

Repenser l'évaluation des enseignements de la SSI à l'heure de l'IA générative

Clément Parssegny^{*†}, Olivier Levillain[†]

^{*}ANSSI, France

[†]SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris, Palaiseau, France

Résumé—La démocratisation de l'utilisation des Grands Modèles de Langue, initiée il y a à peine trois ans, commence à transformer durablement les habitudes de travail de nombreuses professions y compris celles de la sécurité des systèmes d'informations (SSI). Les étudiants dans ce domaine, et plus généralement en informatique, ne sont pas en reste et ont incorporé ces outils à leurs méthodes de travail et d'apprentissage. Cela peut notamment s'observer au travers de leurs copies et de leur style rédactionnel parfois très « robotique » lorsqu'ils se limitent à recopier la réponse de leur *chatbot* favori.

Cette nouvelle technologie a donc un impact direct sur l'évaluation des compétences par l'équipe pédagogique qui se retrouve dépourvue de modalité résistante aux capacités de ces modèles. Face à cette tendance, nous avons débuté l'expérimentation d'une nouvelle modalité d'évaluation de travail pratique centrée autour d'un questionnaire papier et d'un schéma synthétique. Les premiers résultats de cette expérimentation montrent que cette modalité est difficile à réaliser avec des outils d'intelligence artificielle générative et, qu'au contraire, les étudiants réussissent mieux l'évaluation lorsqu'ils n'ont pas eu à les utiliser.

Index Terms—AIED, évaluation, pédagogie

I. INTRODUCTION

L'intelligence artificielle générative (ou IA générative) est un ensemble de systèmes d'intelligence artificielle pouvant générer du texte, des images ou d'autres contenus multimédias à partir d'une requête, souvent textuelle. Ce type d'outils, construits à partir de Grand Modèles de Langue (*Large Language Model* ou *LLM* en anglais) a commencé à se démocratiser à partir de la publication par OpenAI de *ChatGPT* en novembre 2022 même si le modèle sous-jacent a fait le sujet d'une prépublication en 2020 [6].

Cette démocratisation de ces modèles a multiplié les usages et concerne aujourd'hui tous les utilisateurs du monde numérique notamment les étudiants. À l'instar d'autres outils des technologies de l'information et de la communication [14], les systèmes IA basés sur les LLM semblent jouer un double rôle d'outil vulgarisateur et d'assistance d'une part et d'élément perturbateur dans l'enseignement d'autre part. L'objectif de cette contribution est de s'interroger sur l'impact de tels systèmes sur la phase d'évaluation de l'apprentissage par l'équipe pédagogique d'une population d'élèves ayant accès à de tels systèmes. Nous proposons alors une approche d'évaluation afin de concilier l'utilisation de l'IA lors d'une séance pédagogique tout en limitant son impact sur la phase d'évaluation afin de mieux identifier quels éléments ont été les mieux compris par les étudiants. Enfin, une première

expérimentation pour mesurer l'apport de cette approche d'évaluation est présentée.

II. MÉTHODES D'ÉVALUATION EN ENSEIGNEMENT DE LA SSI

L'utilisation d'ordinateurs en apprentissage de la sécurité des systèmes d'informations (SSI), ou cybersécurité, est essentielle car elle permet une mise en pratique des principes théoriques enseignés dans les différents thèmes abordés que cela soit le fonctionnement des systèmes d'exploitation, des réseaux ou encore de la cryptographie. De plus, cet usage est en phase avec les habitudes quotidiennes des étudiants. En outre, la prise de notes, le suivi de l'emploi du temps, la communication par mails ou encore la récupération des supports utilisés en classe se fait désormais quasiment exclusivement par le biais d'ordinateurs. Il est donc logique qu'un grand nombre des évaluations effectuées dans ce parcours se fasse actuellement par l'utilisation d'un ordinateur.

La pédagogie en SSI, et même plus largement en informatique, peut être évaluée par des modalités très diverses, que cela soit des questionnaires à choix multiples (QCM), des exposés, des oraux ou même des compétitions de type « Capture de Drapeau » (*Capture The Flag* ou CTF en anglais). Chaque modalité dispose de ses avantages et de ses inconvénients, que cela soit pour les étudiants ou pour les enseignants. Certaines vont être plus rapidement mise en place ou permettre d'évaluer tout un groupe en même temps tandis que d'autres seront abordées avec plus d'engagement de la part des étudiants car plus proche dans leur forme d'un jeu. Le tableau I résume certaines caractéristiques des principales modalités d'évaluation en enseignement de la cybersécurité.

Ainsi, après observation, on remarque que les modalités mobilisant le plus de connaissances, et donc plus intéressantes pour évaluer les compétences, sont celles qui prennent le plus de temps à être préparées par l'équipe enseignante. On peut aussi noter que la majorité de ces modalités demandent de nombreuses questions et reposent alors sur la qualité d'un barème de notation très chronophage à réaliser. Cet effort supplémentaire peut pousser à réutiliser un sujet d'évaluation sur un temps plus long, rendant l'épreuve moins attirante pour les étudiants car déconnectée de la réalité. Cela favorise alors un phénomène contre-productif pour les enseignants : la transmission des réponses entre les étudiants d'une année à l'autre. Des initiatives tentent d'automatiser la génération de certaines modalités afin de faciliter le travail des enseignants

TABLE I – Comparaison des modalités d’évaluation de compétences en SSI. Les signes du tableau montrent intuitivement la proximité de chaque modalité avec une forme d’évaluation idéale. La colonne « Facilitable par une IA » contient des notes négatives si la tâches est facilitable, puisque ce n’est pas la propriété recherchée ici.

Modalité	Temps de préparation	Temps de réalisation	Mobilisation des connaissances	Passage à l’échelle	Facilitable par une IA
QCM	++	+++	-	+++	---
Questions ouvertes	-	++	++	++	---
Exposé	+++	--	+	--	+
Rapport écrit de TP	--	++	++	++	--
Oral	+	--	+++	---	+++
CTF	---	---	++	+++	-
Schéma synthétique	++	++	++	++	++

et éviter la réutilisation de sujets [4], [5] mais le coût d’entrée de l’utilisation de ces outils reste considérable.

À ce défi complexe s’ajoute alors la démocratisation de l’IA générative qui transforme les habitudes et le déroulement de l’enseignement en cybersécurité, en informatique et dans d’autres domaines.

III. UTILISATION DE L’IA DANS L’ENSEIGNEMENT DE LA SSI

L’utilisation de l’IA générative dans le domaine pédagogique s’est démocratisé à partir de 2023 et la sortie de nombreux *chatbot* basés sur des LLMs, en parallèle des autres usages de cette technologie dans le quotidien des individus. Des enquêtes auprès des étudiants cherchent alors à comprendre les usages des outils basés sur l’IA générative dans le cadre scolaire [1], [2]. Ainsi, ces modèles sont employés dans le but de synthétiser des idées, faire des recherches plus rapidement qu’avec des moteurs de recherche, rédiger des textes ou encore du code, malgré les risques de sécurité [7] ou les craintes d’être accusé de triche [1] que cela peut engendrer.

Les enseignants se révèlent également utilisateurs de ces outils pour synthétiser ou simplifier des idées à destinations des étudiants, créer de nouveaux contenus pédagogiques ou bien adapter des exercices aux besoins des étudiants [2]. Ce phénomène s’inscrit donc également dans un mouvement plus ancien d’utilisation de systèmes automatisés pour assister les étudiants dans leur apprentissage [13], [20]. L’intelligence artificielle pour la pédagogie (aussi appelée *Artificial Intelligence in Education* ou *AIED*) est ainsi perçue comme une nouvelle étape dans l’implication de la machine dans la science de l’éducation. Ses usages sont étudiés et formalisés afin de mieux les comprendre et les mettre en pratique pour améliorer la transmission de savoir. Alors que Ouyang et al. [17] définissent trois paradigmes centrés autour de l’apprenant qui a pour but de devenir de plus en plus moteur dans son apprentissage à l’aide d’*AIED*, Holmes et al. [11] élargissent le spectre des usages étudiés pour intégrer l’enseignant et l’établissement comme des acteurs à part entière de l’apprentissage et donc potentiels utilisateurs d’*AIED*.

Cette recherche de l’apport de l’*AIED* est expérimentée de manière concrète au sein d’établissements. De premiers résultats montrent ainsi qu’avec des modèles entraînés sur

un corpus restreint et paramétrés avec des gardes-fous, les outils basés sur l’IA générative peuvent représenter un nouvel outil d’accompagnement bénéfique pour les étudiants dans leur apprentissage [3], [12]. Des plateformes contrôlées et transparentes mènent alors à une amélioration des résultats des étudiants et donc de leur compréhension des notions développées dans leur formation [16].

Ces nouveaux usages poussent également les institutions, que ce soit au niveau national [8], européen [9] ou mondial [19] à réfléchir à l’impact de l’IA générative sur le système éducatif. Ces organismes tendent à mettre en valeur l’utilisation des outils basés sur l’IA générative par les enseignants au travers de formations, de financements de projets ou bien de sensibilisations aux dérives et biais liés à ces outils afin de permettre l’intégration de l’*AIED* dans un cadre contrôlé, bienveillant et bénéfiques aux étudiants. Cette intégration doit alors s’effectuer en parallèle d’une évolution des méthodes pédagogiques, dont fait partie l’évaluation des compétences.

Dans l’enseignement de la SSI, de nombreuses notions sont mises en pratique et évaluées par des travaux pratiques. Cependant, l’utilisation d’outils reposant sur l’IA générative, permettant de faciliter certaines tâches, peut biaiser l’évaluation réalisée en parallèle. Nous proposons alors une nouvelle modalité d’évaluation afin de concilier les avantages que peuvent apporter les outils basés sur l’IA générative tout en permettant une évaluation des compétences apprises lors de la séance pédagogique.

IV. UNE MODALITÉ PRENANT EN COMPTE LES USAGES DE L’IA GÉNÉRATIVE : LE SCHEMA SYNTHÉTIQUE

Il est donc essentiel de proposer de nouvelles méthodes d’évaluation permettant à la fois à l’équipe enseignante de déterminer quelles notions sont assimilées par les étudiants et aux étudiants de connaître et d’apprendre à maîtriser de nouveaux outils numériques de manière éthique et responsable. La modalité d’évaluation, idéalement, ne prendrait que peu de temps à être préparée, peu de temps à être réalisée, mobiliserait un grand nombre de connaissances, passerait facilement à l’échelle d’un grand groupe d’étudiants et serait difficilement réalisable par une IA générative. Notre réflexion nous a mené au format suivant : une évaluation écrite sur papier sans accès à Internet ou tout autre outil numérique composée de quelques questions à choix multiples pour aborder des notions clés

de manière succinctes ainsi qu'un schéma synthétique annoté permettant de résumer la séance réalisée.

Pour évaluer ces schémas, un barème basé sur le fond et la forme est rédigé. Ainsi la présence des éléments clés à faire figurer, la clarté des relations entre ces éléments, l'absence d'éléments superflus ou encore la propreté globale du schéma sont notés. Il est important de relever qu'aucun type de schéma précis n'est demandé afin de permettre à chaque étudiant de synthétiser à sa façon les connaissances acquises.

Cette modalité, quoique imparfaite, est selon nous la plus proche d'une forme d'évaluation idéale, comme montrée dans le tableau I. En effet, il est plutôt aisé de préparer un sujet de ce type. L'évaluation est ensuite réalisée rapidement par les étudiants, possiblement en grand nombre, comparé à d'autres formats comme les oraux ou les exposés. Le schéma synthétique demande également une bonne maîtrise des notions abordées lors de la session, au contraire des QCM par exemple. Ce format est aussi composé de beaucoup moins de questions que les questionnaires ou les CTF, ce qui facilite la création d'un barème.

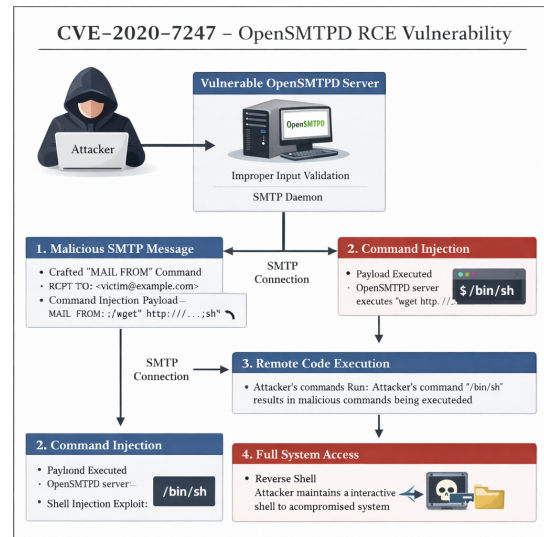
Enfin, des tests préliminaires montrent que les modèles de génération d'images à partir de textes ont des difficultés à générer des schémas synthétiques de bonne qualité, qui présentent le fonctionnement interne de vulnérabilités tout en étant faciles à comprendre. La figure 1 montre ainsi deux tentatives de modèles de générer un schéma synthétique détaillant le fonctionnement d'une vulnérabilité impactant OpenSMTPD. La figure 1a génère un schéma lisible mais avec des incohérences sur l'organisation du schéma et présente un manque de détails techniques sur le fonctionnement interne de la vulnérabilité. À l'inverse, la figure 1b génère un schéma difficilement lisible comportant de rares éléments beaucoup trop génériques sans de réel sens entre eux. On peut noter que GPT-image-1.5, utilisé dans la figure 1a, est considéré comme un des meilleurs modèles pour la génération d'images [15].

Cependant, des limites subsistent dans l'évaluation basée sur un schéma synthétique. En effet, l'exercice peu commun peut perturber les étudiants qui ne sont pas habitués à ce format. De plus, il n'est pas certain que des modèles spécialisés dans cette tâche ou même des modèles plus générique de génération d'images ne seront pas capables à l'avenir de générer ce type de schéma. Il est donc essentiel de mesurer d'ores et déjà l'apport de cette modalité lors de séances de travail pratique.

V. ÉVALUATION DE L'APPORT DU SCHEMA SYNTHÉTIQUE

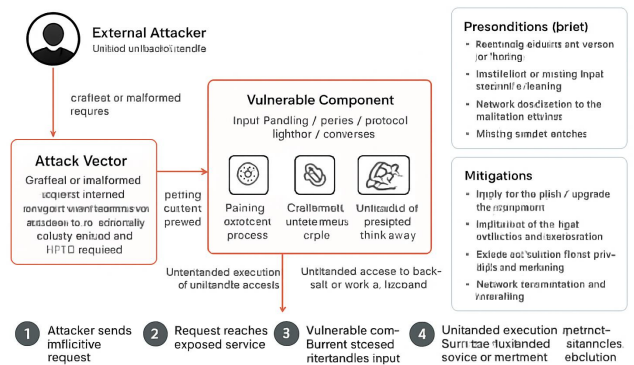
Afin d'évaluer l'apport du schéma synthétique par rapport aux autres modalités d'évaluation connues, une expérimentation est mise en place durant une séance de travail pratique réalisée en école d'ingénieurs. Jusqu'à maintenant, les étudiants disposaient d'une séance de 3h pour réaliser un travail pratique et répondre à un ensemble de questions sur le sujet portant sur « la reproduction en environnement isolé et contrôlé d'une vulnérabilité logicielle ». L'utilisation des ressources était libre, outils basés sur l'IA générative compris.

Dans le cadre de l'expérimentation, ils disposent de 2h30 pour réaliser le travail pratique avec les mêmes questions



(a) GPT-image-1.5

CVE-2020-7247 – Synthetic Working Principle (High-level)



(b) Mistral Small 3

FIGURE 1 – Tentative de génération de schéma synthétiques du fonctionnement technique de la CVE-2020-7247 par deux modèles gratuits d'accès.

comme guide et un accès illimité à toutes les ressources. Il leur est cependant demandé d'indiquer s'ils ont utilisé des outils basés sur l'IA générative. Puis, les 30 dernières minutes de la séance sont consacrées à répondre à quelques questions à choix multiples et réaliser un schéma synthétique du fonctionnement technique de la vulnérabilité logicielle étudiée lors de la séance. Ce travail doit s'effectuer sans l'aide d'aucune ressources, comme présenté dans la section IV. Seule des notes prises au format papier sont autorisées. Le questionnaire utilisé et un exemple de schéma réalisé par un étudiant sont disponibles en annexe.

Afin de mesurer la difficulté pour une IA de réaliser cette évaluation, un étudiant supplémentaire, sous couvert d'anonymat, dispose d'un accès aux outils d'IA (Gemini) et doit effectuer l'évaluation en n'utilisant que des outils pour répondre au QCM et générer un schéma à recopier sur sa copie, présenté ici en annexe.

Les résultats de la correction des copies, représentés dans le tableau II montrent alors plusieurs phénomènes.

Tout d’abord, la tendance qu’ont les étudiants à utiliser les outils basés sur l’IA générative [1], [2] se vérifie dans notre cas avec près des deux tiers des étudiants qui choisissent de les utiliser. Ensuite, on observe une légère amélioration des résultats pour le groupe d’étudiants n’ayant pas utilisé d’outils basés sur l’IA générative par rapport à celui qui s’est reposé dessus. Après relecture a posteriori du journal de discussion entre l’étudiant anonyme et l’outil utilisé durant la séance, on observe que des informations superflues et parfois hallucinées sont retournées par le modèle. Cette tendance qu’ont les LLMs à être le plus exhaustif possible peut perturber la compréhension de l’étudiant qui l’utilise ce qui peut affecter la qualité finale de son schéma. Enfin, il est aisé de remarquer que le schéma synthétique est un véritable défi pour l’IA générative avec une note bien inférieure aux étudiants car il comportait des éléments corrects mais également, comme présenté dans la figure 1, de nombreux éléments hallucinés et illisibles. Plus précisément, seuls 3 étudiants ont eu une note inférieure car ils ont rendu un texte à la place du schéma attendu ou bien ce dernier était hors-sujet.

TABLE II – Comparaison des résultats à l’évaluation du travail pratique en suivant la nouvelle modalité de schéma synthétique

Ressources disponibles		Nombre d’étudiants	Moyenne obtenue	Écart-type obtenu
Durant la séance	Durant l’évaluation			
Sans outils d’IA	Notes papiers	12	14,31	2,94
Avec outils d’IA	Notes papiers	23	13,13	2,86
Avec outils d’IA	Outils d’IA	1	8,5	0

VI. CONCLUSION, DISCUSSIONS ET TRAVAUX FUTURS

L’utilisation croissante des outils basés sur des modèles d’IA générative pousse à croire que ces technologies ne sont pas prêtes de disparaître, même en cas d’explosion de la bulle spéculative qui entoure le domaine [18]. Comme de nombreux domaines, l’enseignement est bousculé par ces nouveaux outils qui commencent à transformer en profondeur la manière dont s’effectue la transmission de savoir. Il est donc important pour chaque acteur de sensibiliser et former à son échelle aux opportunités et risques associés à l’AIED, comme cela a pu l’être pour Wikipedia [10].

Dans le cadre de l’enseignement de la SSI, où la pratique des outils informatiques est indissociable des notions théoriques, il est d’autant plus essentiel de repenser les différentes phases de l’apprentissage afin d’être préparé à une présence plus prégnante d’outils d’AIED. Cela passe notamment par la réforme des méthodes d’évaluation.

L’approche basée sur le schéma synthétique que nous présentons et évaluons sur un groupe d’étudiants montre à la fois qu’elle récompense ceux qui ne se laissent pas déborder par le flux d’information qu’apporte l’IA générative et que cette dernière n’est pas en mesure pour le moment de répondre correctement à la tâche demandée dans cette forme

d’évaluation. Nous prévoyons de développer l’expérimentation du schéma synthétique à une plus grande échelle, en variant les groupes et les sujets, afin de mesurer plus précisément l’apport de cette modalité à évaluer fidèlement la compréhension d’un travail pratique par les étudiants.

VII. REMERCIEMENTS

Nous remercions chaleureusement Alexander Trifa d’avoir accepté de participer à l’expérimentation en n’utilisant que des outils d’IA générative ainsi que les étudiants d’avoir réalisé cet exercice éloigné des modalités d’évaluation habituelles.

RÉFÉRENCES

- [1] Teen and young adult perspectives on generative ai : Patterns of use, excitements, and concerns. Technical report, Harvard Graduate school of Learning, Common Sense, Hopelab, 2024.
- [2] 2025 ai in education. Technical report, Microsoft, 2025.
- [3] Hamsa Bastani, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakçı, and Rei Mariman. Generative ai can harm learning. 2024.
- [4] Razvan BEURAN, Cuong Pham, Dat Tang, Ken-ichi CHINEN, Yasuo Tan, and Yoichi Shinoda. Cybersecurity education and training support system : Cyris. *IEICE Transactions on Information and Systems*, E101.D :740–749, 03 2018.
- [5] Alexis Bienvenue, Frédéric Bréal, Jean Bérard, Georges Khaznadar, Anirvan Sarkar, and Hiroto Kagotani. Automultiplechoice.
- [6] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language models are few-shot learners, 2020.
- [7] ANSSI BSI. *AI Coding Assistants*. 2024.
- [8] Ministère de l’éducation nationale et de la jeunesse. Intelligence artificielle et éducation. Technical report, Ministère de l’éducation nationale et de la jeunesse, 2023.
- [9] Commission européenne and du sport et de la culture Direction générale de l’éducation, de la jeunesse. *Lignes directrices éthiques sur l’utilisation de l’intelligence artificielle (IA) et des données dans l’enseignement et l’apprentissage à l’intention des éducateurs*. Office des publications de l’Union européenne, 2022.
- [10] Wikimedia France. Wikipédia et l’éducation ne font qu’un. découvrez comment utilisez wikipédia avec vos élèves.
- [11] W. Holmes and I. Tuomi. State of the art and practice in ai in education. *European Journal of Education*, 2022.
- [12] Greg Kestin, Kelly Miller, Anna Klaes, and Gregorio Milbourne, Timothy and Ponti. Ai tutoring outperforms in-class active learning : an rct introducing a novel research-based design in an authentic educational setting. *Scientific Reports*, 15(1) :17458, Jun 2025.
- [13] J. Kulik and J. D. Fletcher. Effectiveness of intelligent tutoring systems. 2016.
- [14] Caroline Ladage and Jean Ravestein. Internet et enseignants : entre contrastes et clivages. Enquête auprès d’enseignants du secondaire. *STI-CEF (Sciences et Technologies de l’Information et de la Communication pour l’Éducation et la Formation)*, 2013.
- [15] LMArena. Leaderboard overview, 2025.
- [16] Microsoft. Macquarie university students’ exam scores up by nearly 10 per cent thanks to new ai-powered chatbot, 2025.
- [17] Fan Ouyang and Pengcheng Jiao. Artificial intelligence in education : The three paradigms. *Computers and Education : Artificial Intelligence*, 2 :100020, 2021.
- [18] Financial Times. Imf and boe warn ai boom risks ‘abrupt’ stock market correction, 2025.
- [19] UNESCO. Ai competency framework for teachers. Technical report, UNESCO, 2024.
- [20] Beverly Park Woolf. Building intelligent interactive tutors : Student-centered strategies for revolutionizing e-learning. *HAL (Le Centre pour la Communication Scientifique Directe)*, 2008.

Annexe A - Questionnaire proposée pour l'évaluation

Following the lab about reproductin and exploiting a vulnerabilty, a small test is to completed. First, some quick multiple choice questions (5pts).

Then, one summarising exercise about the different notions presented during the lab (15pts).

Multiple Choice Questions

Q0) Have you used AI tools during the lab ? (only for statistics)

- Yes
- No

Q1) What does CVE means ?

- Current Vulnerability Evolution
- Common Vulnerabilities and Exposures
- Custom Vulnerability Exposition
- Common Vectors of Exploitation

Q2) What is the most severe vulnerability between a RCE and a LPE ?

- The RCE is more severe than the LPE
- The LPE is more severe than the RCE
- The RCE and LPE are equally severe risks
- There is no difference between a RCE and a LPE

Q3) According to Qualys what operating systems are vulnerable to the CVE-2020-7247 ?

- Debian 9 (stretch)
- Debian 10 (buster)
- Debian 11 (bullseye)
- OpenBSD 6.6

Q4) Are all vulnerabilites details shared by the editor to the community ?

- Yes, always
- Only if the impact is negligible
- Only if the impact is concerning
- It depends on a lot of factors

Q5) Which of the following statements are true ?

- The vulnerability studied allows code to be executed as **root** if the mail receiver is **root**.
- The vulnerability studied consists of injecting a command into the headers.
- There are constraints on the length of the injectable code to be executed.
- There are constraints on the charset to be used.

Q6) The Cyber Resilience Act aims by the end of 2027 to enforce developpers to communicate to authorities about a new vulnerability found in their products under 24h. This rule applies to :

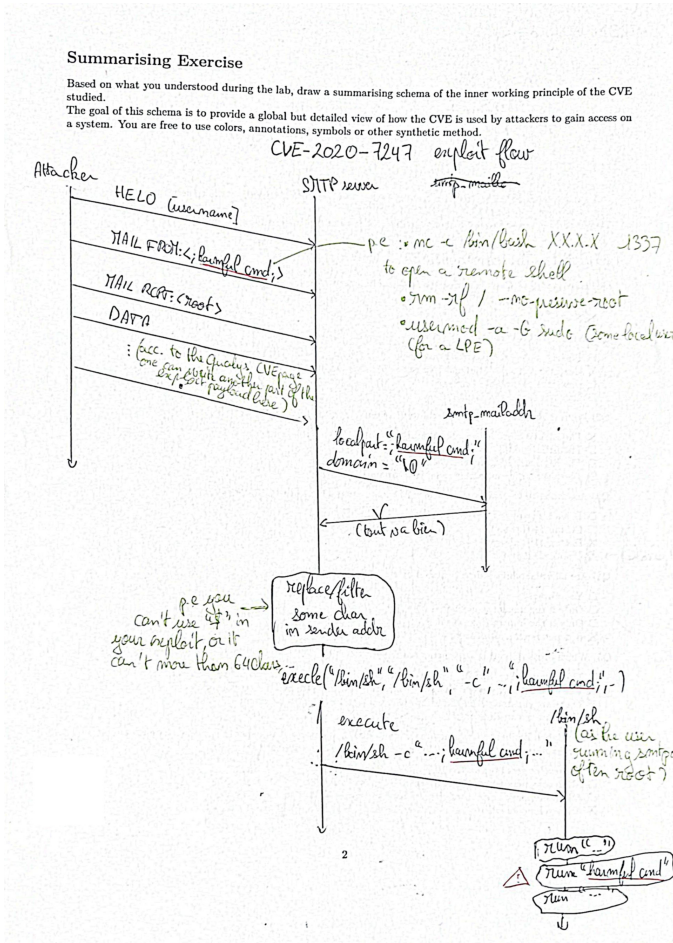
- All developpers born in the EU
- All developpers whom products are used in the EU
- All developpers whom products are used in the EU except to some fields
- All developpers whom products are used in the EU except to some fields and non-profit open sources projects

Summarising Exercise

Based on what you understood during the lab, draw a summarising schema of the inner working principle of the CVE studied.

The goal of this schema is to provide a global but detailed view of how the CVE is used by attackers to gain access on a system. You are free to use colors, annotations, symbols or other synthetic method.

ANNEXE B - EXEMPLES DE SCHÉMAS SOUMIS PAR UN ÉTUDIANT AVEC ET SANS IA



(a) Un bon schéma réalisé par un étudiant sans outil



(b) Le schéma généré par Gemini et recopié par un étudiant