

2023



Memòria d'activitats

OBSERVATORI D'ÈTICA EN INTEL·LIGÈNCIA
ARTIFICIAL DE CATALUNYA

Índex de continguts

1. Breu descripció de les activitats.....	4
2. Relació d'activitats 2023	4
a) III Cicle de Seminaris OEIAC.....	5
b) El model PIO “Obligacions i Drets”	9
c) El model PIO “Salut”	12
d) Bones pràctiques, col·laboracions i reunions sobre IA responsable	13
e) Disseny i difusió de nous recursos pedagògics.....	14
3. Resum d'actuacions i cronograma	24
4. Membres del Consell Assessor	25
5. Persones col·laboradores	26



1. Breu descripció de les activitats

L'Observatori 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.

Aquest 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 dels usos ètics de la IA, 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 i socials 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 cinc línies estratègiques per fomentar i garantir els usos ètics de la IA i, a la vegada, per continuar augmentar les oportunitats dels sistemes d'IA en diferents sectors d'activitat alhora que s'aborden els riscos sociotècnics i reptes d'impacte social que suposa la seva implantació d'una forma generalitzada.

Així doncs, de forma resumida, les activitats que s'han portat a terme són: el III Cicle de Seminaris OEIAC sobre la “Implicació amb la sostenibilitat: De l'Antropocè a la Intel·ligència Artificial”; la creació de capacitat i disseny del nou model PIO “Obligacions i Drets” per l'avaluació ètica i legal de sistemes d'intel·ligència artificial; i la concepció i disseny d'un model PIO “Salut” específic per l'avaluació ètica de sistemes d'intel·ligència artificial amb aplicacions a l'àmbit de la salut; sessions de bones pràctiques amb tots els models PIO disponibles; el disseny i difusió de nous recursos pedagògics sobre els objectius i actuacions de l'OEIAC com a pilar de l'Estratègia CATALONIA.AI; l'edició i publicació d'un llibre de divulgació que recollirà la selecció de les preguntes i respostes donades a la pregunta de cap de setmana de l'OEIAC; i el foment de col·laboracions i intercanvis, especialment internacionals, amb persones i estructures similars.

2. Relació d'activitats 2023



a) III 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 passat va portar a terme un cicle sobre “La intel·ligència artificial i presa de decisions ètiques”. Enguany es celebra la tercera edició del Cicle OEIAC amb el tema “Implicació amb la sostenibilitat: De l'Antropocè a la Intel·ligència Artificial”

El context d'aquest nou cicle cal situar-lo en l'interès doble i generalitzat per l'Antropocè i la Intel·ligència Artificial. L'Antropocè és un període complex marcat principalment per l'impacte humà sobre el canvi climàtic i la destrucció d'hàbitats que ens demostra la importància de pensar de manera crítica i acurada sobre tres elements interdependents: els límits planetaris, la vulnerabilitat i l'agència humana. Mentre l'Antropocè avança sense pausa, la Intel·ligència Artificial (IA) accelera la influència humana, el que suposa que, molt probablement, definirà el camí futur de l'Antropocè en una mesura considerable. I, és clar, ho pot fer d'una manera negativa o positiva. De moment, sabem que els futurs geòlegs trobaran proves en l'estratigrafia d'emissions de gasos d'efecte hivernacle produïdes a gran escala pels humans, radionúclids del planeta a partir de proves d'armes nuclears, centenars de tones mètriques de plàstic produïdes anualment i prou formigó per cobrir la superfície del planeta.

De moment, no podem assegurar que la intel·ligència artificial no sigui un element més d'aquesta estratigrafia. No en va, en el darrer i prestigiós informe “AI Index Report” de la Universitat de Stanford s'apunta que l'ús de la intel·ligència artificial i els centres de dades necessaris impliquen un alt consum d'energia, emissions de gasos d'efecte hivernacle, consum d'aigua i un ús desmesurat de materials escassos associats amb la producció, ús i vida útil dels dispositius i infraestructura informàtica. Així doncs, si bé els sistemes d'intel·ligència artificial també es poden utilitzar per optimitzar els recursos en general i el consum d'energia en particular, aquesta no és actualment la prioritat, tal i com ho demostra l'eclosió de la intel·ligència artificial generativa.

La Intel·ligència Artificial com l'Antropocè han entrat en la consciència pública i són conceptes útils en els quals emmarcar la justícia climàtica. Però també ens fan reflexionar sobre si tots els éssers humans són igualment responsables dels danys derivats de l'Antropocè a la Intel·ligència Artificial. En aquest context, la vulnerabilitat al canvi climàtic i la destrucció d'hàbitats no tan sols poden estar subjectes als intercanvis entre mesures efectives de mitigació, també cal situar-les en el debat sobre la justícia ambiental i social a escala mundial, i això requereix, d'entrada, un ús responsable i sostenible de la intel·ligència artificial que deixa enrera la hiperagència de la humanitat en relació amb el planeta.

Per tal d'aportar algunes respostes i, és clar, també per obrir més preguntes sobre la implicació amb la sostenibilitat: de l'Antropocè a la Intel·ligència Artificial, l'OEIAC va iniciar aquest nou cicle de seminaris de transferència de coneixement amb persones convidades d'arreu que treballen i promouen el desenvolupament d'una intel·ligència artificial responsable i sostenible.

El Cicle de Seminaris OEIAC s'inicià el dijous 25 de maig amb una conferència d'obertura presencial i va finalitzar amb una seminari de cloenda virtual el dia 12 de desembre de 2023. Com és habitual, aquest cicle està dirigit a tothom que tingui interès en l'ús i implantació d'una intel·ligència artificial responsable i sostenible en organitzacions públiques i privades.



A continuació trobareu totes les dates del tercer Cicle de Seminaris OEIAC amb totes les persones convidades i les temàtiques exposades:

-25 de maig a les 11:30h, “AI for Climate: Using Political Philosophy to Reflect on the Political Challenges”, conferència d’obertura a càrrec de Mark Coeckelbergh, Catedràtic de Filosofia dels Mitjans i Tecnologia al Departament de Filosofia de la Universitat de Viena.

-29 de juny a les 11:30h, “Sense IA no hi ha RES!”, a càrrec de Joaquim Meléndez, professor titular d’Universitat del Departament d’Enginyeria Elèctrica, Electrònica i Automàtica de la Universitat de Girona.

-7 de juliol de 10h a 13:30h, “Jornada sobre Feminismes i Intel·ligència Artificial”, amb la direcció tècnica de Sofia Trejo (Barcelona Supercomputing Centre), Ulises Cortés (Universitat Politècnica de Catalunya), Karina Gibert (Universitat Politècnica de Catalunya) i Albert Sabater (Universitat de Girona).

-14 de setembre a les 11:30h, “An Agent-Based Model for Poverty and Discrimination Policy-Making”, a càrrec de Nieves Montes, investigadora de l’Institut d’Investigació en Intel·ligència Artificial (IIIA-CSIC).

-16 de novembre a les 18h, Jornada d’Economia i Empresa “Construint Ecosistemes d’Intel·ligència Artificial Responsable i Sostenible”, amb la col·laboració de la Digital Catalonia Alliance i amb la participació de la mateixa DCA, del Centre of Innovation for Data tech and Artificial Intelligence (CIDAI), de SCHUMPIT, PickleTech, EXPAI i Dear Intelligence.

-12 de desembre a les 16h, “Beyond the AI hype: Balancing Innovation and Social Responsibility”, conferència de cloenda a càrrec de Virginia Dignum, Catedràtica de d’Intel·ligència Artificial Responsable a la Universitat d’Umea (Suècia).

Tots els seminaris foren gratuïts i presencials a la Universitat de Girona (exceptuant el darrer de cloenda que tingué lloc en línia). Com és habitual, es va elaborar una pàgina web per aquest nou cicle que s’ha actualitzat amb tots els recursos compartits per part de les persones que han intervingut al llarg del cicle.

Pàgina web del cicle OEIAC 2023:

<https://www.udg.edu/ca/catedres/oeiac/recursos-i-transferencia/cicle-de-seminaris-oeiac>



III CICLE DE SEMINARIS OEIAC
MAIG-NOVEMBRE 2023



Observatori d'Ètica en Intel·ligència Artificial de Catalunya

IMPLICACIÓ AMB LA SOSTENIBILITAT: DE L'ANTROPOCÈ A LA INTEL·LIGÈNCIA ARTIFICIAL

Conferència d'obertura a
càrrec del professor

MARK COECKELBERGH

25.05.23

11:30

Sala de Graus Pep Nadal

Universitat de Girona

www.oaiac.cat

Una iniciativa de:



Generalitat
de Catalunya
**Departament
d'Empresa i Treball**

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

Universitat
de Girona

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



III CICLES DE SEMINARIS OEIAC
Seminari de cloenda

BEYOND THE AI HYPE: BALANCING INNOVATION AND SOCIAL RESPONSABILITY

12 DESEMBRE
16H
JORNADA VIRTUAL



[Formulari d'inscripció](#)



suport.oeiac@udg.edu



www.oeiac.cat

A CÀRREC DE
Prof. VIRGINIA DIGNUM
(Umeå University)



Generalitat de Catalunya
**Departament d'Empresa
i Treball**



Observatori d'Ètica en Intel·ligència Artificial de Catalunya

Universitat
de Girona

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



b) El model PIO “Obligacions i Drets”

L'OEIAC està elaborant una nova eina de verificació o *checklist* que portarà el títol: “*Model PIO Obligacions i Drets: Un enfocament d'autoavaluació per al compliment de la normativa de dades i sistemes d'IA en el marc de la Unió Europea*”.

Els objectius principals d'aquest treball són:

- Sensibilitzar i proporcionar als diferents agents que utilitzen dades i sistemes d'IA, una sèrie de principis fonamentals sobre els quals reflexionar i una manera d'avançar a través d'una proposta d'autoavaluació que estarà alineada amb la recent legislació sobre intel·ligència artificial de la Unió Europea (AI Act) així com d'altres legislacions sobre dades (GDPR), serveis digitals (Digital Services Act), governança de dades (Data Governance Act), i de les principals guies de recomanació ètica, incloent-hi de la OECD i UNESCO.
- Contribuir a l'aplicació de principis ètics alineats amb la legislació i recomanacions ètiques i d'impacte social associats al desenvolupament i implementació de la IA.

El model PIO “Obligacions i Drets” de l'OEIAC estarà disponible mitjançant una plataforma digital web a través del domini www.oeiac.cat (veure Figura 1) i oferirà informació sobre recomanacions ètiques, un catàleg amb disposicions legals a tenir en compte i, és clar, una eina d'avaluació a través d'un formulari d'autoavaluació de posició. Juntament amb tota aquesta informació, també estarà disponible un catàleg de navegació que es podrà utilitzar abans o després de l'autoavaluació així com informació de compra amb clàusules de recomanació en cas que es vulguin adquirir sistemes d'intel·ligència artificial. Donada la naturalesa més legal d'aquest nou treball, aquesta inclourà referències múltiples a la normativa, a guies de recomanació ètica i nombrosos exemples d'aplicació basant-se en 7 principis ètics fonamentals (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).



El model PIO de l'OEIAC



Figura 1

El disseny d'aquest nou treball exhaustiu sobre l'aplicació ètica de principis es situa en un context en que les organitzacions del sector públic i privat requereixen major informació sobre els usos legals de la intel·ligència artificial en el context de la Unió Europea. El model PIO “Obligacions i Drets” de l'OEIAC s'està desenvolupant a través d'una revisió acadèmica molt rigurosa, de caire tècnic i legal per tal que la seva implementació de principis i metodologies de *checklists* relacionades amb l'ús ètic de la IA sigui efectiva i eficaç.

El desenvolupament del model PIO “Obligacions i Drets” de l'OEIAC segueix organitzat en quatre fases (veure Figura 2: Com funciona I). En primer lloc, les organitzacions hauran de respondre a les preguntes seleccionades estratègicament per a cada principi (Fase 1), de manera que amb les respostes s'obtindrà una puntuació d'adequació amb diferents aspectes dels principis ètics depenent de les respostes obtingudes (Fase 2). En aquest sentit, abans de portar a terme l'autoavaluació serà important escollir la persona o persones responsables dins de les organitzacions públiques o privades que utilitzen sistemes d'IA per realitzar, des d'un punt de vista competencial, una autoavaluació amb la màxima informació disponible. Per fer-ho, es recolliran respostes binàries a més de 100 preguntes sobre cada un dels 7 principis anteriorment descrits, de manera que es recolliran dos valors possibles en les respostes (Si/No | 1/0) o l'opció de “No aplica”. En tots els casos, les preguntes estaran referenciades a diferents disposicions legals com la AI Act i/o a recomanacions ètiques d'organismes supranacionals com la OECD o UNESCO.

Amb la informació obtinguda de l'autoavaluació s'obtindran mètriques per cada un dels 7 principis avaluats i amb elles perfils observables de cada una de les organitzacions que portin a terme l'autoavaluació (Fase 3). Un cop obtingudes les mètriques corresponents, es procedirà a l'elaboració d'un segell de qualitat que serà la imatge gràfica del model PIO representant l'adequació de les dades i del sistema d'IA que s'ha autoavaluat segons els 7 principis ètics. Per realitzar aquest procés es seguirà utilitzant un gradient de color d'acord que representarà



els resultats obtinguts en el qüestionari d'autoavaluació (Fase 4).

El model PIO “Obligacions i Drets” de l’OEIAC permetrà realitzar una autoavaluació completa amb recomanacions segons les disposicions legals de la AI Act i/o formulacions ètiques d’organismes supranacionals com la OECD o UNESCO. Això permetrà a l’equip d’OEIAC, juntament amb les persones col·laboradores externes de l’àmbit legal, analitzar les respostes i poder proporcionar feedback a l’organització que l’hagi realitzat si així ho desitja. Cal tenir en compte que, a partir dels Principis, Indicadors i Observables, es pot proporcionar una descripció de l’orientació i adequació que pren un sistema d’IA no sols per a les persones que el desenvolupen, sinó també per als seus usuaris. D’aquesta manera, a més d’obtenir unes mètriques específiques i un segell de qualitat, les organitzacions que hagin realitzat l’autoavaluació completa rebran recomanacions de “compliance” específiques de millora. Aquest aspecte suposa que l’OEIAC abonarà despeses de consultoria a aquelles persones col·laboradores de l’OEIAC per realitzar informes tècnics de recomanacions expertes relacionades amb l’ús i implantació del model PIO en termes legal i d’ètica aplicada.

Al final del procés d’autoavaluació s’obtindrà un segell de qualitat de tipus “compliance” a través de la nova pàgina web, sent el segell una imatge gràfica que es basa en les mètriques obtingudes sobre l’orientació i adequació del sistema d’IA als diferents principis ètics d’acord amb el model PIO. En aquest sentit, en el seu interior se situen els diferents principis amb un gradient de color d’acord amb les mètriques obtingudes (veure Figura 3: Com funciona II).

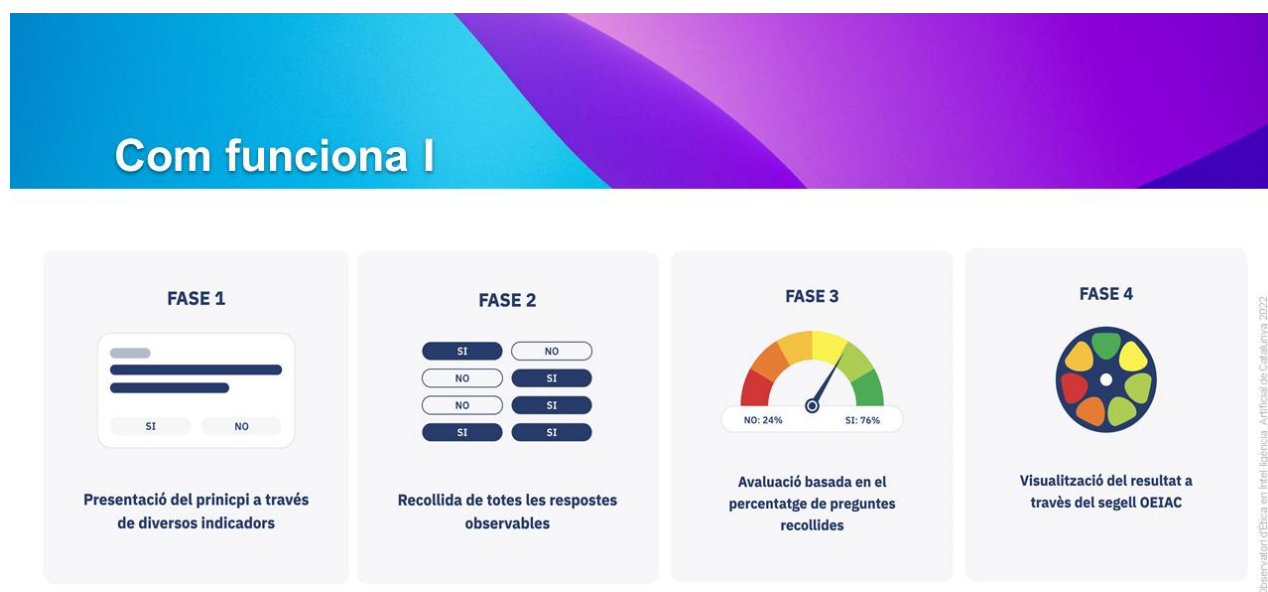


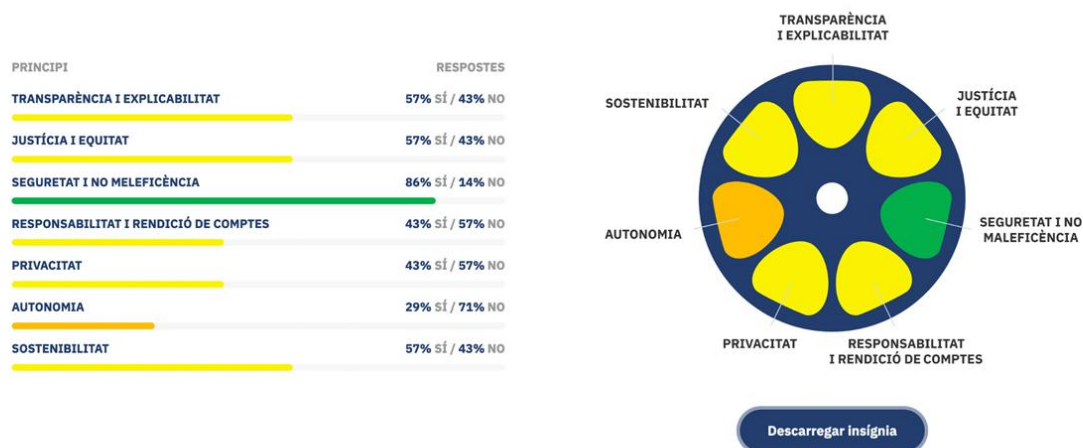
Figura 2

Així doncs, a manera de resum aquestes són les fases que inclou cada model d'autoavaluació:

- Autoavaluació ràpida: FASE 1+ FASE 2 + FASE 3 + FASE 4
- Autoavaluació completa: FASE 1+ FASE 2 + FASE 3 + FASE 4 + Feedback OEIAC



Com funciona II



Observatori d'Ètica en Intel·ligència Artificial de Catalunya 2022

Figura 3

c) El model PIO “Salut”

L'OEIAC està elaborant noves eines de verificació i protocols sota el títol: “*Model PIO Salut: Un model d'autoavaluació sobre l'ús ètic de dades i sistemes d'intel·ligència artificial en l'àmbit de la Salut*”.

L'objectiu principal del model PIO Salut és facilitar l'avaluació i el desplegament dels sistemes d'intel·ligència artificial en l'àmbit de la salut a través d'un model PIO centrat en Salut (PIO-Salut), i destinat als sistemes que s'utilitzen a centres mèdics: hospitals, CAPs i CUAPs, i farmàcies. Per a la primera versió del model específic del model PIO Salut, es consideren dos grups de principis a avaluar: Transparència i Explicabilitat, i Retiment de Comptes i Responsabilitat. Després de portar a terme reunions amb persones expertes, considerem aquests dos conjunts de principis, inclosos dins del model PIO general, com d'especial rellevància, ja que presenten particularitats especials que s'han de tenir en compte per assegurar un desenvolupament segur i fiable de sistemes d'intel·ligència artificial en l'àmbit de la salut.

Prenent el model PIO actual de l'OEIAC com a referència, hem portat a terme una recerca exhaustiva de les solucions de verificació de dades i sistemes actualment disponibles, així com de documents legals sobre la temàtica. També s'ha realitzat una investigació de l'existència d'aquests tipus de models en altres àrees geogràfiques que complementin la informació extreta del territori català, espanyol i europeu. L'objectiu és configurar una eina d'avaluació que, al marge d'abordar els principis relacionats amb la transparència i l'explicabilitat i el retiment de comptes i la responsabilitat, també doni resposta al desplegament legal amb l'arribada de normatives com la AI Act en el context europeu. D'aquesta recerca es prevue que s'extraurà la informació rellevant per adaptar el model PIO general al sector salut.



Juntament amb el formulari específic de verificació o checklist, també es desenvolupa un catàleg d'informació legal i de compra a través de clàusules de recomanació que permeti tant als desenvolupadors de sistemes d'intel·ligència artificial en salut com als usuaris d'organitzacions públiques i privades d'aquests sistemes, una posada en pràctica i desplegament eficient del model PIO Salut tenint en compte els canvis normatius. Es preveu que aquest nou model es provarà i validarà amb diferents persones tècniques així com amb entitats del sistema sanitari que actualment estiguin fent ús de sistemes d'intel·ligència artificial.

d) Bones pràctiques, col·laboracions i reunions sobre IA responsable

L'OEIAC està duent diferents actuacions sobre bones pràctiques, col·laboracions i reunions diverses per promocionar la IA responsable. La IA responsable engloba una àmplia gamma de principis i pràctiques destinades a garantir que les tecnologies d'IA es desenvolupin i s'apliquin de manera ètica, justa i beneficiosa per a la societat. En aquest sentit, les bones pràctiques i la col·laboració són pedres angulars de la IA responsable, ja que reuneixen diverses perspectives i experiències per abordar els complexos reptes i els riscos potencials associats a les tecnologies d'IA.

Un aspecte clau de les bones pràctiques en IA responsable és la promoció de la transparència i el retiment de comptes que fomenta l'OEIAC a través del seu model PIO. És per aquesta raó que des de l'OEIAC es promou, a través de reunions, que les organitzacions i els investigadors que treballen en IA s'esforcin al màxim per ser transparents sobre els seus objectius, mètodes i fonts de dades ja que aquesta transparència ajuda a generar confiança amb les parts interessades i permet un escrutini independent dels sistemes d'IA. La col·laboració té un paper crucial en aquest context, ja que facilita l'intercanvi d'informació, bones pràctiques i lliçons apreses a tota la comunitat d'IA. Els esforços col·laboratius, com ara el desenvolupament de marcs i biblioteques d'IA de codi obert, també permeten als investigadors i professionals treballar junts per crear sistemes d'IA més transparents i responsables.

A més, la promoció de les bones pràctiques, col·laboracions i reunions sobre IA responsable tenen l'avantatge de poder aplicar-se als dominis normatius i polítics. De fet, l'OEIAC treballa perquè els governs, els líders del sector i les organitzacions de la societat civil treballin junts per aconseguir sistemes d'IA responsable. Aquests esforços col·laboratius poden ajudar a garantir que les tecnologies d'IA s'utilitzin de manera que s'alineen amb els valors i les normes de la societat alhora que es minimitzen els riscos. L'OEIAC, en participar en diàlegs entre múltiples parts interessades, promou la transparència, el retiment de comptes i l'equitat en els sistemes d'IA, alhora que aborda consideracions ètiques importants com la privadesa i la protecció de dades.

A continuació recullim la participació en diferents esdeveniments com conferències, debats o tallers sobre intel·ligència artificial, ètica i societat per donar a conèixer l'OEIAC i que formen part de la promoció de bones pràctiques:



Data	Títol
01/01/2024	L'OEIAC contribueix al monogràfic de la revista Valors sobre ètica i IA, https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/39974
21/12/2023	L'OEIAC contribueix al reportatge sobre imatges d'abusos sexuals en bases de dades d'IA, https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/39980
14/12/2023	Participem en la investidura honoris causa del professor Ulises Cortés, membre del Consell Assessor de l'OEIAC, https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/39973
05/12/2023	L'OEIAC contribueix al reportatge sobre ChatGPT i la generació de dades falses en assajos clínics, https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/39979
24/10/2023	L'OEIAC representat al Comitè d'Ètica de les Dades de la Generalitat de Catalunya, https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/39960
17/10/2023	L'OEIAC a la Wales Tech Week amb la delegació del Govern de la Generalitat al Regne Unit, https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/39959
29/09/2023	L'OEIAC participa en la Jornada "Els Drets Humans en l'Era de la Intel·ligència Artificial", https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/35948
07/07/2023	L'OEIAC acull la primera Jornada sobre Feminismes i Intel·ligència Artificial, https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/33221
20/06/2023	L'OEIAC representat al Comitè Directiu del Programa Salut/IA de la Generalitat de Catalunya, https://www.barcelona.cat/festa-de-la-ciencia/ca/activitat/impacte-social-de-la-intelligencia-artificial
18/04/2023	L'OEIAC representat al Consell Assessor d'IA, Ètica i Drets Digitals de l'Ajuntament de Barcelona, https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/29082
17/04/2023	L'OEIAC contribueix al reportatge sobre el sistema predictiu de la Seguretat Social, https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/39978
17/04/2023	L'OEIAC contribueix amb un article d'opinió al Butlletí de Vigilància Tecnològica de CIDAI, https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/29114
23/01/2023	L'OEIAC contribueix al reportatge "Inteligencia artificial generativa: ¿inspiración o plagio?", https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/24920
18/01/2023	La professora Virginia Dignum visita l'OEIAC a la UdG, https://www.udg.edu/ca/udg/detall-noticies/eventid/24811

e) Disseny i difusió de nous recursos pedagògics



Els disseny i difusió de recursos pedagògics al voltant dels usos ètics de la IA es considera important ja que ofereixen diversos avantatges clau. En primer lloc, faciliten la difusió del coneixement i el desenvolupament d'habilitats en el camp de ràpida evolució de la intel·ligència artificial. En aquest sentit, proporcionen materials d'aprenentatge estructurats i accessibles, que permeten a persones obtenir una comprensió profunda dels conceptes, algorismes i aplicacions d'IA que són considerades responsables. Aquesta difusió ajuda en la democratització del coneixement i permet avançar en la superació de la bretxa d'habilitats d'IA. En segon lloc, els recursos pedagògics per a la IA tenen un paper fonamental en la promoció de pràctiques d'IA responsables i ètiques. Dit d'altra manera, no només ensenyen els aspectes tècnics de la IA, sinó que també emfatitzen les consideracions ètiques i les implicacions socials de la tecnologia d'IA. Així doncs, els recursos pedagògics permeten una major formació, entre altres, sobre l'equitat, el biaix, la privadesa i la transparència en els sistemes d'IA, fomentant un sentit de responsabilitat entre els professionals de la IA.

Per tant, es continuarà amb el disseny de nous recursos pedagògics relacionats amb 1) els objectius i actuacions de l'OEIAC com a pilar de l'Estratègia CATALONIA.AI, i 2) l'aplicació del model PIO, tenint en compte els 7 principis ètics fonamentals que ens regeixen (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) per a dades i sistemes d'IA. L'objectiu principal del disseny i difusió d'aquests recursos pedagògics a través d'infografies serà apropar l'activitat de l'OEIAC en termes de divulgació i transferència així com la valorització i simplificació dels continguts dels models PIO per a un públic molt més ampli, incloent-hi persones d'organitzacions educatives.

Juntament amb l'elaboració de noves infografies, també es preveu l'edició i publicació d'un llibre de divulgació que contindrà totes les preguntes de cap de setmana (99) així com les respostes, reflexions i publicacions referenciades.

Recursos pedagògics

Els recursos pedagògics són una sèrie de materials que es van generant de manera paulatina en funció d'aspectes com la irrupció d'una tecnologia d'intel·ligència artificial específica (per exemple, la generativa) i el seu possible impacte a qüestions relacionades amb l'aplicació de tecnologies d'intel·ligència artificial en entorns més o menys sensibles (per exemple, el d'educació) o bé les dues coses. Tots els recursos educatius estan disponibles a la pàgina web de l'OEIAC i són descarregables a través de fitxers PDF o bé es poden consultar en línia a través d'un enllaç públic. L'objectiu de tots els recursos pedagògics és el de facilitar l'aprenentatge sobre els usos ètics de la intel·ligència artificial i l'adquisició de competències bàsiques i/o específiques en aquest àmbit. Aquest any s'ha fet una reorganització dels recursos i també s'inclouen una sèrie d'infografies de guia i resum sobre noves eines d'intel·ligència artificial, especialment generatives, així com de normatives, particularment del context europeu.

Títol	Miniatura
-------	-----------



FAQs i recomanacions sobre generadors de text amb IA a l'aula
https://www.udg.edu/ca/Portals/57/OContent_Docs/Infografia_OEIAC_IAGen_aula.pdf

FAQ'S

GENERADORS DE TEXT AMB IA A L'AULA

FAQ

1 PUC UTILITZAR GENERADORS DE TEXT AMB IA COM CHATGPT O SIMILAR?

SÍ, ÉS CLAR.
PERÒ ABANS DE FER-HO HEU DE SABER QUE HI HA LA POSSIBILITAT DE QUE EN FEU UN MAL ÚS.
I, PER TANT, CAL TENIR EN COMPTE:

- * **EL RISC PER A LA PROFESSIONALITAT I L'ÈTICA**
- * **L'EXPECTATIVA CLAU DE QUÈ ÉS LA INTEGRITAT ACADÈMICA**

UNA CONSIDERACIÓ CLAU:

LA SEVA INTEGRACIÓ EN EL CURS O ASSIGNATURA NO SIGNIFICA, EN CAP CAS, QUE NO HAGUEU DE REALITZAR ELS VOSTRES TREBALLS. AIXÒ VOL DIR QUE PODEU UTILITZAR ELS GENERADORS DE TEXT COM A PART DE PROCÉS D'APRENTATGE PERÒ, EN CAP CAS, PER SUBSTITUIR LA VOSTRA CAPACITAT D'ESCRIBRE ELS TREBALLS ENCOMANATS

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

Universitat de Girona



Infografia sobre la AI Act

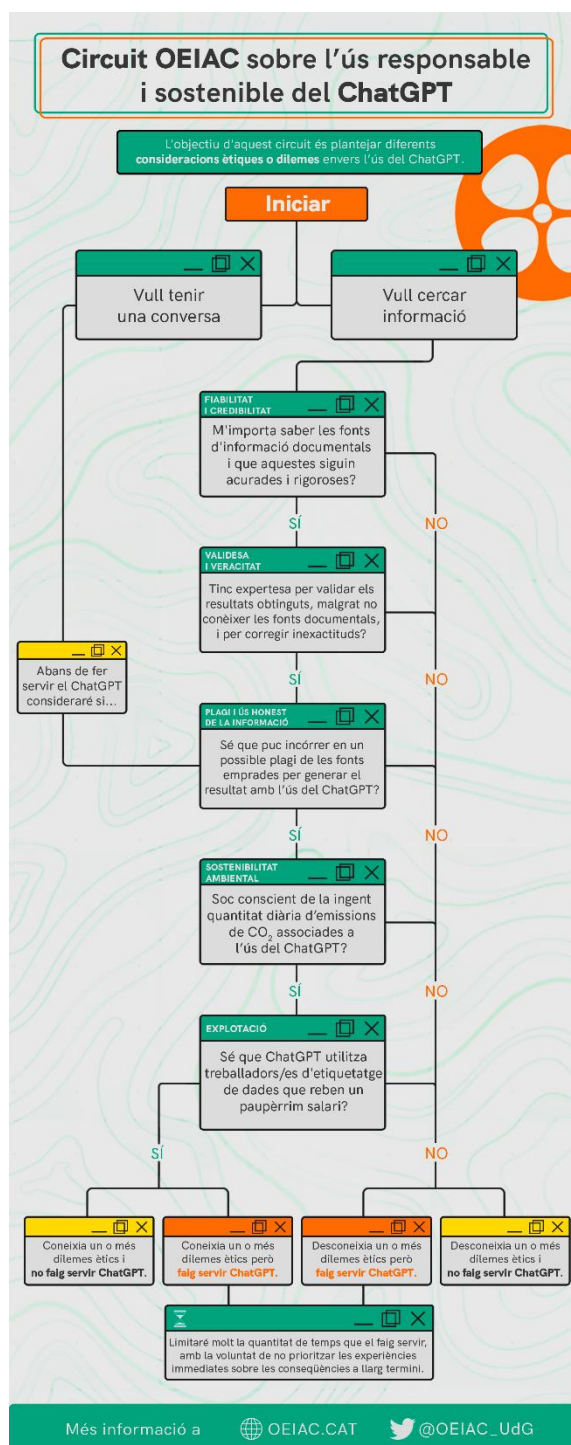
https://www.udg.edu/ca/Portals/57/OContent_Docs/Infografia_OEIAC_AIAct.pdf





Infografia sobre l'ús responsable i sostenible del ChatGPT

https://www.udg.edu/ca/Portals/57/OContent_Docs/Infografia_OEIAC_ChatGPT-1.pdf





Preguntes de cap de setmana

Les preguntes de cap de setmana són enquestes obertes dutes a terme els cap de setmana a través del nostre canal de Twitter sobre debats actuals entorn de la intel·ligència artificial. Cada pregunta es troba disponible en format PDF per a descarregar de manera gratuïta des de la nostra pàgina web. Després d'assolir les 99 preguntes de cap de setmana, s'està elaborant un document en format PDF que les reculli de forma conjunta per poder realitzar una divulgació més enllà de la realitzada a través de Twitter i de la nostra pàgina web.

Títol	Enllaç
Setmana 99 Diverses disciplines de les ciències socials identifiquen la confiança com una característica clau de les societats pròsperes i que funcionen bé. Considereu que la Llei sobre la Intel·ligència Artificial #AIAct contribueix a generar confiança?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_setmana_99-1.pdf
Setmana 98 Les eleccions lliures i la democràcia a Europa poden veure afectades cada cop més per un ús maliciós de la #IA, per produir i difondre desinformació o facilitar la micro-orientació psicogràfica dels votants. Us preocupa aquesta situació?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_setmana_98.pdf
Setmana 97 És entendre la #IA en termes de com és dissenya i funciona un sistema o programari particular que no passa més referència al terme intel·ligència. Esteu d'acord?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_setmana_97.pdf
Setmana 96 Amb la IA algunes empreses "supereficients" creen el potencial de desordre social ja que un excessiu focus en l'eficiència provoca uns alts rendiments per a un nombre cada cop més limitat d'empreses i recompenses més desiguals.	https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_setmana_96.pdf
Setmana 95 Contràriament a la creença popular, els luddites no eren antitecnologia. Més aviat, eren hàbils amb les tecnologies. El seu argument no era amb la tecnologia per se, sinó amb la degradació de les condicions de treball i de vida. Ho sabíeu?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_setmana_95.pdf
Setmana 94 No hi ha conflicte entre el respecte als drets humans i el principi d'innovació. El conflicte és entre el respecte als drets humans i el model de negoci actual de les Big Tech. Esteu d'acord?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_setmana_94.pdf
Setmana 93 Mentre els humans es preocupen pel futur, els sistemes d'IA són 'preocupen' pel passat (les nostres	https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_setmana_93.pdf



dades). Considereu que aquesta situació paradigmàtica ens porta més problemes que solucions?

Setmana 92

Considereu que totes les empreses que desenvolupen els sistemes d'IA generativa haurien de publicar un resum detallat del material amb drets d'autor utilitzat per entrenar models d'IA, com el que hi ha darrere del ChatGPT o midjourney?

https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_setmana_92.pdf

Setmana 91

Considereu que en la nova era d'IA generativa els drets d'autor o les nocions d'originalitat poden acabar sent prescindibles i tenir conseqüències sobre el desempoderament del treball?

https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_setmana_91.pdf

Setmana 90

De quina manera considera el fet que la IA acceleri la influència humana sobre el planeta, particularment en relació al canvi climàtic?

https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_setmana_90.pdf

Setmana 89

Hi ha dos aspectes que defineixen la bona vida per a un ésser humà, des de l'antiguitat i també en l'era de la IA. Quins són?

https://www.udg.edu/ca/Portals/57/OContent_Docs/Pregunta_de_cap_setmana_89.pdf

Setmana 88

La proposta de llei sobre la IA a la UE es complica amb l'arribada del #ChatGPT #BardAi i d'altres IAs generatives. Creieu que aquests s'inclouran en la categoria d'alt risc o en la d'exempció "General Purpose Artificial Intelligence" (GPAI)?

https://www.udg.edu/ca/Portals/57/OContent_Docs/Pregunta_de_cap_de_setmana_88_compressed.pdf

Setmana 87

La societat moderna no ha tingut èxit en crear noves religions més enllà dels cultes de la personalitat. Però a mesura que evoluciona la IA, també apareix un nou tipus de fe generalitzada: les religions basades en la tecnologia. Esteu d'acord?

https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_de_setmana_87.pdf

Setmana 86

La IA generativa ha millorat dràsticament i ara poden produir resultats de text tan realistes i convincts que són difícils de distingir del contingut escrit per humans. Això ofereix una possibilitat major de desinformació i manipulació?

https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_de_setmana_86.pdf

Setmana 85

Amb l'arribada de la IA generativa el "prompt engineering" s'ha convertit en la nova competència EdTech a adquirir. Considereu que el professorat i els alumnes tenen prou indicacions sobre els seus abusos i mals usos (per ex. contingut maliciós)?

https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_de_setmana_85.pdf

Setmana 84

https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_de_setmana_84.pdf



L'interès creixent pels chatbots com #ChatGPT o #DAN (aquest últim sentit moderació de continguts) no radica en la tecnologia sinó amb el nostre profund disseny d'interacció social en una era de soledat i aïllament creixents. Hola, esteu d'acord?

[nt_Docs/Preguntes_de_cap_de_setmana_8_4.pdf](#)

Setmana 83

Fa 50 anys el debat sobre l'absència de cotxes a les ciutats i el foment del transport públic era rebutjat constantment a favor d'una tecnologia en detriment de les persones. És un debat similar amb la IA quan parlem de sostenibilitat?

https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_de_setmana_8_3.pdf

Setmana 82

La "rat race" actual de la #IA funciona al contrari del cercle virtuós ja que tendeix: 1) a més errors sociotècnics (tipus #GoogleBard), 2) en què veiem una IA menys fiable, i 3) en què disminueixi la nostra confiança en la IA. Esteu d'acord?

https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_de_setmana_8_2.pdf

Setmana 81

Una #IA pot oferir solucions inicialment prometedors a curt terminis, però el seu alliberament abans que s'hagin revisat adequadament els seus riscos poden tenir implicacions socials i legals molt àmplies, també a llarg termini. Esteu d'acord?

https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_de_setmana_8_1.pdf

Setmana 80

La indústria tecnològica ha segrestat una varietat de béns comuns i després ens lloga l'accés al que hauria de ser obert per a tothom. Amb l'acceleració de la #IA generativa aquest és un fenomen que veurem accentuat encara més. Esteu d'acord?

https://www.udg.edu/ca/Portals/57/OContent_Docs/Preguntes_de_cap_de_setmana_8_0.pdf

Definicions de cap de setmana

Semblant a les preguntes de cap de setmana, les definicions són explicacions breus sobre termes i conceptes portats a terme el cap de setmana a través del nostre canal de Twitter al voltant de debats actuals entorn de la intel·ligència artificial. Aquestes es van iniciar després d'assolir les 99 preguntes de cap de setmana. Cada definició es troba disponible en format PDF per a descarregar de manera gratuïta des de la nostra pàgina web.

Títol	Enllaç
Definició 21 Sabeu què és el "woke AI"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicio21.pdf
Definició 20 Sabeu què és la "AI Art"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/1ºbloc - Definicions de cap de setmana-2.pdf
Definició 19	https://www.udg.edu/ca/Portals/57/OContent_Docs/1ºbloc - Definicions de cap de setmana-2.pdf



Sabeu què és la diversitat en la IA?	nt_Docs/definicio19.pdf
Definició 18 Sabeu què són els sistemes de puntuació social basats en IA?	https://www.udg.edu/ca/Portals/57/OContent_Docs/definicio18.pdf
Definició 17 Sabeu què és el discurs de la IA?	https://www.udg.edu/ca/Portals/57/OContent_Docs/definicio17.pdf
Definició 16 Sabeu què són els "AI Clones"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/1ºbloc_-_Definicions_de_cap_de_setmana_1_.pdf
Definició 15 Sabeu que és el model "Pay or OK"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/1ºbloc_-_Definicions_de_cap_de_setmana-1.pdf
Definició 14 Sabeu què vol dir "dataficació" i "dataisme"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicio_de_cap_de_setmana_14_.pdf
Definició 13 Sabeu què és una "AI ethics shopping"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicio_cap_de_setmana_13.pdf
Definició 12 Sabeu què és un "AI recall"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicions_de_cap_de_setmana_12.pdf
Definició 11 Sabeu que significa la "AI compliance"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicions_de_cap_de_setmana_11.pdf
Definició 10 Què és l'economia de vigilància?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicions_de_cap_de_setmana_10.pdf
Definició 9 Sabeu què és la "llibertat cognitiva"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicions_de_cap_de_setmana_9.pdf
Definició 8 Què vol dir "jailbreaking"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicions_de_cap_de_setmana_8.pdf
Definició 7 Què és la "manipulació per sistemes d'IA"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicions_de_cap_de_setmana_7.pdf
Definició 6 Què és el "microtargeting"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicions_de_cap_de_setmana_6.pdf
Definició 5 Què és una "IA socialment indiferent"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicions_de_cap_de_setmana_5.pdf



Definició 4 Què significa "Open AI"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicions_de_cap_de_setmana_4-1.pdf
Definició 3 Sabeu la diferència entre un "Shallowfake" i un "Deepfake" i els tipus més comuns?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicions_de_cap_de_setmana_3.pdf
Definició 2 Sabeu que és el "AI Risk"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicions_de_cap_de_setmana_2.pdf
Definició 1 Sabeu que és un "data broker"?	https://www.udg.edu/ca/Portals/57/OContent_Docs/Definicions_de_cap_de_setmana_1-1.pdf

Publicacions en el Blog OEIAC

S'han publicat dos articles d'interès general relacionats amb les xarxes socials i els mòbils així com amb el famós Chat GPT i les seves derivades ètiques i socials. Aquests també son consultables a través de la nostra pàgina web.

- ***Xarxes socials, violència simbólica i adolescència***
Albert Sabater, director de l'Observatori d'Ètica en Intel·ligència Artificial de Catalunya
Data de publicació: 02/03/2023
<https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/26410>
- ***Chat GPT Chat GPT Bang Bang***
Albert Sabater, director de l'Observatori d'Ètica en Intel·ligència Artificial de Catalunya
Data de publicació: 12/02/2023
<https://www.udg.edu/ca/catedres/oeiac/detall-noticies/eventid/26146>



3. Resum d'actuacions i cronograma

a) Nou cicle de Seminaris OEIAC

- a.1. Seminaris presencials del cicle OEIAC.
- a.2. Conferència d'obertura del cicle OEIAC.
- a.3. Acte de cloenda del cicle OEIAC.

b) El model PIO "Obligacions i Drets"

- b.1. Concepció del nou model PIO.
- b.2. Consulta i creació de capacitat del nou model PIO.
- b.3. Disseny del nou model PIO.

c) El model PIO "Salut"

- c.1. Concepció del nou model PIO.
- c.2. Consulta i creació de capacitat del nou model PIO.
- c.3. Disseny del nou model PIO.

d) Bones pràctiques, col·laboracions i reunions sobre IA responsable

- d.1. Foment de reunions i trobades d'IA responsable.
- d.2. Estadets de recerca i col·laboracions nacionals i internacionals.

e) Disseny i difusió de nous recursos pedagògics

- e.1. Confecció i impressió d'infografies pedagògiques OEIAC.
- e.2. Edició i publicació d'un llibre divulgatiu sobre les preguntes de cap de setmana.

Cronograma

Activitat	Mes											
	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												
c.3												
d.1												
d.2												
e.1												
e.2												

(En ombrejat accions continuades)



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 Vilarinho**
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

- **Alba Canal**
MACBA
- **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
- **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 Associada de la Facultat de Turisme, Universitat de Girona
- **Núria Vallès**
Barcelona Science and Technology Studies Group (STS-b), Universitat Autònoma de Barcelona
- **Patricia 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
- **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