**Τμήμα Πληροφορικής**
Εαρινό εξάμηνο 2023-24
**Dr. Παναγιώτης Δεδούσης,**
B.Sc., M.Sc., Ph.D.
e-mail: dedousisp@aueb.gr
website: www.infosec.aueb.gr

# ΚΥΒΕΡΝΟΑΣΦΑΛΕΙΑ – ΓΡΑΠΤΗ ΕΡΓΑΣΙΑ

## Η συμβολή της Τεχνητής Νοημοσύνης στην Κυβερνοασφάλεια

**Θέμα**

Το Νοέμβριο του 2022, η OpenAI παρουσίασε το ChatGPT, το οποίο θεωρήθηκε ως το πλέον προχωρημένο chatbot, ικανό να ανταποκρίνεται σε ερωτήματα που αφορούν σχεδόν κάθε αντικείμενο ενδιαφέροντος. Το ChatGPT βασίζεται σε τεχνικές Τεχνητής Νοημοσύνης (Generative AI) και μεγάλα γλωσσικά μοντέλα (LLMs), τα οποία έχουν εκπαιδευτεί σε μεγάλες ποσότητες δεδομένων κειμένου. Χρησιμοποιώντας αυτά τα μοντέλα, το ChatGPT είναι σε θέση να κατανοήσει το εισερχόμενο κείμενο και να παράγει σχετικές και πλήρεις απαντήσεις σε φυσική γλώσσα.

Η χρήση Generative AI έχει ανοίξει νέους δρόμους για την ενίσχυση της κυβερνοασφάλειας, αλλά έχει επίσης δημιουργήσει νέες προκλήσεις και απειλές. Υπό αυτό το πλαίσιο καλείστε να **διερευνήσετε**, με **κριτική σκέψη, ποια θεωρείται ότι είναι η συμβολή της Generative AI στην κυβερνοασφάλεια**.

Συγκεκριμένα, καλείστε να διερευνήσετε μεθόδους και τεχνικές της κυβερνοασφάλειας στον τομέα της άμυνας (π.χ., δημιουργία ασφαλούς κώδικα, εντοπισμό απειλών και ευπαθειών, αυτοματοποίηση διαδικασιών και συστημάτων έγκαιρης ενημέρωσης συμβάντων, έλεγχος κώδικα κλπ.) και να αναλύσετε πως η Generative AI βελτιώνει την αποτελεσματικότητα τους.

Θα πρέπει επίσης να διερευνήσετε και να αναλύσετε τις προκλήσεις που προκύπτουν από τη χρήση της Generative AI στην κυβερνοασφάλεια, όπως η δημιουργία πολύπλοκων κυβερνοεπιθέσεων (π.χ., Social engineering, ransomware και phishing attacks, κλπ.) με ελάχιστες τεχνικές γνώσεις, τα ηθικά ζητήματα (π.χ. κανόνες λειτουργίας των chatbots), καθώς και η πρόσβαση σε πρωτογενή δεδομένα (π.χ., παραβίασης πνευματικής ιδιοκτησίας).

Τέλος, προτείνετε λύσεις και μεθόδους για την αντιμετώπιση των προκλήσεων που εντοπίσατε και αναλύστε τις μελλοντικές προοπτικές της Generative AI στο τομέα της κυβερνοασφάλειας.

Εκτός από τη θεωρητική ανάλυση, πρέπει να εξερευνήσετε, να παρουσιάσετε και να εξηγήσετε τους **πρακτικούς τρόπους αξιοποίησης** της Generative AI μέσω ερωτημάτων/prompts στα σχετικά εργαλεία (π.χ. ChatGPT, Gemini κλπ.).

**Διευκρινήσεις**

Η εργασία θα αποτελείται από δύο μέρη:

**Μέρος Α: Θεωρητική Ανάλυση και Κριτική Επισκόπηση** που θα περιλαμβάνει τα πιθανά οφέλη, τους κινδύνους και τα ηθικά ζητήματα από την χρήση του Generative AI στην κυβερνοασφάλεια (σύμφωνα με την εκφώνηση της εργασίας). Μελέτη και χρήση τουλάχιστον οκτώ πηγών που αφορούν διάφορες πτυχές της Generative AI στην Κυβερνοασφάλεια, συμπεριλαμβανομένων ακαδημαϊκών εργασιών, αναφορών βιομηχανίας και περιπτωσιολογικών μελετών.

**Μέρος Β: Πρακτική Εφαρμογή:** Καλείστε να δώσετε ως είσοδο περιγραφές ή/και αποσπάσματα κώδικα σε ένα Generative AI chatbot (π.χ. ChatGPT, Gemini κλπ.), να επαληθεύσετε την ακρίβεια της ανάλυσης τους και να αναλύσετε τα αποτελέσματα τους στο πεδίο της Κυβερνοασφάλειας που επιλέξατε. Ενδεικτικά

πεδία αποτελούν ο εντοπισμός και η παραγωγή i) Phishing attacks, ii) Social engineering attacks, iii) Ιομορφικού λογισμικού, iv) παράκαμψη περιοριστικών κανόνων λειτουργίας των chatbots.

Τεκμηριώστε τη διαδικασία και τα αποτελέσματα της **πρακτικής εφαρμογής** με δομημένο τρόπο. Συμπεριλάβετε το σκοπό της μελέτης, τη μεθοδολογία που ακολουθήσατε, τα πορίσματά σας και τους προβληματισμούς σας συνδυάζοντας την εμπειρία που αποκομίσατε μέσω της κριτικής ανάλυσης που κάνατε στο πρώτο μέρος.

Η εργασία μπορεί να είναι στην **ελληνική ή αγγλική γλώσσα**, αλλά η ειδική ορολογία να είναι στα αγγλικά. Η έκταση της εργασίας να κυμαίνεται μεταξύ **3.500 και 4.500 λέξεων** (Γραμματοσειρά: Times New Roman, Μέγεθος: 11, Εκταση: ~12 – 15 σελ.), εξαιρουμένων εξωφύλλου, πίνακα περιεχομένων και βιβλιογραφίας/references.

Η εργασία πρέπει να εκπονηθεί από ομάδες αποτελούμενες από **2 με 3 φοιτητές**. Ατομικές εργασίες δεν θα γίνουν δεκτές. Η σύνθεση των ομάδων επαφίεται στην ευχέρειά σας. Ο βαθμός της εργασίας θα αποτελέσει το **30%** της συνολικής βαθμολογίας του μαθήματος και κατοχυρώνεται, αποκλειστικά και μόνον, για τις (κανονικές) εξεταστικές περιόδους Ιούνη και Σεπτέμβρη 2024.

Σε περίπτωση **αντιγραφής ή χρήσης πλατφόρμας δημιουργίας κειμένου μέσω Generative AI**, πέραν αυτής που χρησιμοποιηθεί για τις ανάγκες της εργασίας (και η οποία θα πρέπει να δηλώνετε ξεκάθαρα στο κείμενο) θα **μηδενίζεται**. Σημειώνεται ότι για να θεωρηθεί ένας φοιτητής επιτυχών πρέπει να αξιολογηθεί με ≥5.0 και στην εργασία και στο εργαστήριο και στην τελική γραπτή εξέταση.

Οι εργασίες θα πρέπει να διαβιβασθούν, ως συμπιεσμένο αρχείο **pdf**, ηλεκτρονικά στο διδάσκοντα (dedousisp@aueb.gr), το αργότερο μέχρι τη **Δευτέρα, 29 Απριλίου 2024,** (<17:00) (σημειώνεται ότι **δεν θα δοθεί καμία παράταση** και όσες εργασίες υποβληθούν μετά την ως άνω ημερομηνία και ώρα **δεν** θα αξιολογηθούν). Η ορθή παραλαβή κάθε εργασίας θα πιστοποιείται **με μήνυμα του διδάσκοντα**. Το όνομα του αρχείου θα πρέπει να περιέχει του αριθμούς μητρώου των συγγραφέων στην εξής μορφή PXXXXX_PXXXXX_PXXXXX.pdf) . Επίσης, **τα ονόματα και οι αριθμοί μητρώου των συγγραφέων πρέπει να αναγράφονται στην πρώτη σελίδα της εργασίας κάτω από τον τίτλο της καθώς και στο ηλεκτρονικό μήνυμα**.

Εφίσταται η προσοχή στην **ορθή και πλήρη περιγραφή** όλων των **βιβλιογραφικών πηγών** που αξιοποιήσατε. Για τη μορφοποίηση των βιβλιογραφικών πηγών να ακολουθηθεί το σύστημα Harvard.

**Καλή επιτυχία!**