



Atelier Unseen

Atelier Unseen

Control uman, utilizare etică și decizii asistate

Supraveghere efectivă, automation bias, responsabilitate și drepturile persoanelor

Obiectivele lecției

Diferențierea deciziilor asistate de deciziile exclusiv automatizate.

Recunoașterea automation bias și a controlului formal.

Stabilirea nivelului de supraveghere în funcție de risc.

Aplicarea principiilor de echitate, proporționalitate, transparență și contestabilitate.

Definirea rolurilor, autorității, override-ului și opririi sistemului.

Control uman real

Competență

Persoana înțelege capacitățile și limitele.

Autoritate

Poate respinge, modifica sau opri rezultatul.

Timp și dovezi

Are timp de verificare și acces la surse.

Trasabilitate

Decizia și raționamentul sunt documentate.

Niveluri de automatizare

Nivel	Rol AI	Rol uman
Asistență	Propune text sau opțiuni	Selectează și verifică
Recomandare	Analizează și sugerează	Evaluează și aprobă
Execuție supravegheată	Execută pași definiți	Monitorizează și oprește
Decizie automatizată	Produce decizia	Rol limitat sau ulterior

Supravegherea umană în AI Act

1

Înțelege

Capacități, limite și context.

2

Monitorizează

Anomalii, erori și performanță.

3

Interpretează

Outputul și instrumentele
explicative.

4

Override

Ignoră, inversează sau modifică.

5

Oprește

Intervine și întrerupe în siguranță.

AI Act și GDPR

AI Act

Pentru sistemele high-risk: supraveghere efectivă, persoane competente, monitorizare, loguri, informarea angajaților și măsuri organizaționale.

GDPR

Articolul 22 protejează persoanele față de decizii exclusiv automatizate cu efect juridic sau impact semnificativ și cere garanții, inclusiv intervenție umană.

Automation bias

Semnal

Scorul sau recomandarea AI este acceptată aproape automat.

Risc

Dovezile contrare sunt ignorate, iar controlul devine formal.

Reducere

Justificare proprie, limite vizibile, dublă verificare, override.

Principii etice

Principiu	Control practic
Echitate	Testezi efectele asupra persoanelor și grupurilor
Proporționalitate	AI este necesară și adecvată scopului
Transparență	Rolul AI este explicat când este necesar
Contestabilitate	Există explicație, corectare și intervenție
Responsabilitate	Rolurile și deciziile sunt documentate
Demnitate	Fără manipulare sau supraveghere disproporționată

Practici interzise relevante

Emoții la locul de muncă

Inferarea emoțiilor este interzisă, cu excepțiile medicale sau de siguranță prevăzute.

Manipulare

Nu folosim tehnici care afectează capacitatea unei persoane de a decide informat.

Social scoring

Nu clasificăm persoane pentru tratament nefavorabil nejustificat.

Decizii asistate pe departamente

Domeniu	Utilizare	Control
HR	Screening și evaluare	Criterii, specialist, contestare
Financiar	Scoring și fraudă	Date, explicație, dublă verificare
Clienți	Prioritizare și răspunsuri	Canal uman și revizie
Securitate	Alerte și incidente	Analiză de specialitate
Management	Previziuni	Ipoteze, scenarii și decizie asumată

Rolurile controlului uman

1

Operator

Folosește și observă anomalii.

2

Evaluator

Analizează outputul și dovezile.

3

Aprobator

Ia decizia finală.

4

Owner

Definește scopul și pragurile.

5

Control

IT, DPO, juridic și conformitate.

Când se oprește utilizarea

Rezultate neașteptate sau inconsistent repetate.

Date necorespunzătoare sau schimbarea scopului.

Risc pentru drepturi, sănătate, siguranță sau reputație.

Lipsa persoanelor competente pentru control.

Incident, plângere, discriminare sau imposibilitatea explicării rezultatului.

Principiul lecției

Omul nu este o semnătură formală peste rezultatul AI. Controlul uman înseamnă să înțelegi, să verifici, să decizi, să documentezi și să oprești.