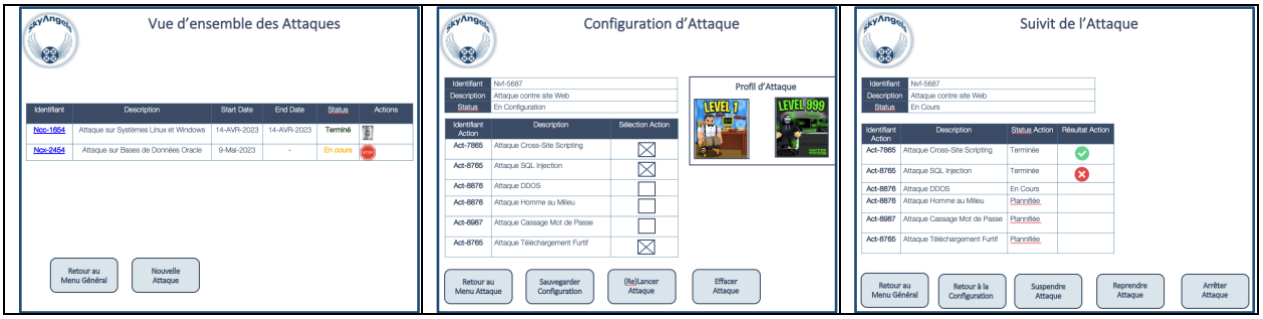


Figure 5: Diagramme d'Etat-Transition décrivant la navigation à travers les 3 écrans supportant la gestion des attaques.

Les Figures 4, 5 et 6 fournissent les maquettes des 3 écrans identifiés dans la Figure 5 :



Figures 4, 5, 6: Maquettes des écrans : 'Vue d'ensemble des Attaques', 'Configuration d'Attaque', 'Suivi de l'Attaque'.

9 CONCLUSION ET TRAVAUX FUTURS

Dans cet article, nous avons pris comme cas d'étude un des thèmes de l'appel à projets Fonds Européen de la Défense (EDF) 2023 qui couvrait l'automatisation, avec de l'Intelligence Artificielle, des Tests d'Intrusion effectués par les administrateurs système eux-mêmes pour mettre à l'épreuve leurs réseaux d'ordinateurs militaires. Nous avons réalisé un état de l'art sur les associations [IHM-Cybersécurité] d'une part et [IHM-Agent Cognitif] d'autre part pour prendre connaissance de leurs tendances actuelles et futures. Nous avons mis en exergue le concept d'HCISec qui s'est développé et qui a pour origine le besoin identifié par les experts en IHM d'améliorer l'utilisabilité des dispositifs de Cybersécurité, oh combien nécessaires par ces temps de Cyberattaques massives contre les entreprises privées et los organisations publiques et militaires Européennes. Nous avons effectué une Analyse Socio-Technique entourant un Agent Cognitif, objet technique d'intérêt, destiné à automatiser les Tests d'Intrusion dans un environnement militaire pour maximiser les chances que cette innovation soit acceptée non seulement par ses utilisateurs militaires finaux mais également toutes les parties prenantes. Nous avons énoncé un processus de Conception Centré Utilisateur approprié pour ce type de projet innovant et nous l'avons mis en perspective d'un Cycle de Vie de Développement Logiciel agile basé sur la tendance AgileUX et qui prend soin de ne pas bâcler, lors de la phase d'élucidation des exigences, les activités cruciales de prise en compte des utilisateurs, de leurs objectifs, caractéristiques et tâches ainsi que, dans la phase de conception, la répartition idoine des fonctions entre utilisateurs et technologie. Nous avons ensuite étudié, pour ce cas d'étude, les Interactions Humain-Machine entre l'acteur humain et l'agent cognitif. Pour ce faire, nous avons bien définit les objectifs de l'acteur humain dans ce contexte, élaboré un Cas d'Utilisation qui spécifie les interactions entre cet acteur humain et l'Agent Cognitif, conçu l'architecture logicielle en mettant en perspective l'Architecture Cognitive Soar avec un Module de Contrôle de tests d'intrusion qui expose des Interfaces Humain-Machine à l'acteur humain. Enfin, nous avons décrit ces interfaces utilisateurs graphiques sous forme d'un diagramme qui montre la navigation à travers ces écrans et conçu les 3 écrans principaux. Nous pensons avoir répondu aux questions de recherche énoncées.

D'un point de vue industriel, les prochaines étapes seront de poursuivre les efforts, dans le cadre de l'appel à projet EDF 2024 dont l'annonce des thèmes est imminente, pour trouver des partenaires industriels, Défense ou non, en France et en Europe (un consortium doit comprendre au minimum 3 partenaires établis dans 3 pays Européens différents) et des partenaires académiques notamment sur les thématiques IHM mentionnées dans cet article (l'IHM combinée à la Cybersécurité est une compétence en développement dans le SkyAngels Lab et la future startup CollosusGuardian Lab). L'Architecture Cognitive OpenCog Hyperon sera explorée comme alternative à Soar (US) pour des raisons de souveraineté Européenne de plus en plus nécessaire ! Enfin, d'un point de vue académique, pourquoi pas une future conférence IHM2x dédiée à la Cybersécurité ? …

REFERENCES

[1] Penetration Testing Execution Standard. Retrieved February 21, 2024 from http://www.pentest-standard.org .

[2] Farah Abu-Dabaseh and Esraa Alshammari. 2018. Automated Penetration Testing: An Overview. Princess Sumaya University for Technology, Amman, Jordan.

[3] Ge Chu. 2021. Automation of Penetration Testing. Ph.D Dissertation. University of Liverpool, Liverpool, UK.

[4] MASFAD 2 EDA CapTech Cyber. Retrieved February 21, 2024 from https://cloud.cylab.be/s/pxp4ToLnqpdn3F3 .

[5] EDF-2023-RA-SI-CYBER-ASPT: Automation of Security Penetration Tests. Retrieved February 21, 2024 from https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/edf-2023-ra-si-cyber-aspt

[6] ISO/IEC 25010:2023. Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Product quality model.

[7] Suzanne Robertson & James Robertson. Addison-Wesley. 2012. Mastering The Requirements Process, 253-256. Third Edition.

[8] ISO 9241-11:2018. Ergonomics of human-system interaction.

[9] Auguste Kerckhoffs. 1883. La Cryptographie Militaire, Journal des sciences militaires, vol. IX, 5-38 , 161-191.

[10] Jeffrey Grant & Chutima and Boonthum-Denecke. 2021. Analysis of Human-Computer Interaction Implication in Cyber Security. ADMI 2021: The Symposium of Computing at Minority Institutions.

[11] R. Kainda, I. Flechais and A.W. Roscoe. 2010. Security and Usability: Analysis and Evaluation, International Conference on Availability, Reliability and Security. Krakow, Poland.

[12] S. Russel & P. Norvig. 2021 .Artificial Intelligence – A Modern Approach, fourth Edition. Pearson,.

[13] Edited by A. Moallem. 2019. Human-Computer Interaction and Cybersecurity Handbook . Chapter 5. CRC Press.

[14] Formalisation de systèmes d'agent cognitif, de la confiance, et des émotions. Retrieved February 22, 2024 from https://www.irit.fr/~Jonathan.Ben-Naim/documents/bennaimLonginLorini1.pdf

[15] H. Sharp, J. Preece, Y. Rogers. 2019. Interaction Design, Beyond human-computer interaction, 5 Th Edition. Wiley.

[16] ISO 9241-210 :2019 – Ergonomics of human-system interaction – Part 210: Human-centered design for interactive systems.

[17] Manifesto for Agile Software Development Retrieved February 22, 2024 from https://agilemanifesto.org

[18] John E. Laird. 2019. *The Soar Cognitive Architecture,* The MIT Press.

[19] John E. Laird, Clare Bates Congdon, Mazin Assanie, Nate Derbinsky and Joseph Xu. 2023. The Soar User's Manual Version 9.6.1. University of Michigan, MI.

[20] D. Paul Benjamin, Partha Pal, Franklin Webber, Paul Rubel and Mike Atigetchi. 2008. Using a Cognitive Architecture to Automate Cyberdefense Reasoning. Edinburgh, UK.

[21] Independent high-level expert group on artificial intelligence set up by the european commission.2019. *The Assessment List For Trustworthy Artificial Intelligence (ALTAI).* European Commission, Brussels, Belgium.