Travail de session et capsule vidéo
INF8430 Hiver 2021

1) Travail en équipe

2) Composition des équipes

a) Taille des équipes : Quatre (4) personnes ;

b) Exceptionnellement, il y aura 1 ou 2 équipes de 3 personnes mais aucune de 5 personnes ;

c) Equipes : les équipes de travail de session et de capsules sont les mêmes que les équipes de laboratoires.

3) Litige entre équipiers

a) En cas de litige, la première étape est d'en discuter ouvertement entre équipiers concernés et de régler le problème ;

b) Si le problème persiste et qu'il s'agisse d'un problème de comportement d'un équipier, parlez-en le plus tôt possible avec le chargé de cours ;

4) Rencontres

a) Rencontres obligatoires : Le chargé de cours rencontrera les équipes deux fois lors de la session :

- i) La première rencontre aura lieu le 28 janvier entre 9h00 et 19h00 :
  - (1) Ce sera 3 jours après avoir soumis votre sujet de travail de session
  - (2) Objectifs : Mettre en place une démarche efficace de réalisation du travail de session et évaluer les besoins de l'équipe en matériel et logiciel
- ii) La deuxième rencontre aura lieu le 11 mars entre 9h00 et 19h00 :
  - (1) Ce sera 3 jours après la remise partielle du travail de session
  - (2) Objectifs : S'assurer que l'équipe comprend les commentaires du correcteur et vérifier le travail qui reste à faire entre ce moment et la remise finale du travail de session, le 29 mars et à vérifier le bon fonctionnement de l'équipe
- iii) Ces rencontres ont lieu à distance (hyperlien Zoom – Rencontres sur le Moodle du cours) et durent entre 25 et 45 minutes
- iv) Ces rencontres sont obligatoires pour tous les membres de l'équipe sous peine des pénalités mentionnées au plan de cours
- v) C'est la responsabilité de l'équipe de prendre rendez-vous avec le chargé de cours dans la semaine qui précède la rencontre. Les plages de rencontres sont 09h00, 09h45, 10h30, 11h15, 12h00, 12h45, 13h30, 14h15, 15h00, 15h45, 16h30, 17h15, 18h00, 18h45. Premier arrivé, premier servi, le cachet horodateur du courriel en faisant foi

b) Rencontre(s) facultative(s) : Le chargé de cours est disponible pour autant d'autres rencontres que nécessaires aux équipes pour rendre un travail exceptionnel.

5) **Sujet du travail**

a) L'orientation du travail doit être résolument « investigation numérique ».

b) Sujet : Trois options sont possibles

i) Choisir un sujet tiré de l'imagination collective de l'équipe. Le sujet traité doit être traité théoriquement et expérimentalement. Le sujet doit être **résolument** relié à la forensique. Par ailleurs, je n'accepterai pas un sujet pour lequel je ne pourrai pas vous aider

ii) Choisir parmi les sujets de la liste de sujets proposée ci-dessous (au numéro 6)

iii) Reprendre et pousser plus loin un sujet traité en 2021 ou 2022 et assujetti à la licence Creative Commons 4.0 (voir items E4_2021 et E4_2022 du Moodle du cours)

c) Non-sujet : Tout ce qui n'est pas relié à la recherche de preuves ou de traces ou à une recherche dans le domaine de la forensique ou à une recherche fortement liée à la forensique. Autrement dit : Un investigateur numérique qui lirait votre travail doit convenir que votre travail est très utile ;

d) Comme il s'agit d'un travail de niveau maîtrise, votre travail doit démontrer une connaissance approfondie du sujet choisi. Vous devez donc éviter de couvrir trop large (c'est-à-dire d'emprunter plusieurs champs de recherche) et plutôt opter pour une couverture d'un seul sujet en profondeur. Par exemple : Opter pour étudier les traces laissées par l'utilisation de tous les logiciels de clavardage serait déraisonnable. Par contre, étudier les traces laissées par l'utilisation de Skype lorsqu'utilisé par un Windows 10 Pro français est super !

6) **Liste des sujets proposés** – Même si ce n'est pas mentionné dans les descriptions suivantes, chaque sujet traité doit être traité théoriquement et expérimentalement. Le sujet doit être **résolument** relié à la forensique. Par ailleurs, je n'accepterai pas un sujet pour lequel je ne pourrai pas vous aider.
ATTENTION : Si vous choisissez un de ces sujets, vérifiez si le sujet fait l'objet d'un document sous licence Creative Commons 4.0 (Items E4 du Moodle)

a) Réaliser une machine virtuelle (Windows ou Linux) permettant de recueillir un site web complet d'une manière forensique et permettant à un huissier d'effectuer un constat sous serment que le site web en question contient des informations constituant des preuves

b) Récupération et manipulation d'un flux IP contenant la voix d'un individu et intercepté (le flux) lors :

i) D'une requête soumise à un assistant numérique personnel, **ou**

ii) D'un entretien en vidéo-conférence (excepté un enregistrement pur et simple de l'entretien)

Les équipes choisissant ce sujet doivent démontrer qu'ils peuvent intercepter le signal, le convertir en un fichier audio de quelque format pratique et l'utiliser pour nuire à celui dont la voix est interceptée

c) La loi de Benford sur les nombres anormaux peut-elle avoir des utilisations forensiques autres qu'en comptabilité forensique ? Voir https://fr.wikipedia.org/wiki/Loi_de_Benford et https://www.reddit.com/r/dataisbeautiful/comments/acow6y/asking_over_8500_students_to_pick_a_random_number/

d) Intercepter et déchiffrer le flux IP d'un ordinateur Apple dans le but de l'interpréter avec Wireshark ou similaire (chose possible avec mitmproxy en Windows). Vous devrez démontrer que vous êtes en mesure d'intercepter les paquets IP d'une communication, de les déchiffrer, de les interpréter en reconstruisant le contenu de la charge utile des paquets d'une communication complète ;

e) Récupération de fichiers partiellement écrasés, soit du domaine audio ou vidéo. Contexte : le jeu des sauvegardes et suppressions de différents types de fichiers fait en sorte que des fichiers contenant des preuves sont souvent effacés en partie seulement. Peut-on, avec l'aide des 512 octets d'un seul secteur, déterminer de quel type de fichier il s'agit, quelles sont ses propriétés (compression, échantillonnage, codec à utiliser) ? Et lorsque c'est identifiable, peut-on « transplanter » un entête à ce qu'on a pour qu'on puisse prendre connaissance de ce contenu ?

f) La 5G apporte la capacité des appareils à relayer un signal provenant d'un autre appareil 5G afin d'étendre la portée du réseau ou de palier à des facteurs de dégradation de la communication. Démontrez empiriquement la possibilité ou l'impossibilité d'interception des signaux ainsi relayés.

g) Traces de vie privée laissées par l'utilisation d'un logiciel de clavardage (ex : Discord) via un clavier, par un système d'exploitation quelconque ;

h) Traces de vie privée laissées par l'utilisation d'un logiciel d'échange voix (ex : Zoom), par un système d'exploitation quelconque ;

i) Dans le cadre de l'identification de la langue utilisée dans un texte écrit en caractères latins, l'approche actuelle est une approche basée sur le dictionnaire. Démontrez qu'il est possible d'utiliser une approche plus efficiente autre que l'approche dictionnaire. Notamment, la quantité des lettres les unes par rapport aux autres (par exemple le w est peu utilisé en français mais beaucoup plus en anglais), le rythme des phrases. Ce sujet est au croisement de la forensique informatique et de la forensique linguistique.

j) Traces laissées sur un ordinateur (station de travail ou portable) par l'utilisation d'un média social connu. Je favoriserai les médias sociaux les plus connus, à l'exception de Facebook (dont on a amplement fait le tour). Donc, vous ne pouvez pas choisir une application obscure utilisée par 5,000 dans un coin reculé de la planète. Où l'information est-elle stockée, stockage en mémoire RAM et de débordement ;

k) Mémoires d'un cellulaire. Extraction et analyse. Données transitoires, semi-permanentes et permanentes. Le cellulaire peut être Android, iPhone etc ;

l) Base de registre et similaires : Quelles sont les clés qui peuvent être utilisées pendant la fouille ou l'examen d'un média ? Comment retracer et extraire les informations effacées ? Peut-on faire la correspondance entre cela et d'autres éléments du média ? Sur Windows, on l'appelle la base de registre. Où sont les informations similaires pour Linux et Mac OS X ?

7) Format du travail

a) Travail de 4,000 à 6,000 mots (environ 8 à 11 pages à 1½ interligne), décompte excluant page-titre, table des matières, index, bibliographie et annexes ;

b) Informations minimales devant être présentes dans le travail (certaines sections peuvent être combinées pour faire un texte plus efficace, plus lisible) :

   i) Page titre avec sa licence Creative Commons (le logo amenuisé de la licence doit se répéter à tous les pieds de page)

   ii) Sommaire (« abstract ») de 100 à 200 mots ;

   iii) Très très brève introduction (de 200 à 400 mots) ;

   iv) Brève revue de littérature spécifique au sujet choisi (environ 750 à 1 000 mots). Attention : un résumé de lecture n'est pas une revue de littérature ;

   v) Protocole expérimental suivi (400 à 500 mots) : le protocole décrit les étapes et articulations de l'expérience (comment l'expérience s'est déroulée dans la réalité, ce qu'il advient lors de moments charnières (par exemple : début ou fin d'une phase, d'une action particulière ou lorsqu'un obstacle est franchi) ;

   vi) Observations et résultats (2 000 à 4 000 mots)

   vii) Conclusion : très très très bref retour sur vos découvertes, 200 à 250 mots (n'oubliez pas que le lecteur vient juste de compléter la lecture de votre rapport). Beaucoup plus important : les points restés sans réponse et les pistes de recherche pour le futur pour 250 à 500 mots ;

   viii) Médiagraphie (**minimum de 5 items** livrés en conformité avec les « Exigences de rédaction ») ;

   ix) Les aperçus d'écran doivent être en annexe.

c) Afin d'amener graduellement les étudiants à écrire un travail de session de haute qualité, les groupes rédigeront le document en cinq étapes :

   i) Remise de la composition de l'équipe : par courriel au chargé de cours avant le 25 janvier 2023 à 19h00

   ii) Rencontre du 28 janvier

   iii) Via le site web du cours : Remise du travail de session partiel du 8 mars 2023 à 19h00. Cette partie comprend la page-titre, la table des matières, une ébauche de sommaire, l'introduction, les objectifs du travail, la revue de littérature et la procédure expérimentale planifiée

   iv) Rencontre du 11 mars

   v) Via le site web du cours au plus tard le 29 mars à 19h00 : Remise de la version finale du travail de session

   vi) Le chargé de cours corrigera les documents intermédiaires dans les 100 heures de la date-heure limite pour le dépôt du rapport d'étape provisoire en indiquant les corrections à apporter pour améliorer le document ;

   NOTE : Comme je me donne la peine de vous lire et d'inscrire des corrections à faire dans vos documents, je m'attends que vous teniez compte de mes commentaires. Lorsqu'un groupe omet de le faire, je me sens porté à trop peu de mansuétude lors de la correction finale. Car si je prends la peine de vous le dire, c'est certainement que je vois un défaut.

d) Les rencontres se dérouleront via le Zoom de Polytechnique via l'hyperlien « Rencontre » sauvegardé sur Moodle.

e) Format des rapports et des artéfacts :

   i) Pour les rapports, format pdf, **1½ interligne** ;

   ii) Le rapport final sera accompagné des artéfacts jugés pertinents par les groupes. Ces artéfacts seront placés dans des conteneurs ad1 ou e01 (donc : aucun dd ni aff) ;

   iii) Le format « Science Direct » ou IEEE ou ACM est fortement suggéré. L'objectif est d'obtenir un produit fini publiable. Pour vous aider, vous trouverez sur le site du cours un gabarit Latec d'un document formaté de manière appropriée et fourni par Monsieur Marc Dib.

8) Déroulement

a) Avant la date-heure indiquée ci-haut, vous devez choisir un sujet ;

b) Le choix du sujet doit être validé par un courriel envoyé au chargé de cours ;

c) Deux équipes peuvent choisir le même sujet mais seront incités à couvrir le sujet sous deux angles différents ;

d) Le chargé de cours répondra au courriel de soumission dans les 24h00 de l'envoi par le groupe et précisera dans quelle mesure le sujet est accepté ;

e) Capsule vidéo de 9m45 à 10m15 : Tous les membres de l'équipe doivent présenter.

f) Remise des artéfacts :

   i) Les artéfacts supportant votre travail ne seront pas à remettre sauf si, lors de la correction, le correcteur se pose des questions qui ne peuvent être résolues que par la consultation de ces artéfacts. Si cela s'avère, l'enseignant contactera tous les membres de l'équipe et discutera avec le représentant désigné de cette équipe pour déterminer une méthode de consultation des artéfacts. Cette méthode peut aller de la simple consultation à distance de l'écran du représentant sur lequel est affiché l'artéfact jusqu'à la transmission de l'artéfact par un moyen déterminé par l'enseignant après discussion avec le représentant. **Ceci dit** : Vous devez obligatoirement rassembler **d'avance** les artéfacts demandés et de la manière prescrite afin de pouvoir les fournir sans délai.

   ii) Ce qui est exigé : Image ou prélèvement forensique des preuves ou traces

   iii) Format : ad1 ou e01

   iv) C'est la responsabilité de l'équipe de prévoir ce qui est nécessaire pour remettre les artéfacts au chargé de cours les artéfacts, sur demande de ce dernier

9) Pénalités

a) Voir plan de cours

b) Il vaut mieux remettre un travail incomplet qu'un travail en retard !

10) Critères de correction du travail

a) Qualité du français : Grammaire, syntaxe, orthographe

b) Qualité du sommaire (« abstract »), de l'introduction et apparence générale du document : pertinence, présentation efficace, utilité pour le lecteur, aspect esthétique, configuration pratique

c) Revue de littérature : pertinence, quantité, notoriété des auteurs, variété des documents. Le fait que c'est bien une revue et non un résumé de lecture

d) Protocoles expérimentaux et plan de travail : Ordre logique, sécurité du protocole face aux éléments externes, niveau de détail (suffisant sans être superflu)

e) Indices extraits des artefacts : Quantité et qualité. Description, illustration, localisation, interprétation des indices, niveau de détail, illustration efficace

f) Conclusion : Qualité du résumé et perspectives

g) Respect des étapes du travail et participation aux rencontres

h) Appréciation subjective du correcteur

11) Autres informations

a) Vous devez faire le pont entre l'investigation numérique, votre formation d'ingénieur et l'emploi que vous espérez exercer à la fin de votre diplôme ;

b) Imaginez-vous employé d'une entreprise d'enquêtes informatiques et vous devez guidez les investigateurs numériques dans leur travail de terrain ;

c) Les travaux ne seront pas remis à vos collègues pour fins de révision mais pourront servir aux étudiants des années prochaines s'ils sont visés par une licence Creative Commons 4.0 portée sur chaque page de votre travail de session final ;

12) Publication du travail

Les travaux rattachés à une licence Creative Commons 4.0 seront mis à la disposition des étudiants qui feront ce cours lors des prochaines années. Les travaux seront conservés pour aussi longtemps qu'ils serviront de sources pour des travaux de session plus une période de battement de cinq années.

13) **Capsule vidéo**

a) Vous devez créer une capsule vidéo d'une longueur de 10 à 12 minutes présentant votre travail de session

b) Tous les membres de l'équipe doivent :

i) Participer aux travaux de mise en place et de création de la capsule

ii) Prendre la parole pour au moins deux minutes dans la capsule

c) La capsule doit être remise via le site du cours à l'heure indiquée au plan de cours

d) La capsule doit être en format interprétable par un système Windows 10, sans qu'il soit nécessaire d'installer quelque logiciel que ce soit (donc, style avi, mp4, wmf,…) ;

e) La section de la présentation des résultats doit durer au minimum 5 minutes

f) Utiliser des couleurs et des arrière-plans sobres

g) La capsule doit présenter une section portant sur les perspectives de recherches consécutives au projet de recherche présenté.

14) **Correction des capsules :** Les capsules seront visionnées par l'enseignant et le chargé de laboratoire et seront visionnées lors d'une séance dont la date est indiquée par le chargé de cours au moment approprié. Les étudiants intéressés à assister à leur diffusion seront invités à le faire. Le contenu des capsules ne sera pas sujet lors de l'examen final.

15) **Licence** : Vous pouvez octroyer au cours INF-8430 une licence Creative Commons visant la capsule vidéo. Elle sera la bienvenue et nous l'utiliserons à titre d'exemple pour les étudiants des années prochaines.

16) **Critères relatifs à la capsule vidéo**

a) Pertinence des éléments de la présentation : Les éléments choisis doivent être les plus importants ressortant du travail, et doivent établir clairement le lien entre génie et investigation numérique

b) Préparation et exécution de la capsule : Le déroulement de la présentation se déroule selon un schéma clairement planifié

c) Participation des membres de l'équipe : Chaque membre de l'équipe apporte quelque chose à la présentation. Equilibre (au point de vue temps et au point de vue difficulté technique de l'explication) entre les membres de l'équipe.

d) Clarté des images et pertinence des items choisis pour illustrer le travail : Les images ne sont pas floues, les transitions se font de façon douce et tranchée, les items choisis pour représenter le travail fait sont pertinents

e) Attitude : Ton des présentateurs, originalité de l'approche pédagogique de la présentation

f) Parler clairement et en bon français : Pour les personnes dont la langue maternelle et d'usage est le français, on s'attend à un très gros effort. Pour ceux de langue maternelle autre que le français : considération de l'effort à parler un bon français

g) Appréciation générale : Comment le correcteur a aimé votre présentation