

GHID PENTRU REALIZAREA UNEI ANALIZE DE IMPACT PRIVIND UTILIZAREA AI ÎN CERCETARE - versiunea februarie 2026

Acest ghid este destinat explicării modului în care cercetătorii Clusterului IMAGO-MOL pot realiza o analiză de impact privind utilizarea AI în cercetări medicale, folosind date obținute prin accesarea platformei pilot <https://3dnet.imago-mol.ro/portal>, utilizată în cadrul proiectului VELES Excellence Hub – Strengthening the South-East Europe Smart Health Regional Excellence and Boosting the Innovation Potential, GAP-101087483.

1. SCOPUL ȘI NECESITATEA ANALIZEI

Platforma prelucrează seturi de date care, pentru cercetători, au fost supuse în prealabil unui proces de anonimizare. Utilizarea algoritmilor de Inteligență Artificială (AI) introduce categorii noi de riscuri care exced simpla protecție a datelor cu caracter personal. Cercetătorul IMAGO-MOL are obligația de a efectua o analiză de impact pentru a asigura conformitatea cu standardele etice și legale (inclusiv Regulamentul UE privind Inteligența Artificială - AI Act¹).

2. EVALUAREA RISCULUI DE CONFIDENȚIALITATE

Analiza va trata riscurile specifice capacității de inferență a modelelor AI. Se va avea în vedere evaluarea riscului ca sistemul AI, prin capacitatea sa de a corela volume mari de informații, să deducă identitatea persoanelor vizate prin încrucișarea datelor din platformă cu surse externe (risc de re-identificare sau *de-anonymization attacks*).

3. EVALUAREA RISCURILOR DE BIAS ȘI ECHITATE

Se va analiza riscul de discriminare algoritmică, luând în considerare că anonimizarea datelor de intrare nu garantează neutralitatea rezultatelor. Se va avea în vedere evaluarea riscurilor:

- Identificarea variabilelor proxy: verificarea faptului că modelul AI utilizat nu utilizează date aparent neutre:
 - unitatea medicală de unde provin datele: Institutul Regional de Oncologie Iași sau Spitalul Clinic de Urgență „Prof. Dr. Nicolae Oblu” Iași;
 - ora și data scanării/investigației: o scanare realizată noaptea poate indica o urgență, iar în timpul zilei, o programare, ceea ce poate indica o gravitate a situației medicale a pacientului;
 - calitatea de asigurat (CAS): poate împărți pacienții direct pe clase sociale;
 - caracteristici biologice secundare: un CT abdominal poate arăta caracteristici precum distribuția grăsimii corporale, care poate fi un proxy pentru stilul de viață;

¹ <https://eur-lex.europa.eu/legal-content/ro/ALL/?uri=CELEX:32024R1689>

- markeri genetici specifici: anumite mutații genetice sunt specifice unor grupuri etnice. Includerea datelor genomice poate duce la identificarea etniei.
- Echitatea rezultatelor: un model AI antrenat pe date istorice anonimizate poate perpetua discriminarea sistemică fără a procesa date cu caracter personal, generând astfel decizii discriminatorii.

4. EVALUAREA RISCULUI PRIVIND HALUCINAȚIILE

Un model AI poate genera rezultate factuale greșite sau periculoase (halucinații) indiferent de datele de intrare. O măsură de reducere a acestui risc constă în evaluarea de către cercetătorul IMAGO-MOL a rezultatelor generate de către AI, prin filtrul propriilor cunoștințe și experiențe.

5. EVALUAREA RISCULUI PRIVIND CONFORMITATEA LEGALĂ ȘI ETICĂ

Regulamentul UE 2024/1689 (AI Act) se concentrează pe riscul sistemului asupra drepturilor fundamentale, sănătății și siguranței, nu doar pe protecția datelor. Dacă rezultatele cercetărilor conduc spre scoring sau triere, aceasta poate fi considerată ca având un risc ridicat privind deciziile medicale ulterioare, aplicabile unor viitori pacienți.

6. EXEMPLE DE MĂSURI DE REDUCERE A RISCURILOR

Reducerea riscurilor vizează implementarea unor mecanisme de control pentru a gestiona erorile specifice AI.

Riscurile pot varia în funcție de specificul fiecărui proiect de cercetare și de tipul de date prelucrate, dar, obligatoriu, nu trebuie furnizate modelului AI mai multe date decât sunt necesare scopului cercetării.

Trebuie menținut un nivel adecvat de control uman (*Human-in-the-Loop*) pentru validarea rezultatelor generate de AI, în special în situațiile în care deciziile pot avea efecte asupra viitorilor pacienți. Aceasta implică verificarea periodică a acurateții rezultatelor pentru a detecta degradarea performanței sau generarea de informații false.