OEIAC

Memòria d'activitats 2024

OBSERVATORI D'ÈTICA EN INTEL·LIGÈNCIA ARTIFICIAL DE CATALUNYA

Fundació Girona Regió del Coneixement

Universitat de Girona

Generalitat de Catalunya

Índex de continguts

1. Breu descripció de les activitats 4

2. Relació d'activitats 2024 5

a) Nou cicle de Seminaris OEIAC 5

b) Desplegament del Model PIO sobre Obligacions i Drets 14

c) Disseny i elaboració dels Model PIO Sectorials (Salut i Educació) 22

d) Sessions de bones pràctiques amb tots els Models PIO disponibles 25

e) Col·laboració amb iniciatives i avaluacions similars 26

f) Despeses generals relacionades amb totes les activitats 30

2. Resum d’actuacions i cronograma 31

3. Resum d’activitats de comunicació 2024 32

4. Membres del Consell Assessor 35

5. Persones col·laboradores 36

1. Breu descripció de les activitats

L'Obsevatori d'Ètica en Intel·ligència Artificial de Catalunya (OEIAC), que pren forma de Càtedra a la Universitat de Girona (UdG), té com a objectiu principal estudiar les conseqüències ètiques, socials i legals i els riscos i oportunitats de la implantació de la Intel·ligència Artificial en la vida diària a Catalunya, i des d'una òptica plenament transversal.

L'objectiu de l'OEIAC i totes les seves activitats s'insereixen en l'Estratègia d'Intel·ligència Artificial (IA) impulsada per la Generalitat de Catalunya que, amb el nom de CATALONIA.AI , realitza un programa d'actuacions per enfortir l'ecosistema català en IA.

CATALONIA.AI inclou un eix sobre "Ètica i Societat", que és on es situa l'OEIAC, per tal de promoure el desenvolupament d'una intel·ligència artificial ètica, que respecti la legalitat vigent, que sigui compatible amb les nostres normes socials i culturals, i que estigui centrada en les persones.

En el desplegament de la seva missió, l'OEIAC porta a terme una sèrie d'activitats relacionades amb 5 línies estratègiques dins CATALONIA.AI:

1) Investigar els impactes ètics, socials i legals de la intel·ligència artificial.
2) Establir directrius i bones pràctiques.
3) Realitzar transferència de coneixement.
4) Mantenir relacions amb organismes internacionals.
5) Col·laborar amb grups d'experts de tot el món.

La relació d'activitats que es detallen a continuació estan totes elles relacionades amb aquestes 5 línies estratègiques, per fomentar i garantir els usos ètics i responsables de la intel·ligència artificial i, a la vegada, per fer augmentar les seves oportunitats i abordar els reptes que suposa la seva implantació d'una forma cada cop més generalitzada.

En aquest sentit, l'OEIAC ha realitzat les següents principals activitats durant aquest any 2024: el IV Cicle de Seminaris OEIAC sobre "Ètica, impacte social i compliment legal"; el desplegament del nou Model PIO sobre Obligacions i Drets, tenint en compte l'aprovació recent de la regulació en materia de riscos de la IA a la Unió Europea (AI Act); el disseny i configuració de dos nous Models PIO sectorials (salut i educació); cursos de formació i capacitació en l'ús i implementació del Model PIO sobre Obligacions i Drets; disseny i difusió de nous recursos pedagògics sobre els objectius i actuacions de l'OEIAC com a pilar de l'Estratègia CATALONIA.AI; i el foment de col·laboracions i intercanvis, especialment internacionals, amb persones i estructures similars.

2. Relació d'activitats 2024

a) Nou cicle de Seminaris OEIAC

L'OEIAC va iniciar el seu primer cicle de seminaris sobre "La intel·ligència artificial en la nostra vida quotidiana" l'any 2021, mentre que l'any 2023 va portar a terme la tercera edició sobre "Implicació amb la sostenibilitat: De l'Antropocè a la Intel·ligència Artificial". L'any passat vam celebrar la quarta edició del Cicle OEIAC amb el tema "Ètica, impacte social i compliment legal".

Aquest nou Cicle OEIAC va tractar de com sovint les consideracions ètiques configuren les lleis i, al mateix temps, de com les normes legals poden influir en consideracions ètiques, generalment amb l'objectiu de promoure el benestar social i protegir drets de les persones. De vegades, però, poden sorgir situacions en què els estàndards legals i les consideracions ètiques entren en conflicte, i en aquests casos es requereix una anàlisi de les necessitats de les parts interessades com a principis ètics i avaluacions d'impacte social sovint més enllà dels requisits legals per atendre les necessitats i expectatives de la societat. Un clar exemple d'això és la recentment aprovada Llei d'Intel·ligència Artificial de la Unió Europea coneguda com a AI Act, que no només hauria de convertir-se en un marc per avaluar i proteger-nos contra els riscos de la IA, sinó també en un instrument per garantir drets individuals, benestar comunitari, valors socials i protecció del medi ambient.

No obstant això, en aquest moment es considera que la AI Act no té una visió prou adequada i centrada en les persones ja que no protegeix suficientment drets fonamentals com la privadesa, la igualtat o la no discriminació. Si bé la normativa inclou avaluacions d'impacte sobre els drets fonamentals, encara que hi ha dubtes sobre quin serà l'abast real que tindran aquestes avaluacions, entre altres perquè els desplegadors de sistemes d'IA només hauran d'especificar quines mesures s'adoptaran. un cop es materialitzin els riscos. D'altra banda, no hi ha obligatòriament d'implicar parts interessades com són la societat civil i les persones afectades per la IA en el procés d'avaluació. Això vol dir que les organitzacions de la societat civil no tindran una manera directa i legalment vinculant a contribuir a les avaluacions d'impacte. Tenint en compte que ha d'existir el dret a una explicació dels processos de presa de decisions individuals, especialment per als sistemes d'IA catalogats com a d'alt risc, això planteja preguntes sobre l'accés a la informació i l'obtenció. d'explicacions per part dels desplegadors de sistemes d'IA. Aquest aspecte es considera especialment rellevant donada l'absència de disposicions com ara el dret a la representació de les persones físiques o la capacitat de les organitzacions d'interès públic per presentar queixes davant les autoritats de control. Finalment, cal remarcar un altre punt important és que la AI Act només fa un primer pas per abordar els impactes ambientals de la IA i, això, malgrat la preocupació creixent en com l'ús exponencial dels sistemes d'IA pot tenir impactes greus en el medi ambient.

Per tal d'aportar algunes respostes i, és clar, també per obrir més preguntes sobre la complexa interacció entre l'ètica, la legislació i la tecnologia, l'OEIAC va dissenyar i portar a terme un nou cicle de seminaris de transferència de coneixement amb persones convidades d' arreu que treballen i promouen els usos ètics i responsables de la IA.

El Cicle OEIAC s'inicià el dimecres 29 de maig i es va allargar fins al 15 de novembre, quan va tenir lloc l'acte de Cloenda. Com és habitual des de la primera edición, aquest cicle està dirigit a tothom que té interès en la IA en general i en els usos ètics i responsables en

particular, així com en persones procedents dels sectors públic i privat.

Les dates específiques i persones ponents convidades han estat les següents:

29 de maig a les 11:00h – Conferència d'obertura presencial sobre "The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements", a càrrec de Alessandro Mantelero, Polytechnic University of Turin and EU Jean Monnet Chair in Mediterranean Digital Societies & Law.

➔ 111 vistes (març 2025): https://youtu.be/BazL0-mNnBA?feature=shared

10 de juny a les 17:00h – Seminari virtual sobre "Inteligencia artificial y derechos humanos en América Latina: Tendencias y reflexiones a partir de estudios de caso", a càrrec de Jamila Venturini, Derechos Digitales (associació de Drets Humans i Tecnologia a Amèrica Llatina).

➔ 70 vistes (març 2025): https://youtu.be/FOr4zvWaRCQ?feature=shared

16 de setembre a les 17:00h – Seminari virtual sobre "Biaixos algorítmics i privadesa de dades: Com podem utilitzar la informàtica teòrica per abordar un problema interdisciplinari?", a càrrec de Sílvia Casacuberta, University of Oxford (Global Rhodes Scholar) & Stanford University.

➔ 188 vistes (març 2025): https://www.youtube.com/live/FKGeHfuGy3s?feature=shared

21 d'octubre a les 17:00h – Seminari virtual sobre "La infrastructuralització de la intel·ligència artificial: Implicacions per als drets fonamentals", a càrrec de Roser Pujadas, Department of Science, Technology, Engineering and Public Policy, University College London. [https://www.youtube.com/live/RSIajpyP3Ew?feature=Shared]

➔ 167 vistes (març 2025): https://www.youtube.com/live/RSIajpyP3Ew?feature=shared

15 de novembre a les 11:00h – Conferència de cloenda presencial sobre "Concerns arising from large-scale datasets and AI systems and approaches to structural change", a càrrec de Abeba Birhane, Mozilla Foundation & School of Computer Science and Statistics at Trinity College Dublin.

➔ 135 vistes (març 2025): https://youtu.be/wA18lmwvKTE?feature=shared

Tots els seminaris foren gratuïts per a tothom i retransmesos en línia (streaming) a través del canal YouTube de la Universitat de Girona. Juntament amb les habituals conferències i seminaris, aquestes van anar acompanyades de dues entrevistes breus a les persones que van protagonitzar les conferències d'obertura i de cloenda. Aquestes entrevistes van anar a càrrec del director de l'OEIAC, Albert Sabater, sobre els principals interessos i accions que porten a terme les persones ponents amb les activitats i visions de futur de l'OEIAC. Aquestes entrevistes estan disponibles a través del mateix canal YouTube de la Universitat de Girona:

Entrevista al professor Alessandro Mantelero:
➔ 131 vistes (març 2025): https://youtu.be/42ZbDkaWVDA?feature=shared

Entrevista a la professora Abeba Birhane:

➔ 62 vistes (març 2025): https://youtu.be/PYpNzEAjn_A?feature=shared

Com és habitual, es va elaborar una pàgina web amb tota la informació i recursos a mida que es va anar actualitzant a mida que el Cicle OEIAC va anar avançant a partir del mes de maig.

La pàgina web del Cicle OEIAC 2024 fou la següent:
https://www.udg.edu/ca/catedres/oeiac/cicle-de-seminaris-oeiac

Juntament amb la pàgina web, també es va habilitar una web per a les inscripcions ja que d'aquesta manera fou possible rebre avisos així com més detalls i l'enllaç de connexió a totes les ponències (en streaming a través del canal YouTube de la Universitat de Girona). Aquesta web per a les inscripcions també fou utilitzada per qüestions logístiques i d'organització ja que, a través d'aquesta, es podia indicar l'assistència o no a la conferència d'obertura i cloenda.

El nombre de persones inscrites per a les conferències d'obertura i cloenda fou de 37 i 32 persones respectívament.

Tota la informació del Cicle OEIAC es va promocionar a través de les webs institucionals i del compte que té l'OEIAC a la plataforma X/Twitter (@OEIAC_UdG).

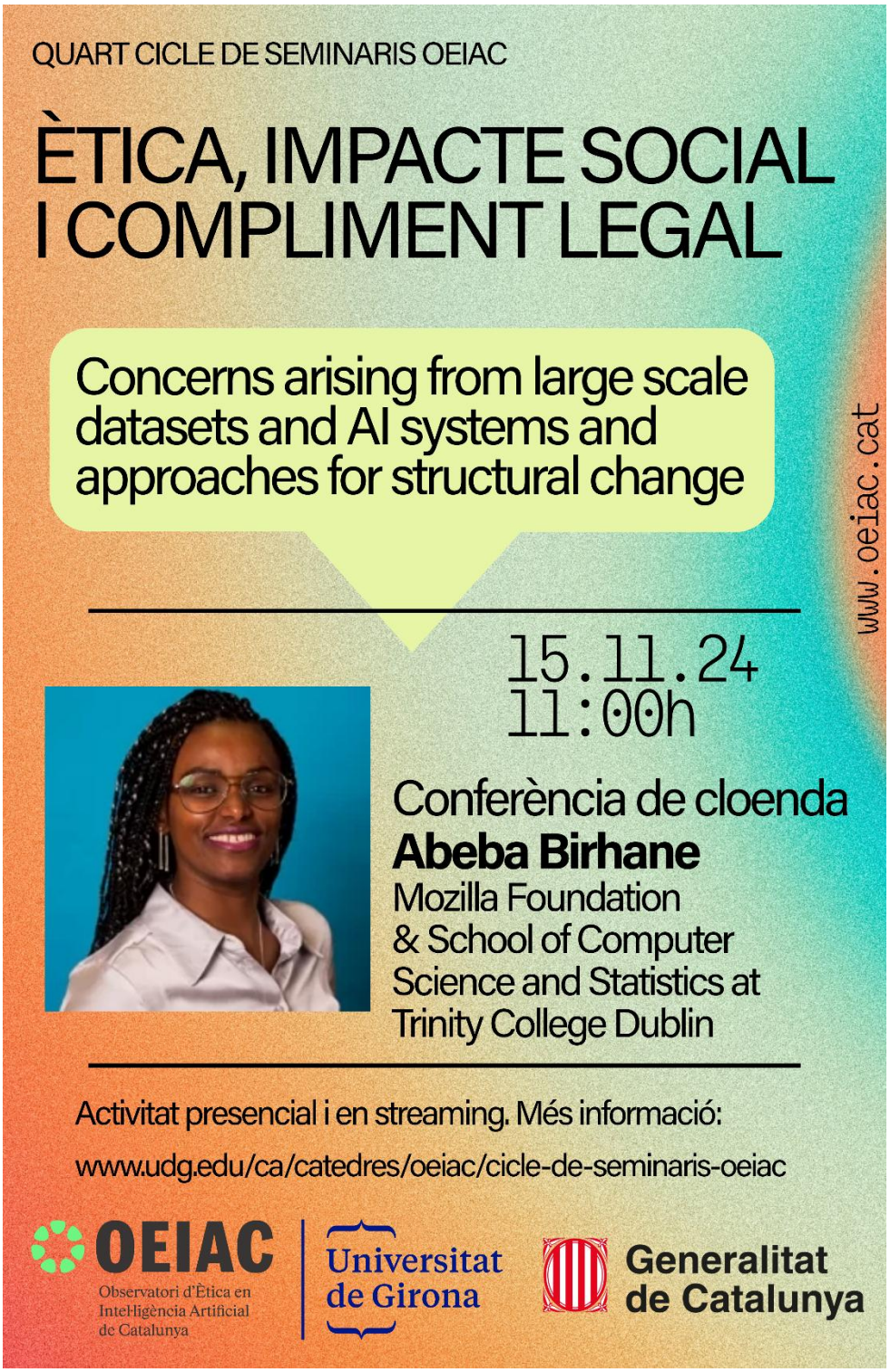
Pòster del IV Cicle de Seminaris OEIAC

Pòster de la conferència d'obertura del IV Cicle de Seminaris OEIAC

Pòster del segon seminari (virtual) del IV Cicle de Seminaris OEIAC

Pòster del tercer seminari (virtual) del IV Cicle de Seminaris OEIAC

Pòster del quart seminari (virtual) del IV Cicle de Seminaris OEIAC



Pòster de la conferència de cloenda del IV Cicle de Seminaris OEIAC

b) Desplegament del Model PIO sobre Obligacions i Drets

Després d'una revisió exhaustiva del contingut i disseny Model PIO beta (veure Figura 1) i a la llum de l'aprovació recent de la normativa en materia de riscos de la IA a la Unió Europea (AI Act), l'OEIAC ofereix un nou Model PIO sobre Obligacions i Drets (veure Figura 2) que té els següents objectius:

- Afavoreix el compliment de la normativa i regulació actual sobre riscos associats a la intel·ligència artificial a través d'un procés de verificació exhaustiu.
- Permet identificar accions adequades o inadequades i sensibilitzar a la quàdruple hèlix a través dels usos ètics i responsables de les dades i sistemes d'intel·ligència artificial.
- S'alinea amb la creixent adopció internacional de principis ètics i d'estàndards d'alt nivell en el disseny, implementació i ús de dades i sistemes d'intel·ligència artificial.

Figura 1

El Model PIO sobre Obligacions i Drets substitueix el present Model PIO beta a través del domini www.oeiac.cat, si bé segueix basant-se en 7 principis ètics fonamentals, com són la transparència i explicabilitat; la justícia i equitat; la seguretat i no maleficència; la responsabilitat i retiment de comptes; la privacitat; l'autonomia; i la sostenibilitat (veure Figura 2a i Figura 2b).

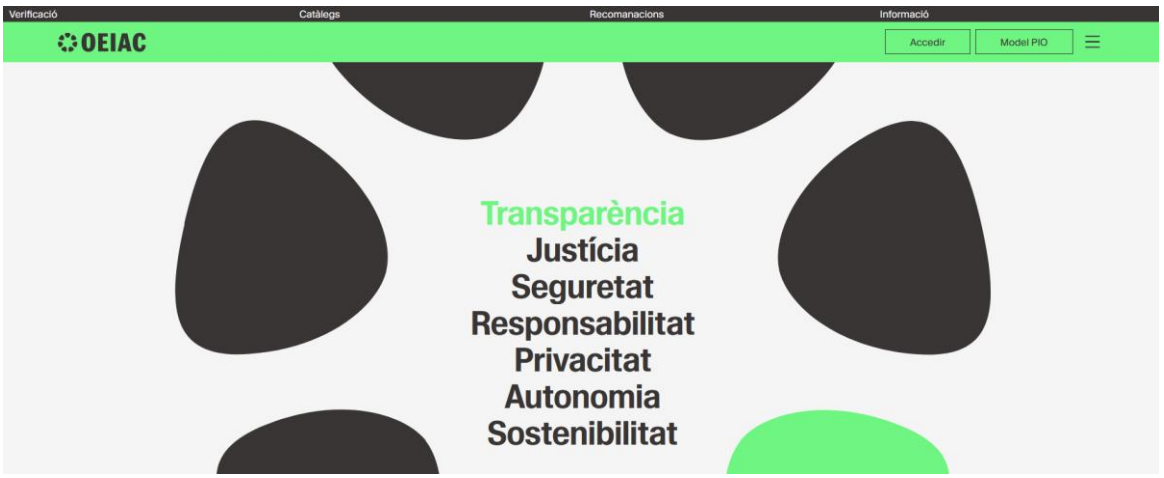


Figura 2a

Figura 2b

El nou Model PIO sobre Obligacions i Drets permet realitzar una preavaluació segons les categories de risc contemplades en la normativa sobre intel·ligència artificial de la UE (AI Act) i classificar els sistemes d'IA segons si són considerats de risc inacceptable a partir de la normativa de la AI Act i, després, d'acord amb la categoria de risc:

- Risc alt: aquest fa referència a un nombre limitat de sistemes d'IA que poden generar un impacte advers en la seguretat de les persones o els seus drets fonamentals segons la Carta dels Drets Fonamentals de la UE i, per tant, es consideren d'alt risc.
- Risc limitat: aquest fa referència als riscos associats a la falta de transparència en l'ús de la IA i, per tant, a la necessitat d'obligacions específiques de transparència per garantir que els humans estiguin informats quan sigui necessari, fomentant la

confiança.

- Risc mínim: aquest fa referència a tots els altres sistemes d'IA que es poden desenvolupar i utilitzar d'acord amb la legislació vigent sense obligacions legals addicionals. Voluntàriament, els proveïdors d'aquests sistemes poden optar per aplicar principis ètics com els del Model PIO.

El resultat d'aquesta preavaluació centrada en l'avaluació de riscos de la AI Act, permet posicionar als usuaris del Model PIO Obligacions i Drets en un punt de partida sobre l'avaluació d'alguns dels principals requisits legals de la normativa central d'IA de la Unió Europea.

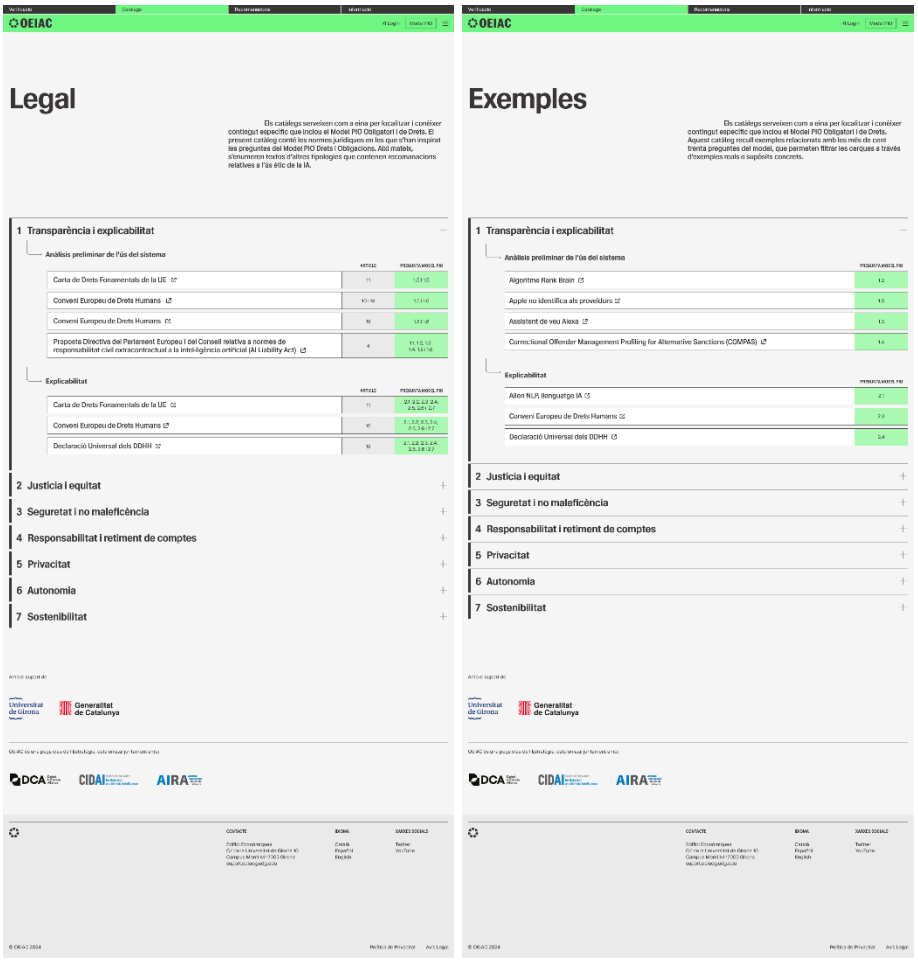
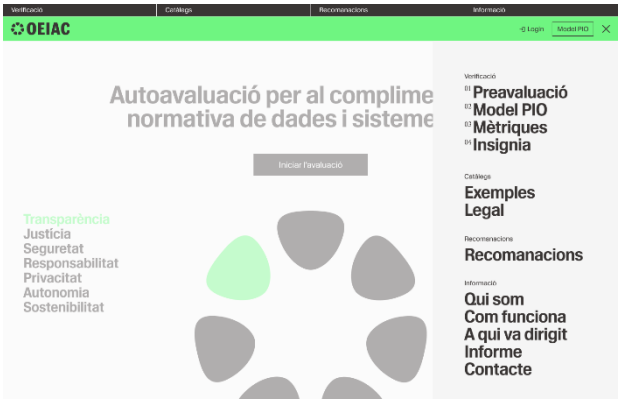


Figura 3

El Model PIO sobre Obligacions i Drets incorpora una sèrie de recursos nous, com catàlegs que serveixen per localitzar i conèixer contingut específic que inclou el mateix model. En aquest sentit, s'han desenvolupat dos catàlegs (veure Figura 3): el primer relatiu a la normativa legal i a les recomanacions ètiques existents que han configurat les qüestions plantejades al model. El segon catàleg recull exemples relacionats amb les més de cent trenta preguntes del model, que permeten filtrar les cerques a través d'exemples reals o supòsits concrets.

Ambdós catàlegs segueixen l'ordre del Model PIO Obligacions i Drets, és a dir, la categorització a partir dels set principis i la classificació dels apartats corresponents. No obstant, la informació classificada també està ordenada de manera alfabètica per facilitar la navegació. Pel que fa al catàleg legal, cada normativa referenciada dona accés al text corresponent, a l'article concret citat i a la pregunta del Model PIO Obligatorietat i Drets on s'esmenta. Al catàleg d'exemples, es troben enllaços a l'article periodístic o acadèmic que recolza o exemplifica el contingut, així com la pregunta del Model PIO Obligatori i Drets on apareix.

Aquest enfocament estructurat garanteix que les consideracions ètiques, semblants a les destacades per marcs com la Llei d'IA de la UE i les directrius acadèmiques sobre l'ètica de la IA, s'integrin en la comprensió i aplicació del contingut, reforçant la importància del compliment legal i la innovació ètica en el desenvolupament i implantació de sistemes d'IA.



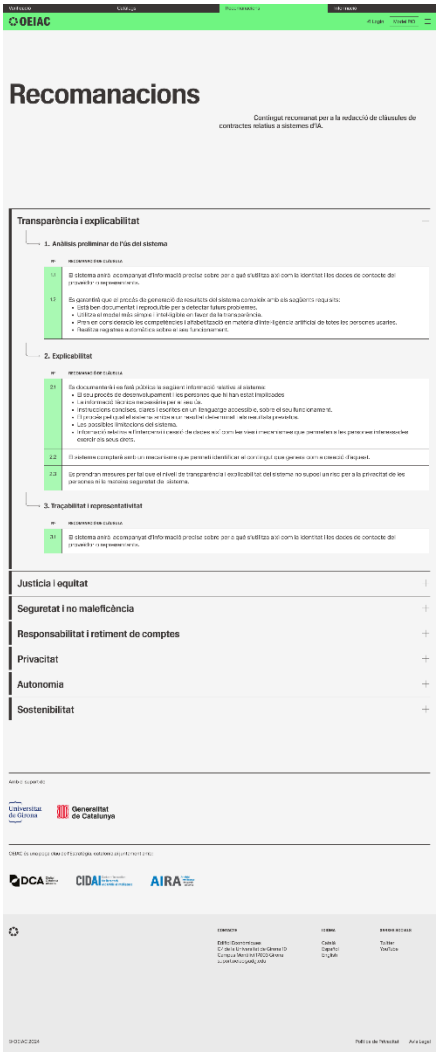


Figura 4

Juntament amb els catàlegs, el Model PIO sobre Obligacions i Drets proporciona clàusules de recomanació ètica per a la compra i adquisició de sistemes d'IA (veure Figura 4). Aquestes clàusules, que incorporen els requisits legals i els estàndards i recomanacions ètiques, serveixen com a pautes perquè les organitzacions públiques i privades les incorporin en els seus processos de contractació. Aquestes clàusules de recomanació giren al voltant dels 7 principis bàsics del Model PIO Obligacions i Drets.

En aquest sentit, conté clàusules de transparència i requereix que els proveïdors proporcionin una documentació completa sobre el disseny, els algorismes i els processos de presa de decisions del sistema d'IA; clàusules d'equitat que estipulen que els sistemes d'IA s'han de sotmetre a processos de detecció i mitigació de biaix, demostrant el compromís amb la justícia i l'equitat en els resultats per a tots els usuaris; clàusules de seguretat que exigeixen proves d'avaluacions de riscos rigoroses i la implementació de mesures de seguretat per prevenir danys; clàusules de retiment de comptes que insisteixen en marcs de responsabilitat clars dels venedors, que han de detallar els rols i les responsabilitats en el desenvolupament i el desplegament de sistemes d'IA; clàusules de protecció de privadesa que exigeixen el compliment de les lleis de protecció de dades (p. ex., GDPR) i la implementació dels principis de privadesa des del disseny, salvaguardant la privadesa dels usuaris; clàusules de preservació de l'autonomia per assegurar que els sistemes d'IA incorporen funcions que permetin la supervisió, el control i la intervenció humana per mantenir l'autonomia de l'usuari; i clàusules de sostenibilitat que obliguen a l'avaluació de l'impacte ambiental dels sistemes d'IA i a un compromís amb pràctiques de compromís social que reflecteixen el principi de sostenibilitat de l'Agenda 2030 de les Nacions Unides.

Finalment, en l'apartat de recursos, el Model PIO sobre Obligacions i Drets també inclou informació relativa a les característiques principals que ha de tenir una avaluació de conformitat, que està relacionada amb els sistemes d'IA d'alt risc per demostrar que un sistema compleix els requisits obligatoris per a una IA fiable (per exemple, qualitat de les dades, documentació i traçabilitat, transparència, supervisió humana, precisió, ciberseguretat i robustesa); a les característiques principals que ha de tenir una avaluació d'impacte en els drets fonamentals, que està relacionada amb els sistemes d'IA d'alt risc per demostrar si existeix o no un impacte en els drets fonamentals i notificar els resultats a l'autoritat nacional; i a les característiques principals que han de tenir les obligacions de transparència, tenint en compte que aquestes estan relacionades amb tots els sistemes d'IA.

En general, la recollida d'informació del Model sobre Obligacions i Drets es realitza a través de respostes binàries (afirmatives o negatives) i que no apliquen segons la informació que tenen els mateixos usuaris del model per realitzar, a posteriori, un còmput i fotografia del grau de compliment en termes d'obligacions i drets segons els requisits legals i estàndards ètics (veure Figura 5a , Figura 5b i Figura 5c).

Amb la informació obtinguda a través de la preavaluació i avaluació, s'obtindran mètriques globals, matrius de riscos i punts de millora relacionats amb el grau d'adequació a cada un dels 7 principis avaluats i també segons els temes i categories de risc de les dades i sistemes d'IA avaluats. Aquests resultats, supeditats a la pròpia declaració de responsabilitat, mitjançant la qual, l'organització es fa responsable de la veracitat de les respostes aportades, serà possible crear perfils observables de cada una de les entitats que portin a terme l'avaluació de les seves dades i sistemes d'IA.

El darrer pas un cop obtingudes les mètriques corresponents, és el de l'elaboració d'una insígnia que és la imatge gràfica que representa l'adequació de l'entitat als requisits legals i estàndards ètics segons el Model PIO Obligacions i Drets a través dels set principis del model. La mateixa insígnia és descarregable per tal que es pugui reflectir i evidenciar l'adequació i capacitació d'un sistema d'IA d'acord amb el Model PIO Obligacions i Drets.

El Model PIO Obligacions i Drets s'elabora amb la simplicitat en el seu nucli, girant al voltant d'una pregunta fonamental i eficaç que qualsevol organització pública o privada implicada en el desenvolupament, la gestió o la supervisió de projectes que incorporen dades i sistemes d'IA pot entendre fàcilment: Ho hem fet?

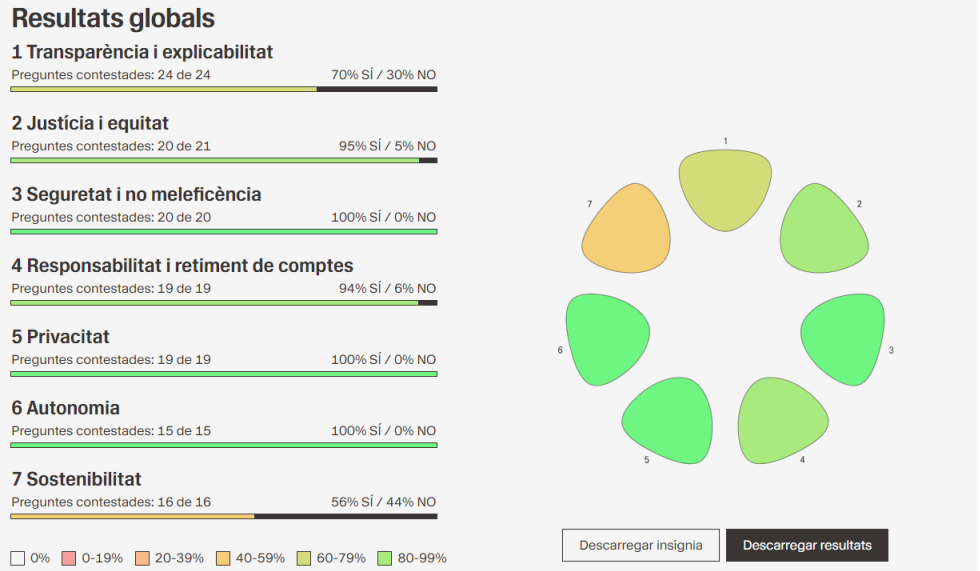


Figura 5a

	SEVERITAT SEGONS LA NORMATIVA O AI ACT		
PROBABILITAT DE PORTAR A TERME UNA ACCIÓ	Mínima	Limitada	Alta
Molt alta	1	0	4
Alta	2	0	4
Ni alta ni baixa	1	0	0
Baixa	1	1	2
Molt baixa	0	0	0

Figura 5b

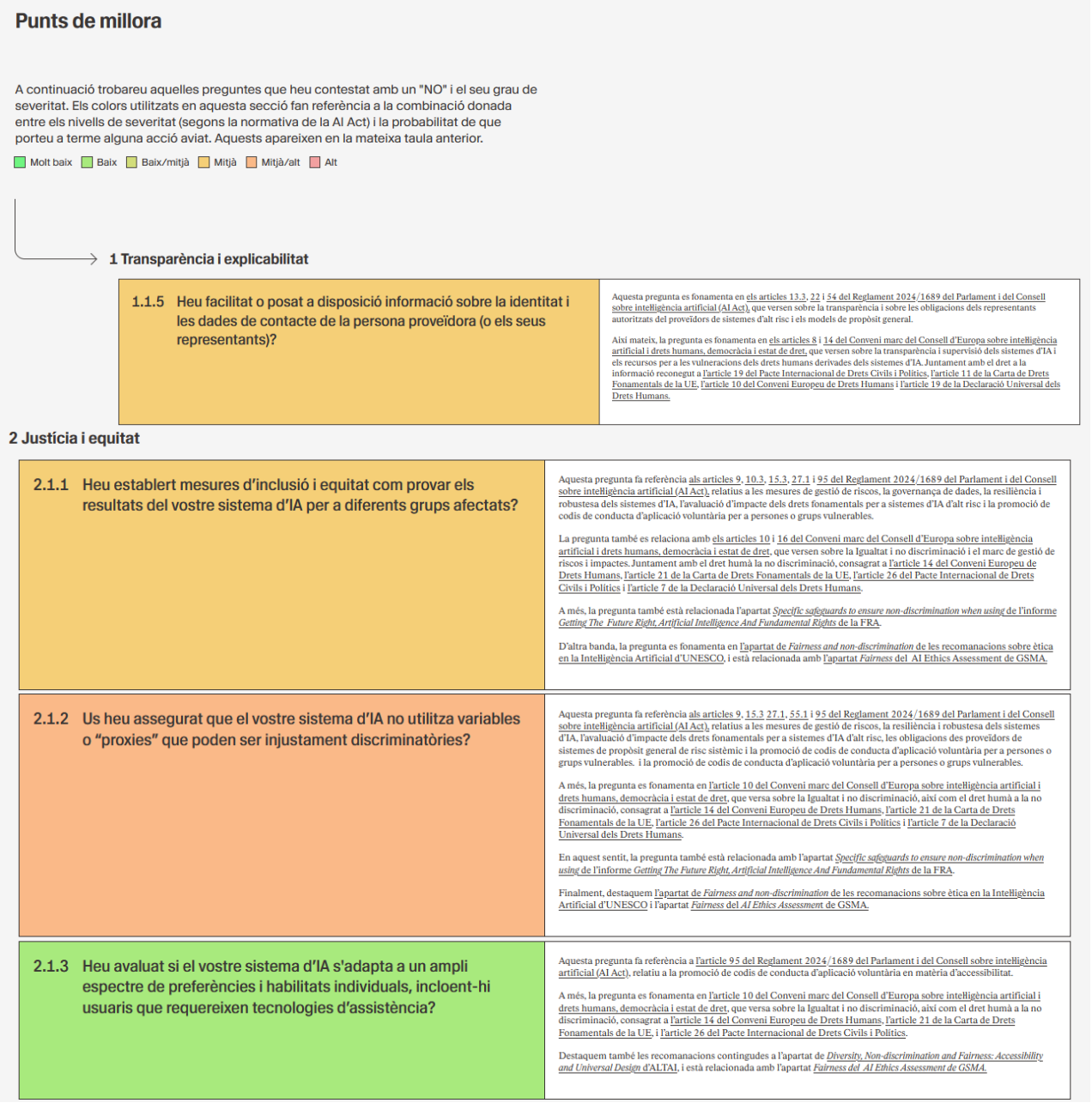


Figura 5c

Partint de la premissa que una de les avaluacions més crítiques que poden dur a terme les organitzacions públiques i privades es refereix a la seva utilització de dades i sistemes d'IA, el Model PIO posa l'accent en la importància d'aquestes avaluacions. Independentment de si les decisions organitzatives es perceben (tant internament com externament) com a molt positives o negatives, hi ha una necessitat creixent que aquestes decisions estiguin alineades amb els principis ètics rellevants per al context. A més, és imprescindible que aquesta alineació es comuniqui amb transparència i s'implementi amb la màxima precisió possible, especialment a la llum de la nova legislació sobre IA com la AI Act.

Des del punt de vista de la responsabilitat social corporativa, el Model PIO es presenta com una eina versàtil, aplicable a diverses etapes del cicle de vida de les dades i sistemes d'IA. Això abasta la fase de disseny i modelització, que inclou la planificació, la recollida de dades

i la construcció del model; la fase de desenvolupament i validació, que inclou la formació i la prova de models; i la fase de desplegament, seguiment i perfeccionament, que implica la resolució de problemes i la millora contínua. Tal i com es justifica en l'informe "El Model PIO (Principis, Indicadors i Observables) sobre Obligacions i Drets: un enfocament d'autoavaluació per al compliment de la regulació en dades i sistemes d'intel· ligència artificial en el marc de la Unió Europea", aquest model promou el compliment i la preparació per a l'ús responsable de la IA mitjançant una llista de verificació d'autoavaluació completa. També identifica les accions adequades a través de diverses qüestions vinculades a normatives i estàndards ètics, i consciencia a la quàdruple hèlix sobre l'ús responsable de les dades i els sistemes d'IA. A més, cal dir que es tracta d'una llista de verificació completa que ajuda a classificar el nivell de risc d'un sistema d'IA i facilita la comprensió del que s'ha de fer en virtut de la Llei d'IA de la UE, que es pot descarregar de forma gratuïta a través de la pàgina web de l'OEIAC (www.oeiac.cat) i de la següent referència:

Lillo, A.; Brufau, A. & Sabater, A. (2025) El Model PIO (Principis, Indicadors i Observables) sobre Obligacions i Drets: un enfocament d'autoavaluació per al compliment de la regulació en dades i sistemes d'intel· ligència artificial en el marc de la Unió Europea. Girona: Universitat de Girona, https://dugi-doc.udg.edu/handle/10256/26104.

➔ Des de la seva publicació, aquest treball a generat 501 vistes i 128 descàrregues a través de la pàgina web del servei de publicacions de la Universitat de Girona.

➔ Des de la seva posada en funcionament a finals de juny de 2024, s'han portat a terme un total de 256 autoavaluacions completes o finalitzades del Model PIO a través de la interfície web.

Per tal de promocionar el nou Model PIO s'han realitzat una sèrie de vídeos de presentació sobre les característiques principals d'aquesta eina per avançar en l'avaluació ètica de dades i sistemes d'IA a través d'un formulari de verificació o checklist.

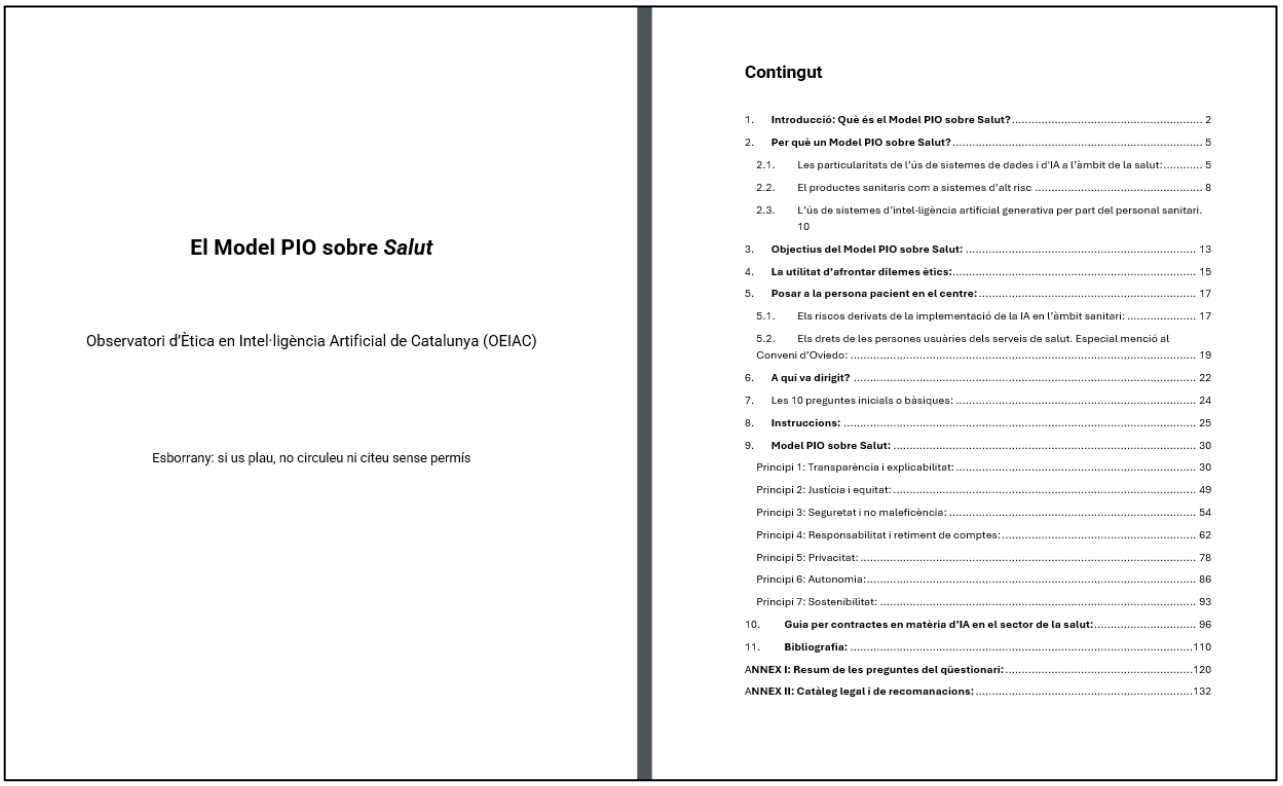
➔ Des de la seva publicació, el vídeo promocional (que ara s'ha traduït al castellà i anglès) ha tingut 163 vistes a través del canal YouTube de la Universitat de Girona: https://youtu.be/w7QWbWnSsUw?feature=shared

c) Disseny i elaboració dels Model PIO Sectorials (Salut i Educació)

Sota el paraigua del Model PIO, l'OEIAC ha dissenyat i elaborat dos Models Sectorials que, seguint mateix format i metodologia, posen el focus en diferents aspectes relacionats amb l'ús de la IA per tal d'adaptar el model segons les necessitats, especialment en matèries i àmbits clau com són el de la salut i l'educació.

En aquest context sorgeix el Model PIO sobre Salut, presentat en aquest informe. I és que l'àmbit de la salut, lluny de ser una excepció, també experimentant l'aparició de diferents eines que utilitzen, o són en si mateixes, sistemes d'IA, les quals tenen com a objectiu assistir, o fins i tot automatitzar, diferents aspectes de la prestació de serveis sanitaris. D'aquesta realitat, sorgeixen riscos particular derivats de la implementació de la IA en àmbit sanitari, que requereixen d'un abordament especialitzat, com s'explicarà en seccions posteriors.

Així doncs, el Model PIO sobre Salut promou una interrogació detallada i continuada d'aspectes, errors i resultats inesperats, abans i durant el desplegament d'un sistema d'IA. El model d'autoavaluació i les eines associades, anàlogament a altres avaluacions com les auditories, proposen una metodologia per treballar en el desenvolupament d'una IA de confiança. Com descrivim més endavant, el procés d'autoavaluació requereix diferents usuaris i perspectives, per aprofitar habilitats tècniques, així com el coneixement del context, per poder així anticipar-se millor als efectes potencials un cop el sistema es desplega i les vulnerabilitats queden exposades de manera més evident, però amb un risc ja massa elevat i inacceptable (veure Figura 6).



El Model PIO sobre *Salut*

Observatori d'Ètica en Intel·ligència Artificial de Catalunya (OEIAC)

Esborrany: si us plau, no circuleu ni citeu sense permís

Contingut

1. **Introducció: Què és el Model PIO sobre Salut?** ... 2
2. **Per què un Model PIO sobre Salut?** ... 5
   2.1. Les particularitats de l'ús de sistemes de dades i d'IA a l'àmbit de la salut: ... 5
   2.2. El productes sanitaris com a sistemes d'alt risc ... 8
   2.3. L'ús de sistemes d'intel·ligència artificial generativa per part del personal sanitari. 10
3. **Objectius del Model PIO sobre Salut:** ... 13
4. **La utilitat d'afrontar dilemes ètics:** ... 15
5. **Posar a la persona pacient en el centre:** ... 17
   5.1. Els riscos derivats de la implementació de la IA en l'àmbit sanitari: ... 17
   5.2. Els drets de les persones usuàries dels serveis de salut. Especial menció al Conveni d'Oviedo: ... 19
6. **A qui va dirigit?** ... 22
7. Les 10 preguntes inicials o bàsiques: ... 24
8. **Instruccions:** ... 25
9. **Model PIO sobre Salut:** ... 30
   Principi 1: Transparència i explicabilitat: ... 30
   Principi 2: Justícia i equitat: ... 49
   Principi 3: Seguretat i no maleficència: ... 54
   Principi 4: Responsabilitat i retiment de comptes: ... 62
   Principi 5: Privacitat: ... 78
   Principi 6: Autonomia: ... 86
   Principi 7: Sostenibilitat: ... 93
10. **Guia per contractes en matèria d'IA en el sector de la salut:** ... 96
11. **Bibliografia:** ... 110

ANNEX I: Resum de les preguntes del qüestionari: ... 120
ANNEX II: Catàleg legal i de recomanacions: ... 132

Figura 6

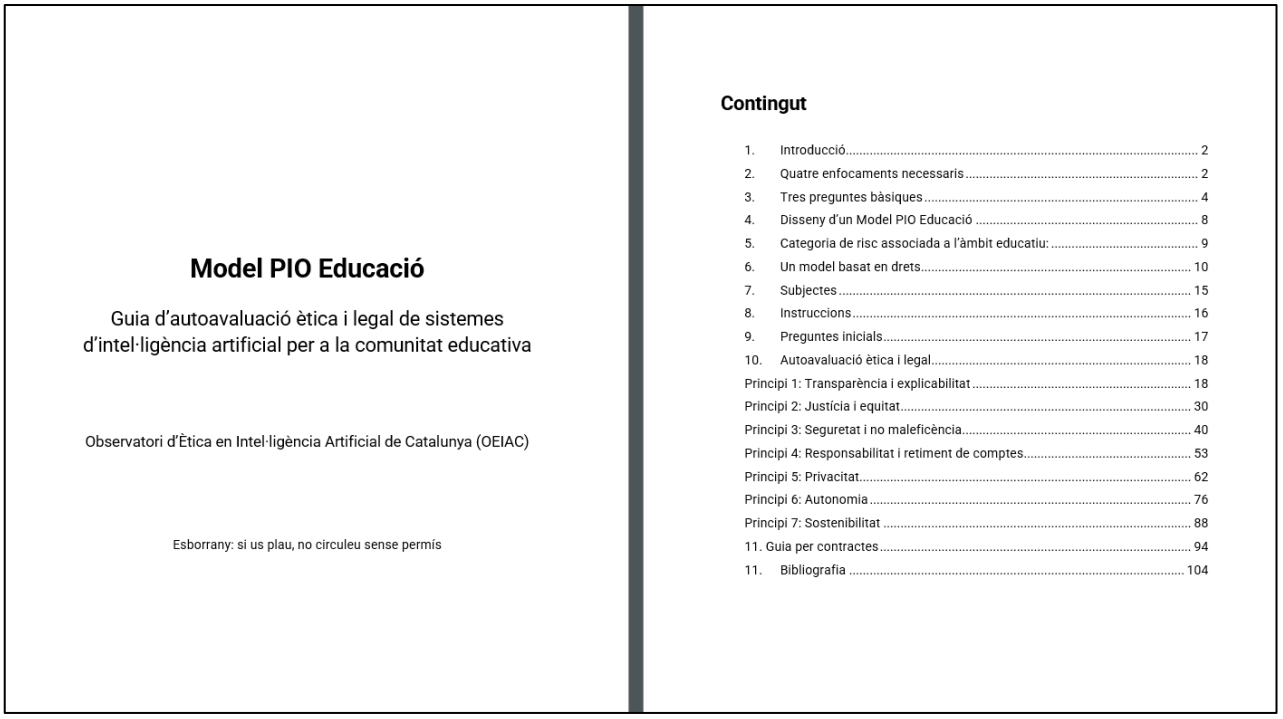
D'altra banda, un dels principals factors que fan que molts sistemes de dades i d'IA desenvolupats mai arribin a tenir un impacte efectiu és el conegut com a "*performance gap"*: la diferència de rendiment entre el sistema d'IA quan es prova al laboratori i quan s'avalua ja en un entorn de desplegament real. Aquesta diferència és deguda a causes diverses segons cada cas, però té generalment un impacte notable en sectors com el de la salut. Necessitem eines i estratègies per identificar i prevenir aquesta diferència de rendiment, i el Model PIO sobre Salut és una de les metodologies reconegudes per identificar el problema i trobar vies per mitigar-lo.

Els models en què estan basats els sistemes d'IA cobreixen moltes tipologies, i no només les xarxes neuronals i d'aprenentatge profund (*Deep Learning*), amb els quals darrerament es tendeix a associar qualsevol sistema d'IA. El Model PIO és aplicable a qualsevol sistema d'IA, independentment de si es tracta d'un sistema d'aprenentatge automàtic simple, basat en regles fixes; si és un model lineal, o basat en arbres de decisió; o si, en contrast, és un sistema més complex com els de les grans xarxes neuronals profundes, incloent-hi els grans models de llenguatge. En aquest sentit, el Model PIO té l'avantatge de ser flexible i d'abast general.

A més de les consideracions ètiques i aspectes sobre la robustesa tècnica d'un sistema d'IA, els Models PIO prenen com a principal fonament el contingut del Reglament europeu d'Intel·ligència Artificial, que entrà en vigor el 2 d'agost de 2025, més conegut com a RIA o AI Act. A part d'aquesta norma, es van prendre com a referència altres textos legislatius rellevants, com per exemple el Reglament general de protecció de dades (RGPD), entre molts altres.

En el cas del Model PIO sobre Salut, a les regles ja introduïdes a al Model PIO sobre Obligacions i Drets, s'ha afegit regulacions específiques sobre els dispositius mèdics, en concret, el Reglament 2017/745, sobre els productes sanitaris, i el Reglament 2017/746 sobre productes sanitaris per a diagnòstic in vitro. Així mateix, s'han pres en consideració normes relatives als drets de les persones usuàries dels serveis de salut, com Llei bàsica reguladora de l'autonomia del pacient i de drets i obligacions en matèria d'informacions i documentació clínica; regles de funcionament del sistema sanitari com la Llei General de Salut Pública i la Llei General de Sanitat; i altres lleis i reglaments que afectaven d'una o una altra manera la prestació dels serveis de salut.

D'aquesta manera, el Model PIO sobre la Salut, més enllà de d'avaluar si els sistemes s'adhereixen a les directrius ètiques determinades per autoritats en la matèria, permet fer una primera revisió sobre la conformitat del sistema respecte del marc regulador actual.



Model PIO Educació

Guia d'autoavaluació ètica i legal de sistemes d'intel·ligència artificial per a la comunitat educativa

Observatori d'Ètica en Intel·ligència Artificial de Catalunya (OEIAC)

Esborrany: si us plau, no circuleu sense permís

Contingut

1. Introducció ... 2
2. Quatre enfocaments necessaris ... 2
3. Tres preguntes bàsiques ... 4
4. Disseny d'un Model PIO Educació ... 8
5. Categoria de risc associada a l'àmbit educatiu: ... 9
6. Un model basat en drets ... 10
7. Subjectes ... 15
8. Instruccions ... 16
9. Preguntes inicials ... 17
10. Autoavaluació ètica i legal ... 18

Principi 1: Transparència i explicabilitat ... 18
Principi 2: Justícia i equitat ... 30
Principi 3: Seguretat i no maleficència ... 40
Principi 4: Responsabilitat i retiment de comptes ... 53
Principi 5: Privacitat ... 62
Principi 6: Autonomia ... 76
Principi 7: Sostenibilitat ... 88
11. Guia per contractes ... 94
11. Bibliografia ... 104

Figura 7

En el context dels Models PIO Sectorals també s'ha dissenyat i elaborat el Model PIO sobre Educació (veure Figura 7) ja que, si bé des de fa temps la tecnologia ha estat una eina clau en l'evolució de l'educació, facilitant l'accés al coneixement i la diversificació de mètodes d'ensenyament, l'arribada de la IA marca un punt d'inflexió. Per què? Doncs perquè no tan sols introdueix capacitats i potencials fins ara inèdits, sinó que també exigeix una reflexió profunda sobre el seu ús ètic i el seu impacte social. En aquest sentit, el Model PIO sobre Educació és cabdal per establir uns principis ètics sòlids que orientin la integració de la tecnologia en l'àmbit educatiu, garantint que la seva aplicació beneficia tant al professorat com a tot l'alumnat de manera justa i equitativa. Aquests principis del Model PIO sobre Educació permeten portar a terme un esforç col·lectiu, no només a través de les polítiques públiques i el sistema escolar que tenen un paper protagonista, sinó també mitjançant les persones que desenvolupen tecnologies d'IA, els quals han de ser conscients de l'impacte social i educatiu de la tecnologia.

Mentre la responsabilitat de guiar aquesta transformació recau en tota la comunitat educativa, famílies i la societat en general, és evident que en aquesta etapa inicial cal posar l'accent especialment en l'elaboració i implementació de polítiques educatives adaptades a un context en que el rol del professorat és cabdal, ja que són aquest grup els que han de liderar la transformació digital a les aules. Sense aquest enfocament serà molt difícil aprofitar algunes de les avantatges de la IA, alhora que minimitzem els seus riscos i fomentem un desenvolupament tecnològic alineat amb els enfocaments de pensament ètic i educatiu.

Però això requereix d'una autoavaluació constant i sistemàtica aplicant una sèrie de principis ètics abans d'utilitzar la IA a l'aula. Aquests principis són (sempre) els mateixos del Models PIO (transparència, justícia, no maleficència, responsabilitat, privacitat, autonomia i sostenibilitat), que podem descriure de la següent manera en l'àmbit d'interès:

- Transparència: Per assegurar que tant el professorat com l'alumnat entenguin com funciona la IA, les dades que utilitza i com arriba a les seves conclusions. Això pot incloure

qüestions sobre les capacitats de la tecnologia així com la disponibilitat de recursos que expliquen els algoritmes en termes accessibles.

- Justícia: Per garantir que les aplicacions de la IA a l'aula no reprodueixin biaixos existents ni creïn noves desigualtats. Això implica revisar i ajustar els algoritmes per tal d'assegurar-nos que tracten de manera justa a tot l'alumnat, independentment del seu origen, capacitat o situació socioeconòmica.
- Seguretat: Per prioritzar la seguretat de l'alumnat en l'ús de la IA, garantint que les aplicacions no tenen efectes perjudicials en el seu benestar físic, psicològic o emocional. Això requereix una avaluació rigorosa dels riscos i la implementació de mesures de protecció adequades.
- Responsabilitat: Per establir mecanismes clars de responsabilitat per a les persones desenvolupadores i usuàries de l'IA en l'educació. Aquest principi inclou l'assignació clara de les responsabilitats i els processos que es porten a terme per abordar qualsevol problema o incidència que pugui sorgir.
- Privacitat: Per protegir la privacitat de l'alumnat mitjançant la gestió sensible de les seves dades. Això comporta l'ús de protocols de seguretat de dades i assegurar que l'alumnat i el professorat tenen control sobre com les dades personals es recopilen, utilitzen i comparteixen.
- Autonomia: Per respectar l'autonomia de l'alumnat i professorat en la presa de decisions educatives, assegurant que la IA sigui una eina de suport i no un determinant social de les trajectòries educatives. En aquest sentit, cal desenvolupar la capacitat de qüestionar o rebutjar les recomanacions fetes per sistemes basats en IA.
- Sostenibilitat: Per abordar l'impacte ambiental de les tecnologies d'IA i treballar per minimitzar la seva petjada ecològica. Això pot incloure plantejaments sobre la presència i absència de la tecnologia així com l'optimització de l'eficiència energètica dels sistemes d'IA i la selecció de programari d'IA que sigui respectuós amb el medi ambient.

d) Sessions de bones pràctiques amb tots els Models PIO disponibles

Després del desenvolupament del Model PIO sobre Obligacions i Drets, les sessions de bones pràctiques han constituït un pas necessari i imprescindible per introduir i formar els participants en l'avaluació ètica de les dades i els sistemes d'IA. Aquestes sessions, que han variat de format, des de tallers d'una hora fins a mòduls de formació de mig dia, han assegurat la flexibilitat per adaptar-se a diferents públics i necessitats. Més enllà de la instrucció tècnica, també han servit de plataforma per explicar els antecedents del Model PIO i les activitats més àmplies de verificació ètica que duu a terme l'OEIAC. Amb aquesta iniciativa de bones pràctiques hem posat de manifest la necessitat creixent de marcs d'IA ètics, no només com a directrius teòriques, sinó com a eines accionables amb un important valor social i econòmic, tant a Catalunya com a escala internacional.

Les sessions han estan dissenyades per afavorir una comprensió més profunda de l'ètica de la IA a la pràctica. Això significa que s'han cobert temes com la mitigació de biaixos, la transparència, la responsabilitat i el compliment de les normatives com la Llei d'IA de la UE. Mitjançant la integració d'estudis de casos reals i alguns exercicis, els participants de les sessions de bones pràctiques han obtingut una experiència rica en l'avaluació de sistemes d'IA

a través d'una perspectiva d'IA responsable i ètica. A més, aquests tallers han destacat el paper de l'OEIAC en l'auditoria d'aplicacions d'IA, la promoció dels estàndards de certificació i la col·laboració amb entitats del sector públic i privat en aquesta matèria. En aquest sentit, el Model PIO no és només un document estàtic sinó una forma de treballar al costat dels avenços tecnològics i normatius.

Evidentment, el context més ampli d'aquest esforç rau en posicionar els usos ètics de la IA ètica com un imperatiu competitiu i social. A mesura que la IA s'integra cada cop més en sectors crítics (sanitat, finances, administració pública), el seu desplegament ètic és essencial per prevenir danys, generar confiança pública i fomentar la innovació responsable i sostenible. El Model PIO, a través d'aquestes sessions de formació, vol esdevenir un estàndard de referència, i que sigui adoptat no només pels actors locals sinó també per les organitzacions internacionals que busquen directrius ètiques sòlides. La posició proactiva de l'OEIAC i de Catalonia.AI en general en aquest camp pot suposar una millora en termes d'excel·lència com a centre per al desenvolupament responsable de la IA, que faci possible una major atracció d'inversions i associacions de valor afegit així com a referent per a altres persones i territoris.

Totes les sessions de bones pràctiques estan recollides en l'Annex d'activitats, on s'especifiquen els dies i detalls de cada sessió. No obstant, volem destacar les següents:

- ➔ 22 de febrer: Model d'autoavaluació PIO: ètica i intel·ligència artificial. Com estem? Avaluem-nos. Jornada de Govern Obert i Bon Govern.
- ➔ 4 de juny: El Model PIO sobre Obligacions i Drets. Jornada Masterclass CIDAI.
- ➔ 3 d'octubre: eBusiness: Ús ètic de la IA. Jornada Cambra de Comerç de Barcelona.
- ➔ 8 d'octubre: eBusiness: Ús ètic de la IA. Jornada Cambra de Comerç de Tarragona.
- ➔ 16 d'octubre: Jornada eBusiness: Ús ètic de la IA. Jornada Cambra de Comerç de Girona.
- ➔ 17 d'octubre: Jornada eBusiness: Ús ètic de la IA. Jornada Cambra de Comerç de Lleida.
- ➔ 22 de novembre: Jornada sobre IA i la protecció de dades personals. Escola d'Administració Pública de Catalunya.

e) Col·laboració amb iniciatives i avaluacions similars

Aquest any 2024 s'han fomentat les col·laboracions amb estructures consolidades en l'àmbit català com la portada a terme amb l'Ateneu Barcelonès per a la realització de l'acte de cloenda del Cicle OEIAC i, molt especialment, les de caire internacional per tal de visibilitzar les actuacions i principis de l'OEIAC mitjançant els Models PIO disponibles. Aquestes col·laboracions són rellevants per dissenyar i desenvolupar nous projectes de transferencia i recerca conjuntament amb altres organitzacions similars així com per establir acords i avançar cap a l'excel·lència en els usos ètics de les dades i de la intel·ligència artificial alineats amb l'Estratègia CATALONIA.AI.

D'aquestes destaquen la col·laboració amb la Disruptive & Emerging Technology Alliance (DETA), una aliança de governs regionals que representen els principals centres tecnològics del món i que s'han unit en un laboratori de polítiques global únic per treure el màxim profit de la innovació i alhora protegir els drets humans, la democràcia i la diversitat cultural. En aquest context, hem impulsat el "Manifest per avançar en una IA fiable" en el marc de l'Estratègia d'Intel·ligència Artificial de Catalunya, Catalonia.AI que impulsa el Govern a través de la Secretaria de Polítiques Digitals. Aquest manifest és un document sense

precedents a escala global impulsat pel Govern català en col·laboració amb l'OEIAC.

El Manifest presentat i ratificat a principis de novembre de 2024, fa una diagnosi dels reptes i oportunitats de la integració de la IA en la vida quotidiana per establir una sèrie de compromisos basats en uns valors fonamentals (dignitat humana, democràcia i estat de dret) i fer una crida a l'acció per avançar en una IA fiable.

Entre els compromisos que estableix, hi ha el desenvolupament ètic i sostenible; la governança transparent; la innovació inclusiva; i la rendició de comptes i responsabilitat. A partir d'aquí, el document fa una sèrie de recomanacions per fomentar un entorn on la IA es desenvolupi i es governi d'acord amb els valors fonamentals i els compromisos establerts, i apropar-se així a l'objectiu d'assolir una IA "millor"" que reforci la justícia social i contribueixi als objectius més amplis de l'Agenda 2030 per al desenvolupament sostenible.

El document es vehicula a través de la DETA, que agrupa actualment governs com el de la Província de Buenos Aires (Argentina), Baviera (Alemanya), Catalunya (Espanya), Costa Rica, Emília-Romanya (Itàlia), Flandes (Bèlgica), Gyeonggi (Corea del Sud), Hessen (Alemanya), Massachusetts (EUA), Kyoto (Japó), Occitània (França), el Quebec (Canadà), Escòcia (Regne Unit), South Australia (Austràlia), Gal·les (Regne Unit), i Western-Cape (Sud-àfrica).

➔ Podeu veure tot el contingut del manifest a través de la següent web de la DETA: https://detalliance.com/a-manifesto-for-advancing-trustworthy-ai/

Altres col·laboracions internacionals a destacar són les protagonitzades amb la professora Petra Ahrweiler (Johannes Gutenberg-Universität Mainz) per l'elaboració de la publicació "Empowering Societal Engagement in AI: Aligning Ethics, Sustainability, and Development within the Digital Economy", un Policy Brief de la T20 per la trobada del G20 a Brasil (veure Figura 8). Aquest Policy Brief va formar part del conjunt de documentació escollida a presentar als funcionaris del G20 en l'àmbit de potenciar el rol i compromís de la societat en el present i futur de la IA. Malgrat que no es va efectuar el desplaçament previst, si que es van produir nombroses trobades online amb la coautora i amb diversos representats de la T20.

➔ Sabater, A. & Ahrweiler, P. (2024). Empowering Societal Engagement in AI: Aligning Ethics, Sustainability, and Development within the Digital Economy, Policy Brief for T20 Brasil (Río de Janeiro), TF5 Inclusive Digital Transformation, G20 engagement group from G20 members.

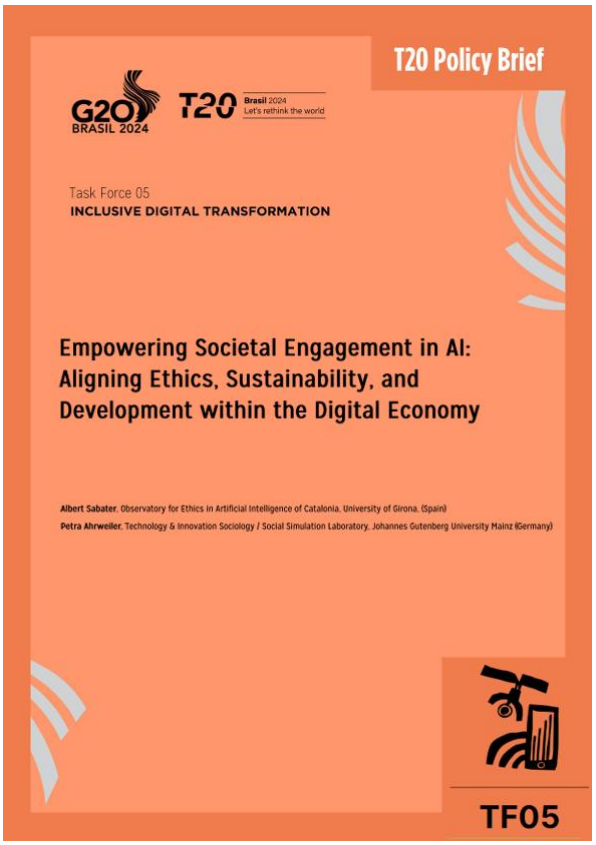


Figura 8

En la mateixa línia, també s'ha col·laborat en la iniciativa "Reclaiming digital sovereignty: A roadmap to build a digital stack for people and the planet", encapçalada per la professora Cecilia Rikap (UCL Institute for Innovation and Public Purpose), per avançar en una digitalització democràtica i dirigida públicament; elaborar una agenda de recerca centrada en desenvolupaments digitals que puguin resoldre problemes col·lectius; i fonamentar la sobirania digital en un internacionalisme ecològic, que es veu com a antídot a la vigilància i als abusos de poder, dos aspectes que minimitzen els recursos necessaris per construir la mateixa digitalització democràtica i pública.

➔ Rikap C. et al. (2024). Reclaiming digital sovereignty: A roadmap to build a digital stack for people and the planet. Disponible a https://www.ucl.ac.uk/bartlett/public-purpose/publications/2024/dec/reclaiming-digital-sovereignty

A collective proposal by the Democratic and Ecological Digital Sovereignty Coalition

Reclaiming Digital Sovereignty

A roadmap to build a digital stack for people and the planet

LEAD AUTHORS

Cecilia Rikap, University College London
Cédric Durand, University of Geneva
Edemilson Paraná, LUT University
Paolo Gerbaudo, Universidad Complutense de Madrid
Paris Marx, Host of Tech Won't Save Us

December 2024

Español | Português

Contributors

Contributing authors

Aaron Benanav, Department of Global Development, College of Agriculture and Life Sciences, Cornell University
Albert Sabater, Observatory for Ethics in Artificial Intelligence of Catalonia, University of Girona
Alessandra De Rossi, University of Turin
Ana Carolina Machado Arrolo
Ana Carolina Sousa Dias
Ana Valdivia, Oxford Internet Institute and UCL Centre for Capitalism Studies
Anahid Bauer, Institute Mines Telecom-Business School, LITEM
Atahualpa Blanchet, Institute of Advanced Studies of the University of Sao Paulo
Benjamin Bürbaumer, Sciences Po Bordeaux
Danilo Spinola, Birmingham City University
David Adler, Progressive International
David Nemer, University of Virginia
Deirdre McQuillan, Technological University Dublin
Dóra Piroska, Central European University
Emilio Cafassi, Universidad de Buenos Aires
Emrys Schoemaker, Caribou Digital / Graduate Institute, Geneva
Fausto Gernone, University College London, Institute for Innovation and Public Purpose
Felipe Lauritzen, Ph.D. candidate, Department of Economics, Sciences Po Paris
Gaia Lucia Gonzalez Barrios, Virginia Tech
Helena Maria Martins Lastres, RedeSist - Research Network on Local Innovation and Production Systems, Federal University of Rio de Janeiro
Jasa Veselinovic, Vrije Universiteit Amsterdam
Jason Hickel, ICTA-Universidad Autónoma de Barcelona and London School of Economics
Javier Diaz-Noci, Pompeu Fabra University
Jonas C. L. Valente, Oxford Internet Institute
José Eduardo Cassiolato, Economics Institute, Federal University of Rio de Janeiro & RedeSist
Juan Carlos De Martin, Nexa Center for Internet & society at Politecnico di Torino
Koen Frenken, Utrecht University
Leila Mucarsel, CEII Universidad Nacional de Cuyo/CONICET
Manuel Scholz-Wäckerle, Vienna University of Economics and Business
Mark Graham, Oxford Internet Institute
Martin Hernacá, Surrey Centre of Digital Economy & Laboratorio de Simulación de Eventos Discretos ICC-UBA/CONICET
Mohammed Amir Anwar, University of Edinburgh
Niels van Doorn, University of Amsterdam
Oscar Arruda d'Alva, Instituto Brasileiro de Geografia e Estatística (IBGE)
Pablo Nabarrete Bastos, Postgraduate Programme in Media and Everyday life of Federal Fluminense University
Paola Ricaurte, Tecnológico de Monterrey
Paul Nemitz, College of Europe
Rafael Cardoso Sampaio, Universidade Federal do Paraná
Rafael Evangelista, Unicamp
Rafael Grohmann, University of Toronto
Rodrigo Moreno Marques, Universidade Federal de Minas Gerais
Rodrigo Santaella Gonçalves, LUT University
Rose Marie Santini, Netlab, Federal University of Rio de Janeiro
Roser Pujadas, University College London
Scott Timcke, Research ICT Africa
Simona Levi, Xnet, Institute for Democratic Digitalisation
Stefano Azzolina, Scuola Superiore Sant'Anna
Tulio Chiarini, Instituto de Pesquisa Economica Aplicada (Ipea)
Ugo Pagano, Università di Siena
Valeria Brusco, Social Sciences Department, National University of Cordoba
Yuri Oliveira de Lima, Federal University of Rio de Janeiro

Figura 9

Finalment, s'ha col·laborat amb els organitzadors de la 17ª edició del Pujiang Innovation Forum, especialment en una ronda de suggeriments de temes a tractar en l'apartat de la cultura de la innovació i del fòrum de discussió. Com que la cultura està estretament lligada a la ciència i la tecnologia, un entorn cultural específic no només és propici per al procés d'innovació tecnològica, sinó que també és fonamental per a la formació d'una cultura sociotècnica. Aquest fòrum va ser un dels moments destacats del 17ª edición del Pujiang Innovation Forum, i es va celebrar al Zhangjiang Science Hall de Xangai (Xina) el passat 8 de setembre.

➔ Podeu trobar més informació sobre aquest esdeveniment anual a la següent adreça: https://english.shanghai.gov.cn/en-17thPujiangInnovationForum/index.html

Figura 10

f) Despeses generals relacionades amb totes les activitats

Per portar a terme totes les activitats presentades, l'OEIAC compta amb tres persones que realitzen el suport tècnic, administratiu i econòmic (Alexandra Lillo, Arlet Brufau i Sònia Senent) i una persona que s'encarrega de la direcció, supervisió i desenvolupament inicial i final de múltiples tasques (Albert Sabater).

L'augment de tasques realitzades en els apartats a), b) c) i d) durant l'any 2024, conjuntament amb una incapacitat de despesa a finals d'aquest mateix any per tasques com el disseny i difussió de nous recursos pedagògics, ha fet que s'hagi prioritzat la despesa en manteniment de personal i en una dedicació plena de l'equip OEIAC en les tasques esmentades.
Per aquest mateix motiu, també ha augmentat la dedicació del director de l'OEIAC al pressupost inicial així com el de la persona tècnica administrativa (Grup A2), que desenvolupa una gran part de les tasques d'operativa (gestió i administració econòmica) de l'OEIAC.

2. Resum d'actuacions i cronograma

a) Nou Cicle de Seminaris OEIAC

a.1. Conferència d'obertura presencial del Cicle OEIAC.
a.2. Seminaris virtuals del Cicle OEIAC.
a.3. Conferència de cloenda presencial del Cicle OEIAC.

b) Desplegament del Model PIO sobre Obligacions i Drets

b.1. Renovació de la interfície web per l'allotjament del nou Model PIO.
b.2. Actualització del Model PIO beta a través d'una remodelació sociolegal.
b.3. Elaboració i publicació de l'informe del nou Model PIO.

c) Disseny i elaboració dels Models PIO Sectorials

c.1. Definició i configuració dels Models PIO Sectorials (Salut i Educació).
c.2. Incorporació beta dels Models PIO Sectorials a la interfície web.

d) Sessions de bones pràctiques amb tots els models PIO disponibles

d.1. Disseny i preparació de seminaris i tallers d'introducció i capacitació als Models PIO.
d.2. Organització amb diferents entitats de seminaris i tallers als Models PIO.

e) Col·laboració amb iniciatives i avaluacions similars

e.1. Cerca d'observatoris o estructures relacionades amb els usos ètics de la IA.
e.2. Cerca de projectes de recerca relacionats amb l'avaluació ètica de sistemes d'IA.
e.3. Intercanvis, estades i "capacity-building" amb les estructures i projectes detectats.

Cronograma

	Mes											
Activitat	1	2	3	4	5	6	7	8	9	10	11	12
a.1												
a.2												
a.3												
b.1												
b.2												
b.3												
c.1												
c.2												
d.1												
d.2												
e.1												
e.2												
e.3												

3. Resum d'activitats de comunicació 2024

Mes	Data	Tema	Medi / Plataforma	Enllaç
Gener	19	I és un gos i és un gat... i aquest conte s'ha acabat? La importància creixent de l'etiquetatge en l'era de la IA	ESPAICRÀTER	https://espaicrater.com/explora/iniciatives/els-grans-interrogants-de-la-ciencia/albert-sabater-coll-i-es-un-gos-i-es-un-gat-i-aquest-conte-sha-acabat-la-importancia-creixent-de-letiquetatge-en-lera-de-la-ia/
	26	SISCAT and research center managers delve into the field of artificial intelligence in health	Generalitat de Catalunya - Salut	https://iasalut.cat/en/noticia/gerents-del-siscat-i-de-centres-de-recerca-sendinsen-en-el-camp-de-la-intelligencia-artificial-en-salut/
Febrer	5	Intel·ligència artificial al servei de la ciutadania: social, ambiental i econòmica	TEATRE MUNICIPAL DE GIRONA	https://www.girona.cat/agenda/cat/agenda_fitxa.php?id=45141&ta=age
	22	L'OEIAC presenta el Model d'autoavaluació PIO: ètica i intel·ligència artificial. Com estem? Avaluem-nos.	Govern Obert i Bon Govern	https://governobert.diba.cat/node/2275
Març	14	Ètica i intel·ligència artificial	Youtube	https://www.youtube.com/watch?v=9Ifz-u1XC0c
Abril	s/d	Intel·ligència artificial a l'àmbit social - Conversa amb Albert Sabater	DIXIT Centre de Documentació de Serveis Socials	https://dixit.gencat.cat/ca/04recursos/01dossiers_tematics/dossiers-tecnics/intelligencia_artificial_ambit_social/index.html
	s/d	La intel·ligència artificial a Catalunya	Acció	https://www.accio.gencat.cat/web/.content/bancconeixement/documents/informes-tecnologics/ACCIO-ia-catalunya-informe-complet-2024.pdf
	26	Reflexions sobre l'ètica en la Intel·ligència Artificial a l'educació: Albert Sabater al FIET 2024	Youtube	https://www.youtube.com/watch?v=7UAK6tvsbMo
Maig	14	Albert Sabater: "Busquem consideracions ètiques que ens permetin fer servir la IA d'una manera responsable"	Verificat	Albert Sabater: "Busquem consideracions ètiques que ens permetin fer servir la IA d'una manera responsable" - Verificat
	29	Conferència d'obertura a càrrec del professor Alessandro Mantelero	Youtube	Conferència d'obertura a càrrec del professor Alessandro Mantelero - YouTube
	31	La intel·ligència artificial en la música és una eina o una amenaça?	ENDERROCK	https://www.enderrock.cat/noticia/28089/intelligencia-artificial-musica-eina-amenaca
Juny	4	Masterclass: El Model PIO sobre Obligacions i Drets	CIDAI	https://cidai.eu/masterclass-model-pio-sobre-obligacions-i-drets/
	4	Masterclass: El Model PIO sobre Obligacions i Drets	Youtube	https://www.youtube.com/watch?v=kASsncejafg
	10	Segon Seminari del Cicle OEIAC a càrrec de Jamila Venturini	Youtube	https://www.youtube.com/watch?v=FOr4zvWaRCQ
	12	Sabater, A. & Varo, A. (2024). "Dramatizing the future: Theater as a tool for enhancing AI Literacy and education in sustainable futures". In Abegglen, S., Nerantzi, C., Martínez-Arboleda, A., Karatsiori, M., Atenas, J., & Rowell, C. (eds), Towards AI Literacy: 101+ Creative and Critical Practices, Perspectives and Purposes, #creativeHE. https://doi.org/10.5281/zenodo.11613520.	Zenodo	Towards AI Literacy: 101+ Creative and Critical Practices, Perspectives and Purposes
Juliol	8	La IA, un repte i una oportunitat per la llengua	EL TEMPS	https://www.eltemps.cat/opinio/60775/la-ia-un-repte-i-una-oportunitat-per-la-llengua
	8	«El principal perill és pensar que la IA fa les coses millor que nosaltres»	EL TEMPS	https://www.eltemps.cat/article/60780/el-principal-perill-es-pensar-que-la-ia-fa-les-coses-millor-que-nosaltres
	12	Presentada la nueva Unidad ELLIS Barcelona	UAB	https://www.uab.cat/web/sala-de-prensa/detalle-noticia/presentada-la-nueva-unidad-ellis-barcelona-1345830290069.html?detid=1345921419605
	12	Catalunya, pionera en l'impuls d'una IA ètica i innovadora d'acord amb la nova legislació europea	Govern.cat - Polítiques Digitals	https://govern.cat/salapremsa/notes-premsa/625162/catalunya-pionera-limpuls-duna-ia-etica-innovadora-dacord-nova-legislacio-europea
	29	Nou Model d'autoavaluació PIO (Principis, Indicadors i Observables)	Youtube	https://www.youtube.com/watch?v=w7QWbWnSsUw
	31	El OEIAC crea un modelo para favorecer el cumplimiento de la primera ley de IA de la UE	LA CIUTAT C!	https://laciutat.cat/es/laciutatdegirona-es/oeiac-crea-modelo-primera-ley-ia

Mes	Data	Tema	Medi / Plataforma	Enllaç
Agost	1	L'observatori català OEIAC controlarà el compliment de la llei europea sobre IA	EL NACIONAL	https://www.elnacional.cat/oneconomia/ca/on-ia/observatori-catala-oeiac-controlara-compliment-llei-europea-sobre-ia_1261265_102.html
	1	Alicia Cañellas-Mayor, PhDAlicia Cañellas-Mayor, PhD	LinkedIn	https://www.linkedin.com/posts/acanelma_oeiac-observatori-d%C3%A8tica-en-intellig%C3%A8ncia-activity-7217468012491112451-wcLk/
	1	Polítiques Digitals - TIC Catalunya	LinkedIn	https://www.linkedin.com/posts/tic-catalunya-7730342b_oeiac-observatori-d%C3%A8tica-en-intellig%C3%A8ncia-activity-7218626426730647552-HpZ4/?originalSubdomain=es
Setembre	16	Tercer Seminari del Cicle OEIAC a càrrec de Sílvia Casacuberta	Youtube	https://www.youtube.com/watch?v=FKGeHfuGy3s
	26	Barcelona alumbra un manifiesto para una inteligencia artificial feminista y pacifista	elPeriódico	https://www.elperiodico.com/es/tecnologia/20240926/barcelona-alumbra-manifiesto-inteligencia-artificial-feminismo-regulacion-ia-imagenes-108604599
Octubre	3	Jornada eBusiness: Ús ètic de la IA	Gencat - Treball	https://treball.gencat.cat/ca/detalls/activitatagenda/activitatAgenda_01375
	7	La Ley de IA Europea Abre una Era	BNEW	https://www.bnewbarcelona.com/es/session.php?id=32
	7	Conferència "Ètica i Intel·ligència Artificial en la intervenció social"	Youtube	https://www.youtube.com/watch?v=UM-4hXyvjyA
	7	Conferència "Ètica i Intel·ligència Artificial en la intervenció social"	DIXIT Centre de Documentació de Serveis Socials	https://dixit.gencat.cat/ca/detalls/Article/conf_etica_ia_intervencio_social_dixit-tv
	8	Estem parlant de l'ús ètic de la intel·ligència artificial a les empreses amb Albert Sabater Coll, professor de la Universitat de Girona i director de l'Observatori d'Ètica en Intel·ligència Artificial de Catalunya.	Instagram	https://www.instagram.com/cambratarragona/p/DA203TOt1M8/?img_index=3
	8	Jornada ebusiness. Ús ètic de la intel·ligència artificial	Cambra de Comerç de Tarragona	https://www.cambratgn.com/inici/esdeveniments/8589-1641-jornada-ebusiness-us-etic-de-la-intel-ligencia-artificial
	13	La inteligencia artificial sacude los Premios Nobel: ¿por qué el triunfo de Google preocupa a algunos científicos?	elPeriódico	La inteligencia artificial sacude los Premios Nobel: ¿por qué el triunfo de Google preocupa a algunos científicos?
	16	Ús ètic de la IA	Cambra de Comerç de Girona	https://digital.cambragirona.cat/activitats/us-etic-ia/
	17		Cambra de Comerç de Lleida	https://cambralleida.org/producte/17-10-2024-us-etic-de-la-ia-capacitacio-tecnica/
	18	Taula rodona: El costat fosc de la IA	17a edició de la Catosfera	https://www.catosfera.cat/programa/costat-fosc-ia
	21	Quart Seminari del Cicle OEIAC a càrrec de Roser Pujadas	Youtube	https://www.youtube.com/watch?v=RSIajpyP3Ew
	21	IV Cicle de Seminaris de l'Observatori d'Ètica en Intel·ligència Artificial de Catalunya (OEIAC)	Comitè d'ètica de la UPC	https://comite-etica.upc.edu/ca/actualitat/esdeveniments/iv-cicle-de-seminaris-de-lobservatori-detica-en-intel-ligencia-artificial-de-catalunya-oeiac
	23	Google crea un sistema de marcas de agua para detectar textos creados con IA	La Vanguardia	Google crea un sistema de marcas de agua para detectar textos creados con IA
	30	14th ERION EARMA Thematic Group event. Artificial Intelligence and Ethics, the Big Wave	EARMA	https://earma.org/conferences/14th-erion-earma-thematic-group-event/
	31	L'impacte de la intel·ligència artificial, tema central de les X Jornades Catalunya Futura	X Jornades Catalunya Futura	https://plataformacti.cat/limpacte-de-la-intelligencia-artificial-tema-central-de-les-x-jornades-catalunya-futura/

Mes	Data	Tema	Medi / Plataforma	Enllaç
Novembre	12	LAM24 Taula rodona: Els límits de la IA	Youtube	https://www.youtube.com/watch?v=mFFfNCiYD9Q
	15	Dra. Abeba Birhane. Cloenda del IV Cicle de Seminaris de l'Observatori d'Ètica en Intel·ligència Artificial de Catalunya (OEIAC)	Comitè d'ètica de la UPC	https://comite-etica.upc.edu/ca/actualitat/esdeveniments/dra-abeba-birhane-cloenda-del-iv-cicle-de-seminaris-de-lobservatori-detica-en-intel-ligencia-artificial-de-catalunya-oeiac
	15	Cloenda del IV Cicle de Seminaris de l'Observatori d'Ètica en Intel·ligència Artificial de Catalunya (OEIAC)	Ateneu Barcelonès	https://ateneubcn.cat/activitats/cloenda-cicle-oeiac/
	15	Conferència de cloenda a càrrec de la professora Abeba Birhane	Youtube	https://www.youtube.com/watch?v=wA18lmwvKTE
	22	Experts debaten els impactes de la IA en la privacitat	Gencat - ACPD	https://apdcat.gencat.cat/ca/sala_de_premsa/notes_premsa/noticia/Noticia-jornada-IA-EAPC24
	22	Jornada sobre intel·ligència artificial i la protecció de dades personals (10402/2024-1)	Escola d'Administració Pública de Catalunya	https://formacio.eapc.gencat.cat/infoactivitats/AppJava/DetalleActividad.do?codi=10402&ambit=1&edicio=1&any=2024
	22	Jornada sobre intel·ligència artificial i la protecció de dades personals	Youtube	https://www.youtube.com/live/wd6KauY0g8s?app=desktop&cbrd=1
	27	I Congreso Internacional Salud Digital CIP 2024 - Día 1 (Parte 1)	Youtube	https://www.youtube.com/watch?v=6SpQytJgJCA
Desembre	12	Cicle DonaTIC a Girona	Col·laboratori Catalunya	https://colabscatalunya.cat/esdeveniment/cicle-donatic-a-girona/
S/D	s/d	Sabater, A. & Ahrweiler, P. (2024). Empowering Societal Engagement in AI: Aligning Ethics, Sustainability, and Development within the Digital Economy, Policy Brief for T20 Brasil, TF5 Inclusive Digital Transformation, G20 engagement group from G20 members.	T20 Brasil	TF05_ST_05_Empowering_Societal66cf6a3318459.pdf

4. Membres del Consell Assessor

- **Artur Serra**
  Fundació i2CAT
- **Carme Torras**
  Institut de Robòtica i Informàtica Industrial, Consejo Superior de Investigaciones Científicas, Universitat Politècnica de Catalunya
- **David Pereira**
  Head of Data and Intelligence at NTT Data Europe & Latam
- **Elisabet Golobardes**
  Vicerectora d'Ordenació i Qualitat Acadèmica, Universitat Ramon Llull
- **Fernando Vilariño**
  Centre de Visió per Computador, Universitat Autònoma de Barcelona
- **Itziar de Lecuona**
  Observatori de Bioètica i Dret, Universitat de Barcelona
- **Joan Mas**
  Centre of Innovation for Data tech and Artificial Intelligence (CIDAI)
- **Judith Membrives**

  Lafede.cat - Organitzacions per a la Justícia Global
- **Karina Gibert**
  Intelligent Data Science and Artificial Intelligence Research Center, Universitat Politècnica de Catalunya
- **Karma Peiró**
  Periodista experta en noves tecnologies digitals i comunitats virtuals
- **Liliana Arroyo**
  Directora General de Societat Digital, Generalitat de Catalunya
- **Marc Pérez-Batlle**
  Institut Municipal d'Informàtica, Ajuntament de Barcelona
- **Miquel Domènech**
  Departament de Psicologia Social, Universitat Autònoma de Barcelona
- **Montse Guàrdia**
  Directora d'Estratègia del Mobile World Capital i Responsable de l'Àrea de Societat de l'Entitat
- **Nuria Oliver**
  Directora del ELLIS Alicante Foundation i Chief Scientific Adviser at the Vodafone Institute
- **Ramon López de Mántaras**
  Institut d'Investigació en Intel·ligència Artificial, Consejo de Investigaciones Científicas, Universitat Autònoma de Barcelona
- **Ramon Trias**
  AIS Group
- **Ricardo Baeza-Yates**
  Web Science & Social Computing Group, Universitat Pompeu Fabra
- **Ulises Cortés**
  Barcelona Supercomputing Center, Centro Nacional de Supercomputación
- **Xavier Trabado**
  m4Social, Taula del Tercer Sector

5. Persones col·laboradores

- **Abeba Birhane**

  AI Accountability Lab & Trinity College Dublin
- **Alba Canal**
  MACBA
- **Alessandro Mantelero**

  Polytechnic University of Turin and EU Jean Monnet Chair in the Mediterranean Digital Societies & Law
- **Alexa Hagerty**
  Leverhulme Centre for the Future of Intelligence, Centre for the Study of Existential Risk, University of Cambridge
- **Alicia de Manuel**
  NTT DATA
- **Ana Valdivia**
  Investigadora postdoctoral a la Universitat King's College London
- **Ariadna Font**
  Director of Engineering, Cortex Machine Learning Platform, Site Lead, NYC, Twitter
- **Àtia Cortés**
  Investigadora al Barcelona Supercomputing Center (BSC) i co-directora del AI4EU Observatory on Society and AI
- **Beatriz López**
  Coordinadora de l'àrea Medicine and Healthcare del grup de recerca eXIT (Control Engineering and Intelligent Systems), Universitat de Girona
- **Carissa Véliz**
  Associate Professor at the Faculty of Philosophy and the Institute for Ethics in AI, University of Oxford
- **Carlos Castillo**
  ICREA Research Professor a la Universitat Pompeu Fabra i líder del Grup de Recerca sobre Web Science and Social Computing
- **Daniel Innerarity**
  Catedràtic de Filosofia Política, investigador “Ikerbasque” a la Universidad del País Vasco i titular de la càtedra Inteligencia Artificial y Democracia al Instituto Europeo de Florencia
- **Daniel Santanach**
  Centre de Telecomunicacions i Tecnologies de la Informació
- **Jamila Venturini**

  Derechos Digitales
- **Joan Colomer**
  Coordinador del grup de recerca eXIT, Universitat de Girona
- **Joan Manuel del Pozo**
  Professor emèrit i ex-Síndic de la Universitat de Girona
- **Joan Vergés**
  Director de la Càtedra Ferrater Mora de Pensament Contemporani, Universitat de Girona
- **Joana Marí**
  Delegada de Protecció de Dades i Responsable de Projectes Estratègics, Autoritat Catalana de Protecció de Dades

- **Joaquim Melendez**
  Director de la Càtedra Girona Smart City, Universitat de Girona
- **Júlia Pareto**
  Investigadora a l'Institut de Robòtica i Informàtica Industrial, Universitat Politècnica de Catalunya
- **Lorena Jaume-Palasi**
  The Ethical Tech Society
- **Louise Hickman**
  Minderoo Centre for Technology and Democracy, University of Cambridge
- **Lucía Velasco**
  School of Transnational Governance, European University Institute
- **Manel Poch**
  Lequia (Laboratori Enginyeria QUímica I Ambiental), Universitat de Girona
- **Mark Coeckelbergh**
  Professor of Philosophy of Media and Technology, University of Vienna
- **Natàlia Ferrer**
  Professora Asociada de la Facultat de Turisme, Universitat de Girona
- **Núria Vallès**
  Barcelona Science and Technology Studies Group (STS-b), Universitat Autònoma de Barcelona
- **Patrícia Ventura**
  Autora de l'informe del Consell de la Informació de Catalunya "*Algoritmes a les redaccions: reptes i recomanacions per dotar la intel·ligència artificial dels valors ètics del periodisme"*
- **Pere Simón**
  Professor de Dret Constitucional, Universidad Internacional de la Rioja UNIR
- **Rafael Garcia**
  Director Científic del Campus Robòtica, Universitat de Girona
- **Roser Pujadas**

  Department of Science, Technology, Engineering and Public Policy (STEaPP), University College London
- **Sílvia Casacuberta**
  University of Oxfort (Global Rhodes Scholar) & Stanford University
- **Toni Lorente**
  Investigador predoctoral a la Universitat King's College London
- **Petra Ahrweiler**
  President of the European Social Simulation Association, and Director of the Technology and Innovation Sociology / Social Simulation Lab, Johannes Gutenberg University Mainz
- **Wanda Muñoz**
  Experta ante la Alianza Global sobre Inteligencia Artificial, Red Seguridad Humana en América Latina - SEHLAC