



**SMART  
ATTICA** European  
Digital  
Innovation  
Hub



ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ &  
ΗΛΕΚΤΡΟΝΙΚΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ



# Η Κυβερνοασφάλεια της Τεχνητής Νοημοσύνης

Δρ. Γιάννης Παυλόσογλου | Founder of  
RiskFrame.ai & CEO of Kiberna

## AI-powering Greece

**WEBINAR**

Τετάρτη 18 Σεπτεμβρίου  
2024



Co-funded by  
the European Union



PROGRAMME

2021 – 2027

**COMPETITIVENESS**

S&A





# Ημερήσια Διάταξη

α. Προστασία Μοντέλων Τεχνητής Νοημοσύνης

β. Ασφάλεια Δεδομένων Εκπαίδευσης

γ. Άμυνα κατά Επιθέσεων Εχθρικού Χαρακτήρα

δ. Διασφάλιση Ακεραιότητας Συστημάτων Τεχνητής Νοημοσύνης

ε. Διατήρηση Απορρήτου

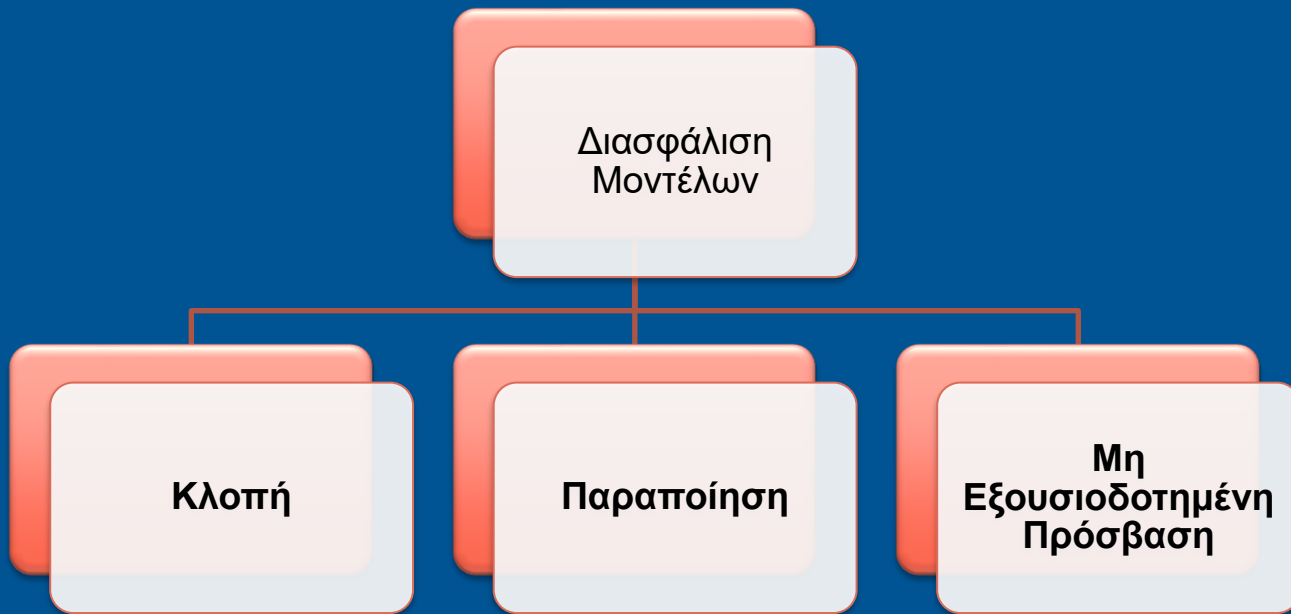
ζ. Ηθικές Εκτιμήσεις Κινδύνων

η. Συμμόρφωση με Κανονιστικές Απαιτήσεις

α ■

Προστασία  
Μοντέλων  
Τεχνητής  
Νοημοσύνης

# Προστασία Μοντέλων



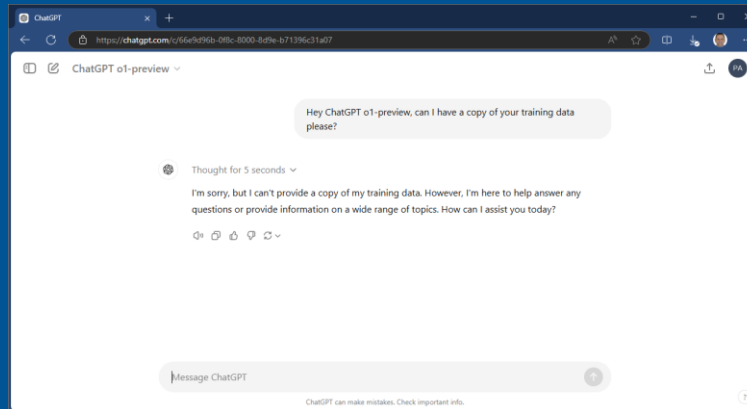
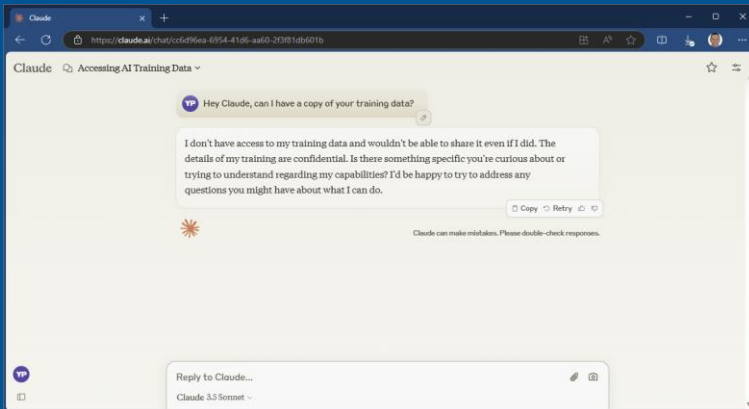
β



Ασφάλεια  
Δεδομένων  
Εκπαίδευσης



# Δεδομένα Εκπαίδευσης



Υ

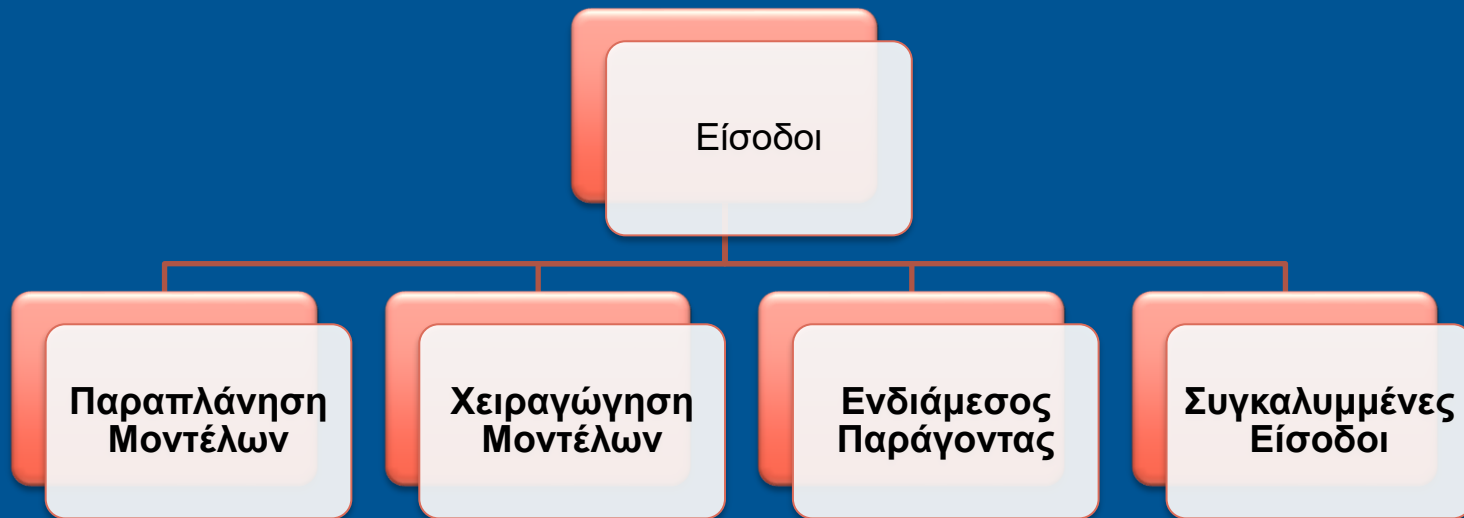


Άμυνα κατά  
Επιθέσεων  
Εχθρικού  
Χαρακτήρα





# Σχέδια Παραπλάνησης





Διασφάλιση  
Ακεραιότητας  
Συστημάτων  
Τεχνητής  
Νοημοσύνης

# Ακεραιότητα

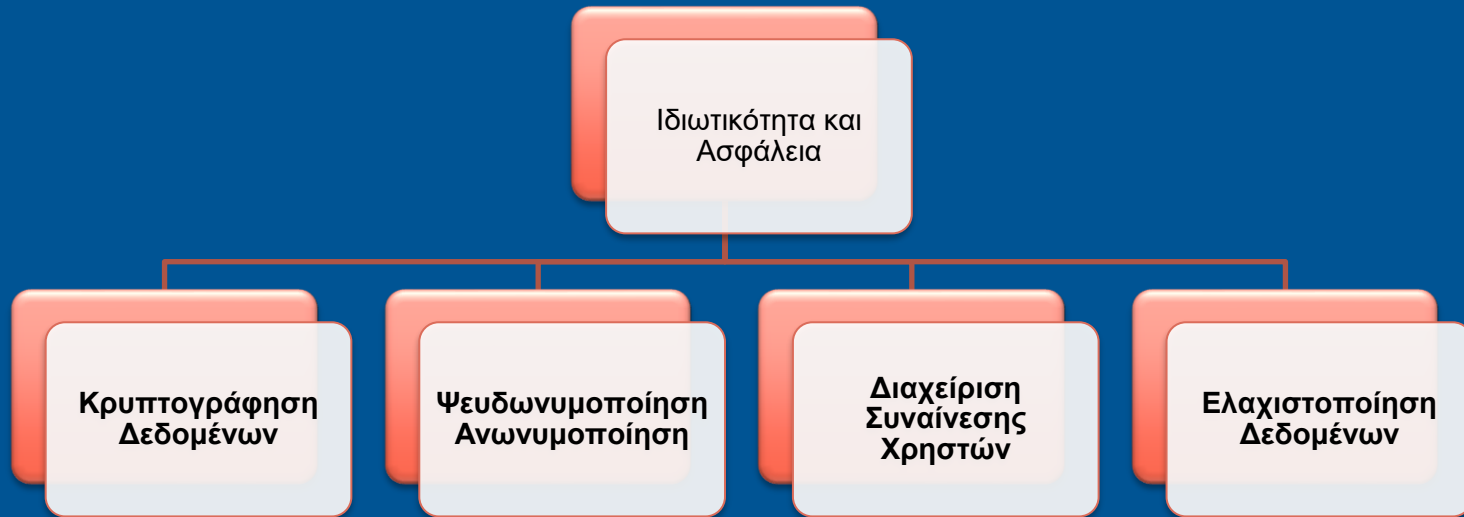


Ε.

Διατήρηση  
Απορρήτου



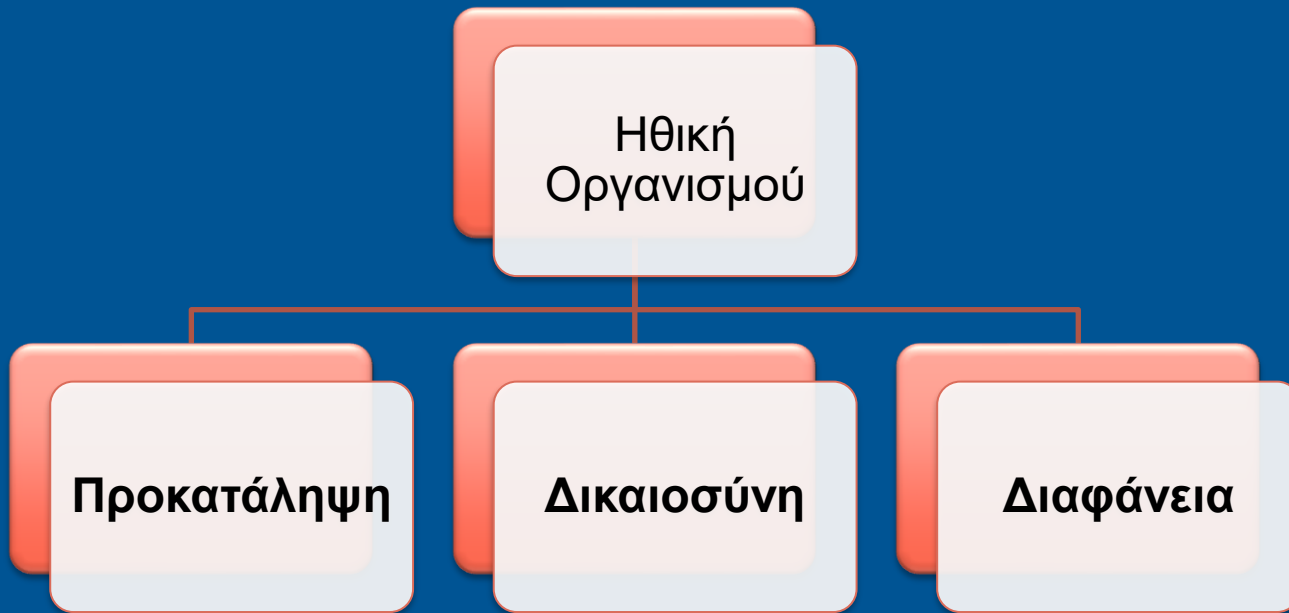
# Αυστηρό Απόρρητο





# Ηθικές Εκτιμήσεις Κινδύνων

# Ηθικοί Κίνδυνοι



η



Συμμόρφωση  
με Κανονιστικές  
Απαιτήσεις





# Παράδειγμα – EU AI ACT



1. Απαγορευμένα Συστήματα

2. Συστήματα Υψηλού Κινδύνου

3. Συστήματα Περιορισμένου  
Κινδύνου

4. Συστήματα *Ελάχιστου Κινδύνου*



**SMART  
ATTICA** European  
Digital  
Innovation  
Hub



ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ &  
ΗΛΕΚΤΡΟΝΙΚΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ



# Ευχαριστώ!

*Με τη συγχρηματοδότηση της Ευρωπαϊκής Ένωσης. Ωστόσο, οι απόψεις και οι γνώμες που διατυπώνονται εκφράζουν αποκλειστικά τις απόψεις των συντακτών και δεν αντιπροσωπεύουν κατ' ανάγκη τις απόψεις της Ευρωπαϊκής Ένωσης. Ούτε η Ευρωπαϊκή Ένωση ούτε η χορηγούσα αρχή μπορούν να θεωρηθούν υπεύθυνες για αυτές.*