

4D FRAMEWORK FOR RISKS & ETHICS

DE CE AVEM NEVOIE DE GVERNANȚĂ?

Viteza fără direcție este doar un accident care așteaptă să se întâmple.

În entuziasmul adopției AI, organizațiile riscă să sară direct la execuție și să ignore fundația etică și operațională. De aceea, apoi se confruntă cu halucinații, cu date sensibile expuse, sau cu soluții care încă nu știu ce problemă rezolvă.

CEI 4 PILONI

1D: DELEGARE

Înainte de a scrie un prompt, definim mandatul AI-ului pentru acel task.

Problema pe care o rezolvă AI-ul trebuie să fie extrem de clar definită, iar rezultatul dorit trebuie să fie măsurabil și ușor de evaluat.

2D: DESCRIERE

Un prompt vag sau generic generează rezultate mediocre.

Este important să oferim AI-ului contextul complet pentru succesul acelui task și să îi furnizăm un golden set de exemple pozitive și negative.

3D: DISCERNĂMÂNT

Prin discernământ validăm calitatea procesului delegat către AI.

Trebuie să definim cum măsurăm succesul, ce volum de erori e acceptabil, ce facem dacă modelul greșește, cine auditează, când.

4D: DILIGENȚĂ

E important ca inovația să nu apară mână în mână cu lipsa de conformitate. Aici validăm existența unor bias-uri ascunse în deciziile modelului, definim felul în care protejăm datele personale și construim un audit trail reproductibil.

RITUALUL DE GVERNANȚĂ

Matricea 4D se folosește ca un audit preliminar pentru toate inițiativele de adopție AI.

Nicio astfel de inițiativă nu trebuie implementată fără să clarificați următoarele:

- Ce date NU au voie să părăsească infrastructura proprie sub nicio formă? (ex: date financiare neauditate, date HR).
- Cine este *omul din buclă* (human-in-the-loop) care validează output-ul AI-ului?
- Respectăm normele legislative (ex: EU AI Act, GDPR) în această instanță?