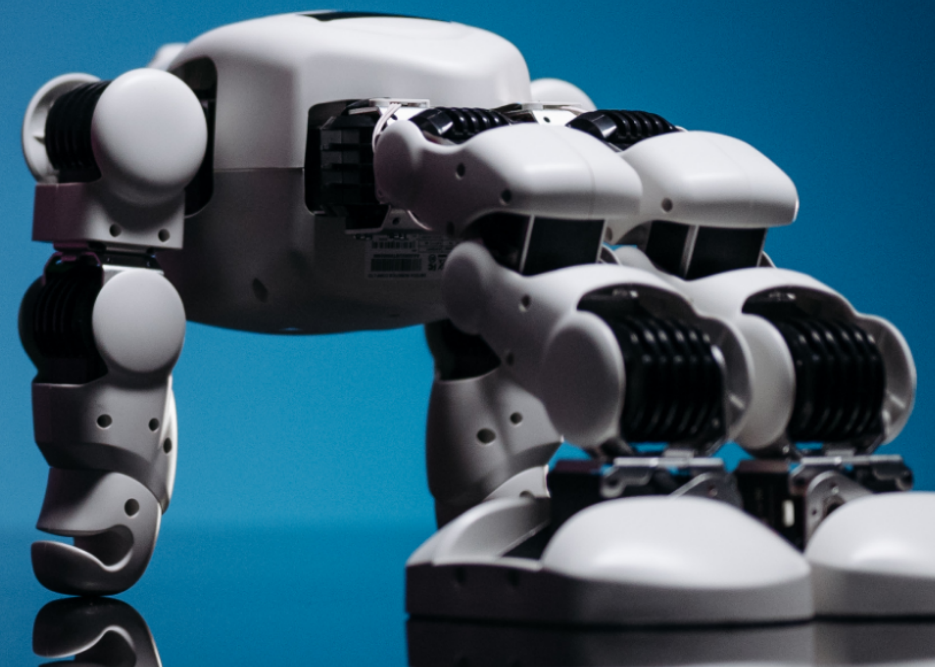


Mònica Navarro-Michel, Oriol Pujol, Carlos Lopezosa, Jordi Vitrià

Generació i Difusió de Mitjans Sintètics en els Mitjans de Comunicació Audiovisuals Públics



UNIVERSITAT DE
BARCELONA

Càtedra UB en Intel·ligència
Artificial i Mitjans

Autors:

Mònica Navarro-Michel (m.navarro@ub.edu), Facultat de Dret.
Oriol Pujol (oriol_pujol@ub.edu), Facultat de Matemàtiques i Informàtica.
Carlos Lopezosa (lopezosa@ub.edu), Facultat d'Informació i Mitjans Audiovisuals.
Jordi Vitrià (jordi.vitria@ub.edu), Facultat de Matemàtiques i Informàtica.

Data:

19.02.2025

Finançat per la Càtedra UB en Intel·ligència Artificial i Mitjans & 3Cat



Objectiu del document

Fer una proposta de directrius clares per a l'ús responsable d'imatges, vídeo i àudio sintètics a un mitjà de comunicació audiovisual públic, assegurant la transparència, l'ètica i la protecció dels drets dels individus. Aquestes directrius han de minimitzar els riscos associats a la intel·ligència artificial i, alhora, maximitzar-ne els beneficis.

Context

Els avanços en intel·ligència artificial han facilitat la generació d'imatges, àudio i vídeos sintètics amb un alt nivell de realisme. Aquest progrés obre noves oportunitats creatives, però també comporta riscos significatius, especialment en l'àmbit de la informació i el periodisme. En aquest context, els mitjans públics, com a institucions compromeses amb els valors democràtics, han de garantir un ús ètic i transparent d'aquestes tecnologies en els diversos escenaris d'aplicació, com ara els serveis informatius, els programes d'entreteniment i les produccions de ficció.

En l'àmbit de l'entreteniment i la ficció, els mitjans sintètics generats per intel·ligència artificial ofereixen noves possibilitats creatives. Aquestes tecnologies permeten la creació de personatges digitals, la recreació d'escenaris històrics o fantàstics, així com la generació d'efectes visuals i sons innovadors que poden enriquir la narrativa audiovisual i millorar l'experiència de l'audiència.

D'altra banda, en l'àmbit de la informació, l'impacte de les tecnologies sintètiques resulta especialment crític. Aquestes eines poden ser útils per il·lustrar conceptes complexos mitjançant recreacions visuals, generar mapes interactius o crear animacions explicatives que facilitin la comprensió de les notícies. No obstant això, també presenten el risc de desinformació i manipulació, especialment si són susceptibles de ser usades per tercers. Per aquest motiu, és imprescindible establir protocols estrictes que assegurin la verificabilitat del contingut generat per intel·ligència artificial i la seva identificació explícita com a material sintètic.

Context

En aquest escenari, la prioritat fonamental d'un mitjà de comunicació audiovisual públic ha de ser preservar la confiança del públic i garantir que les informacions es presentin amb el màxim rigor, imparcialitat i fidelitat als fets.

L'ús d'imatges i àudio sintètics en un mitjà públic pot implicar dues responsabilitats diferenciades segons el Reglament Europeu d'Intel·ligència Artificial (RIA): la de productor/proveïdor i la de desplegador. Aquest doble rol determina les obligacions ètiques i legals que l'organització ha d'assumir per garantir un ús responsable de la intel·ligència artificial en els seus continguts audiovisuals.

Com a productor/proveïdor, la televisió pública genera continguts que poden incorporar elements sintètics creats amb IA, com ara recreacions històriques, personatges digitals o doblatges automàtics. En aquest sentit, actua com a proveïdor d'un sistema d'IA segons el RIA, fet que implica responsabilitats específiques:

- **Transparència i etiquetatge:** El contingut generat amb IA ha d'estar clarament identificat com a sintètic per evitar confusions i desinformació.
- **Mitigació de riscos:** Cal establir mecanismes per minimitzar biaixos, errors o manipulacions potencials en els materials generats.
- **Garantia de qualitat i fiabilitat:** Els continguts han de respectar els estàndards ètics i professionals del servei públic audiovisual.

Context

Com a desplegador, la televisió pública utilitza eines d'IA desenvolupades per tercers per a la creació, edició o distribució de continguts. Això comporta responsabilitats associades a l'ús d'aquests sistemes:

- Supervisió humana: La IA no ha de substituir la presa de decisions editorial, sinó complementar-la sota la supervisió de professionals humans.
- Avaluació d'impacte: L'ús d'IA en processos informatius o de creació de continguts ha de ser revisat per assegurar-ne la transparència i el respecte pels drets fonamentals.
- Compliment del RIA: Segons el reglament europeu, les aplicacions d'IA en mitjans de comunicació poden ser considerades d'alt risc en funció del seu impacte en la desinformació o la manipulació de l'opinió pública. Això exigeix mecanismes de control, auditoria i transparència en el seu desplegament.

Aquest doble enfocament obliga les televisions públiques a definir directrius clares per a l'ús d'imatges i àudio sintètics. L'adopció d'aquestes directrius no només respon a un imperatiu legal segons el RIA, sinó també a la necessitat de mantenir la confiança de l'audiència i reforçar el paper de la televisió pública com a referent en informació veraç i continguts de qualitat.



Riscos Identificats

Desinformació

- L'ús d'imatges, vídeos o àudios generats per IA pot confondre l'audiència, fent que percebi com a reals continguts que són totalment o parcialment sintètics.
- Si no es garanteix una identificació clara del material generat artificialment, es pot erosionar la confiança del públic en la veracitat de la informació proporcionada per la televisió pública.
- En l'àmbit de les notícies i els programes informatius, un ús inadequat d'aquestes tecnologies pot contribuir a la difusió de notícies falses o a la manipulació de la percepció pública sobre determinats esdeveniments.

Manipulació i ús malèvol

- La modificació d'imatges o àudios de persones pot donar lloc a la creació de deepfakes, que poden ser utilitzats per part de tercers per a la difamació, la falsificació de declaracions o la manipulació de l'opinió pública.
- La possibilitat d'alterar declaracions o generar àudios sintètics de figures públiques sense el seu consentiment pot derivar en greus problemes ètics i legals.
- Un ús irresponsable d'aquestes tecnologies pot afavorir la creació de narratives enganyoses que perjudiquin individus o institucions.

Riscos Identificats

Impacte ètic i vulneració de drets

- L'explotació de dades biomètriques per generar continguts sintètics pot vulnerar el dret a la privacitat i la protecció de dades de les persones.
- La creació de personatges sintètics basats en trets biomètrics reals pot suposar un ús inadequat de la imatge d'individus sense el seu consentiment explícit.
- L'automatització de veus sintètiques pot generar preocupacions sobre la pèrdua d'autenticitat en la comunicació audiovisual i la deshumanització de la informació.

Pèrdua de llocs de treball

- L'ús creixent de models d'IA per generar contingut audiovisual pot reduir la necessitat de professionals humans en àrees com el doblatge, l'edició de vídeo, la il·lustració digital o la producció de continguts informatius.
- La substitució de locutors per veus sintètiques o d'actors per avatars generats per IA pot afectar el sector audiovisual i els seus professionals.
- Un ús indiscriminat d'aquestes tecnologies pot afavorir la precarització laboral i reduir la diversitat creativa en els mitjans de comunicació.

Riscos Identificats

Propietat intel·lectual i drets d'autor

- La generació de continguts amb IA planteja qüestions sobre la titularitat dels drets d'autor: qui és el propietari d'una obra creada per una IA?
- L'ús de material preexistent per entrenar models d'IA sense el consentiment dels autors pot vulnerar els drets de propietat intel·lectual.
- La reutilització d'imatges, àudios o estils artístics sense permís pot derivar en litigis per infracció de drets d'autor.

Risc reputacional

- Si l'audiència percep que una televisió pública utilitza IA d'una manera opaca o enganyosa, la seva credibilitat es pot veure greument afectada.
- Un ús inadequat d'imatges o àudios sintètics en programes informatius pot generar escàndols mediàtics i crisi de confiança.
- La manca de polítiques clares de transparència sobre l'ús d'IA en els continguts audiovisuals pot derivar en una pèrdua d'audiència i una disminució de la legitimitat de la televisió pública.

Riscos Identificats

Riscos legals i regulació

- L'ús de tecnologies sintètiques ha d'adaptar-se a la legislació vigent, especialment al Reglament Europeu d'Intel·ligència Artificial (RIA), que classifica els sistemes d'IA en funció del seu grau de risc: inacceptable, alt, limitat i mínim.
- Si una televisió pública no implementa mecanismes de verificació i supervisió, pot incórrer en responsabilitats legals relacionades amb la difusió de desinformació o la vulneració de drets.
- En cas d'ús inadequat d'àudios o imatges generades per IA, poden sorgir reclamacions per vulneració del dret a la intimitat, a l'honor, per l'ús indegut de la imatge o incompliment de normatives sobre protecció de dades.

Llista de possibles danys derivats de l'ús de mitjans sintètics que busquem mitigar en una televisió pública:

Suplantació d'identitat

Reproduir de manera enganyosa la veu o la imatge d'individus públics o privats sense consentiment.

Crear deepfakes que puguin desinformar l'audiència.

Desinformació i manipulació informativa

Generar continguts sintètics que falsegin esdeveniments, declaracions o dades.

Utilitzar IA per simular fonts periodístiques sense verificació humana.

Manipulació política i electoral

Crear continguts que indueixin a error els votants sobre processos electorals.

Difondre imatges, vídeos o declaracions falses que perjudiquin o beneficiïn actors polítics.

Fraus i enganys comercials

Usar IA per generar testimonis falsos en anuncis o continguts patrocinats.

Manipular informació sobre productes o serveis de manera no ètica.

Discursos d'odi i discriminació

Crear continguts que puguin ser usats per fomentar la discriminació per motius de raça, gènere, religió o altres condicions.

Manipulació de menors

Crear continguts potencialment enganyosos destinats a infants sense supervisió.

Manipulació de l'opinió pública

Utilitzar IA per fabricar comentaris o interaccions en plataformes digitals.

Pèrdua de confiança en els mitjans públics

Aplicar IA sense transparència en continguts informatius o de servei públic.

No advertir clarament quan un contingut ha estat generat o modificat per IA.

Extorsió i difamació

Difondre continguts sintètics que poden ser usats per coaccionar persones o institucions.

Crear deepfakes que puguin perjudicar la reputació de persones públiques o privades.

Ús inadequat en continguts informatius i culturals

Substituir la feina de periodistes i creadors sense justificació editorial.

Fer servir veus i imatges sintètiques sense indicar-ho clarament a l'audiència.



Principis Rectors

El principi de **transparència** estableix que qualsevol contingut sintètic, ja sigui una imatge, un vídeo o un àudio generat mitjançant IA, ha de ser clarament identificat com a tal. Això significa que s'han d'incloure etiquetes o marques visuals i auditives que permetin a l'audiència reconèixer immediatament la naturalesa artificial del material. La transparència és essencial per evitar la desinformació i garantir que els espectadors puguin diferenciar entre contingut real i contingut generat artificialment.

El **consentiment** és un altre pilar fonamental, especialment quan es tracta de la creació de mitjans sintètics que impliquen la veu, el rostre o altres trets identificatius d'una persona. Per utilitzar aquestes dades, cal obtenir l'autorització prèvia de la persona en qüestió (es considerarà cada cas específic, incloent-hi personatges públics, ciutadans, menors, entre altres). A més, només es poden emprar dades biomètriques recollides de manera legal i ètica, garantint el compliment de les normatives sobre protecció de dades i drets fonamentals. En cas de mort, el consentiment, si no hi ha persona especialment designada pel difunt, s'hauria d'obtenir dels hereus o persones vinculades per raons familiars o de fet.

La **responsabilitat editorial** és un principi clau que recau sobre els professionals dels mitjans públics, especialment els editors i responsables de contingut. Han de garantir que els mitjans sintètics no puguin fer servir amb finalitats manipuladores ni per enganyar l'audiència. Això implica establir mecanismes de control per evitar la difusió de deepfakes, recreacions enganyoses o qualsevol altre ús que pugui alterar la percepció de la realitat per part del públic. L'ús d'aquestes tecnologies ha d'estar orientat a complementar i enriquir la informació, mai a distorsionar-la.

La **responsabilitat social** comporta l'obligació de respectar les normatives de protecció de dades i assegurar que no es vulnerin drets fonamentals com la privacitat, la propietat intel·lectual i la transparència. L'ús d'eines d'IA ha de ser compatible amb Reglament europeu d'IA (RIA), que estableix criteris per a la supervisió i regulació d'aplicacions d'IA en àmbits d'alt impacte, com el periodisme i la comunicació audiovisual. Això significa que les televisions públiques han de fer una avaluació contínua dels riscos associats a aquestes tecnologies, garantint-ne un ús ètic i alineat amb els valors del servei públic.



weather-app > src > JS App.js > App > render

```
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
  
getWeather = async () => {  
  const apiCall = fetch(`${apiBase}q=${this.state.userinputCity}`)  
  .then(res => res.json())  
  .then(res => {  
    console.log(res);  
    let city = res.name;  
    let country = res.sys.country;  
    let weatherDescription = res.weather[0].main;  
    let currentTemp = res.main.temp;  
    let maxTemp = res.main.temp_max;  
    let minTemp = res.main.temp_min;  
  
    console.log(weatherDescription);  
    console.log(currentTemp);  
    console.log(maxTemp);  
    console.log(minTemp);  
  
    this.setState({  
      city: city,  
      country: country,  
      weatherDescription: weatherDescription,  
      currentTemp: currentTemp,  
      maxTemp: maxTemp,  
      minTemp: minTemp  
    })  
  })  
}
```

Directius Específiques

Avisos a l'audiència

La transparència és essencial per mantenir la confiança del públic en els mitjans de comunicació, especialment en una televisió pública. Per aquest motiu, sempre que es faci ús de contingut sintètic, s'ha d'incloure una etiqueta visual o auditiva que indiqui clarament que el material ha estat generat mitjançant intel·ligència artificial. Aquest avís ha de ser visible i comprensible per a tota l'audiència, evitant confusions sobre l'autenticitat del contingut.

A més, en programes informatius i reportatges, cal explicar el propòsit i el mètode utilitzat per generar aquests materials. Això pot incloure una descripció durant l'emissió o en els crèdits del programa, detallant com s'han creat les imatges o els àudios sintètics i quin ha estat el seu paper en la narrativa. Aquest tipus d'explicacions són especialment importants en casos en què els mitjans sintètics s'utilitzen per il·lustrar escenaris històrics, recreacions hipotètiques o elements didàctics.

Directius Específiques

Control de qualitat

Per prevenir la desinformació i garantir que el material sintètic es faci servir de manera ètica, cal implementar mecanismes de verificació interna que assegurin que el contingut generat per IA no sigui enganyós ni es pugui utilitzar amb finalitats manipuladores. Aquests mecanismes poden incloure l'anàlisi prèvia per part d'un equip editorial especialitzat, la revisió per part d'experts en ètica periodística i la supervisió de continguts per professionals independents quan es tracti de temes d'alt impacte.

També és fonamental dur a terme revisions ètiques abans de l'emissió per avaluar si l'ús de mitjans sintètics és adequat en el context en què es vol aplicar. Aquestes revisions han de valorar la possibilitat d'interpretacions errònies per part del públic, així com l'impacte potencial en la percepció dels fets representats. L'objectiu és assegurar que el contingut sintètic contribueixi a la informació de manera clara i honesta, sense generar ambigüitats o dubtes sobre la seva autenticitat.

Directius Específiques

Prohibicions

L'ús d'intel·ligència artificial en la generació de continguts audiovisuals ha d'estar subjecte a restriccions estrictes per evitar abusos i possibles impactes negatius. Una de les prohibicions fonamentals ha de ser la imitació o suplantació de persones en declaracions que no hagin fet. L'ús d'àudio sintètic per simular la veu d'algú o la creació de deepfakes per posar paraules en boca de figures públiques sense el seu consentiment constitueixen pràctiques inacceptables que poden comprometre greument la credibilitat de la televisió pública.

De la mateixa manera, cal evitar l'ús de contingut sintètic en notícies que tractin temes sensibles o controvertits, llevat que hi hagi una justificació rigorosa per fer-ho. En casos en què es consideri necessari utilitzar aquests mitjans per il·lustrar fets complexos, la televisió ha de proporcionar un context clar, assegurant-se que el públic entén la naturalesa del material mostrat.

Directius Específiques

Formació

La integració de tecnologies sintètiques en els processos de producció audiovisual requereix que els professionals de la televisió pública estiguin formats en el seu ús ètic i responsable. Per això, és imprescindible capacitar els equips de treball en aspectes tècnics i deontològics relacionats amb la creació i aplicació de continguts generats per IA. Aquesta formació ha d'abastar tant el personal editorial com els responsables de producció i tecnologia, garantint que tothom tingui els coneixements necessaris per gestionar adequadament aquestes eines.

A més, és fonamental fomentar la comprensió de l'impacte social i ètic d'aquestes tecnologies, posant èmfasi en els riscos de manipulació, els drets de privacitat i la protecció de dades. Els mitjans públics han d'assumir un paper actiu en la reflexió sobre els efectes de la IA en el periodisme i la comunicació, promovent una cultura de responsabilitat i transparència en l'ús d'aquestes eines.

Aquest conjunt de directrius ha de servir com a marc normatiu per a la implementació de mitjans sintètics en els mitjans públics, assegurant que el seu ús es faci sempre amb criteris d'ètica, rigor i respecte pels drets de la ciutadania.



Implementació

Comissió d'IA Responsable

Un element central en la implementació d'aquestes directrius és la creació d'una comissió especialitzada que pugui avaluar l'ús d'imatges i àudio sintètics en els continguts audiovisuals. Aquesta comissió ha d'estar formada per experts en intel·ligència artificial, periodisme, ètica i dret, així com per representants de la societat civil.

Les funcions principals d'aquesta comissió inclouen:

- Supervisar l'aplicació de les directrius i garantir-ne el compliment.
- Avaluar casos concrets on l'ús de mitjans sintètics pugui generar dilemes ètics o conflictes amb la normativa vigent.
- Emetre recomanacions per millorar les pràctiques internes i adaptar-les a nous desafiaments.
- Establir un protocol per a la revisió de continguts sensibles abans de la seva emissió. Aquest protocol podria emprar-se com un instrument d'avaluació heurístic simple i dicotòmic (sí/no = 1/0), a mode de test, per a determinar la viabilitat d'ús en funció d'un percentatge mínim d'acceptació.
- Publicar informes periòdics sobre, entre d'altres, l'ús de mitjans sintètics en la televisió pública, amb l'objectiu d'informar la ciutadania i reforçar la confiança en el mitjà. Desenvolupar un informe d'auditoria anual on s'identifiquin totes les accions de l'ús de IA duts a terme durant l'any, amb un enllaç que porti, d'una banda a la peça sintètica final i per un altre, al document de control ètic. Aquest informe estaria pensat, no sols per a la societat, sinó també per als treballadors de l'ens públic.

Implementació

Actualització Constant de les Directrius

Les tecnologies d'intel·ligència artificial evolucionen ràpidament, i amb elles, els riscos i les oportunitats associades al seu ús en els mitjans de comunicació. Per aquest motiu, és fonamental establir mecanismes de revisió i actualització constant de les directrius.

Aquesta revisió ha de realitzar-se de manera periòdica, amb una anàlisi de:

- L'evolució tecnològica: noves eines, models i aplicacions d'IA que puguin impactar la creació de continguts audiovisuals.
- L'impacte social: possibles canvis en la percepció pública sobre l'ús de mitjans sintètics i les expectatives en matèria de transparència i ètica.
- L'entorn normatiu: adaptació a noves regulacions, especialment en el marc del Reglament europeu d'IA (RIA) i altres normatives sobre desinformació, privacitat i propietat intel·lectual.

A més, s'ha d'implementar un sistema de seguiment continu que permeti detectar usos problemàtics de la IA en la producció audiovisual i actuar de manera proactiva per corregir-los abans que generin impactes negatius.

Implementació

Redacció d'un manual de bones pràctiques o d'un decàleg

- **Transparència obligatòria**
 - Informem clarament l'audiència quan un contingut ha estat generat o modificat amb mitjans sintètics.
 - Utilitzem etiquetes visibles, notes de context i altres mecanismes de divulgació.
- **Autenticitat i verificació periodística**
 - No utilitzem mitjans sintètics per crear o manipular informació de notícies, reportatges o continguts periodístics sense una verificació clara i contextualitzada.
- **Respecte per la identitat i la dignitat de les persones**
 - Evitem la suplantació d'identitat i l'ús de deepfakes sense el consentiment explícit de les persones implicades.
 - Protegim la privacitat i l'ètica en la representació d'individus en continguts sintètics.
- **Límit a la manipulació de la realitat**
 - No representem persones dient o fent accions que no han fet en realitat, especialment en contextos informatius i de servei públic.
 - Quan recreem escenes amb IA, deixem clar que es tracta d'una representació no autèntica.
- **Ús responsable en l'entreteniment i la ficció**
 - Respectem els drets d'autoria i evitem la perpetuació d'estereotips o biaixos de gènere, ètnia o cultura.
- **Evitar biaixos algorítmics i discriminació**
 - Revisem els models d'IA que usem per evitar biaixos discriminatoris en la representació de gènere, ètnia, edat o altres factors.
- **Protecció dels col·lectius vulnerables**
 - No utilitzem mitjans sintètics per manipular, explotar o desinformar. Cal ser especialment curosos per que la referència als menors, víctimes d'abusos o altres grups en situació de vulnerabilitat.
- **Traçabilitat i rastreig de continguts sintètics**
 - Incorporem mecanismes de verificació com marques d'aigua digitals, metadades rastrejables o tecnologies de verificació d'autenticitat (per exemple, C2PA).
- **Correcció immediata d'errors**
 - Si es detecta un ús inadequat de mitjans sintètics o una divulgació insuficient, ho corregim i ho comuniquem a l'audiència de manera transparent.
- **Compromís amb una televisió pública ètica i confiable**
 - Publiquem i mantenim una política clara sobre l'ús de mitjans sintètics, adaptant-la a les noves

Implementació

Col·laboració Externa per a la Definició d'Estàndards Comuns

El desenvolupament d'un marc ètic per a l'ús d'imatges i àudio sintètics no pot fer-se de manera aïllada. Per garantir coherència i consistència en l'aplicació de bones pràctiques, els mitjans públics han de col·laborar amb altres mitjans, associacions professionals i organismes reguladors.

Aquesta col·laboració pot incloure:

- Participació en iniciatives internacionals per l'establiment d'estàndards compartits en l'ús d'IA en els mitjans de comunicació.
- Intercanvi d'experiències i bones pràctiques amb altres televisions públiques per identificar estratègies efectives de control i transparència.
- Diàleg amb experts acadèmics i organitzacions de drets digitals per anticipar nous reptes ètics i normatius.
- Aliances amb institucions de recerca en IA i periodisme per desenvolupar eines de verificació i detecció de manipulacions sintètiques.

Aquestes accions contribuiran a harmonitzar les polítiques d'ús de mitjans sintètics en l'àmbit audiovisual i a reforçar la credibilitat dels mitjans públics en la seva tasca informativa i educativa.

Implementació

Foment d'una Cultura de Transparència i Col·laboració Interna

A nivell intern, la implementació d'aquestes directrius requereix una estratègia activa per fomentar la participació dels professionals de la televisió pública i crear un entorn de comunicació oberta i col·laboració.

Per facilitar aquest procés, és important establir canals específics de comunicació interna dedicats a reportar inquietuds o suggeriments sobre l'ús de mitjans sintètics. Això pot incloure:

- Correus electrònics dedicats o formularis anònims perquè els treballadors puguin informar sobre possibles riscos o mal ús d'aquestes tecnologies.
- Plataformes internes per compartir bones pràctiques, dubtes i casos d'ús problemàtics relacionats amb IA.
- Sessions periòdiques de debat i reflexió on els equips de producció, edició i direcció puguin discutir dilemes ètics i compartir experiències.
- Memòria anual sobre l'ús de la IA durant l'any, amb enllaços a totes les peces sintètiques finals i al document de control ètic.

A més, és fonamental reconèixer i celebrar les contribucions dels treballadors que identifiquin riscos o proposin solucions innovadores per millorar l'ús ètic de la IA en els continguts audiovisuals. Aquest reconeixement pot materialitzar-se en premis interns, mencions en reunions de direcció o la possibilitat d'implicar aquests professionals en el desenvolupament de noves polítiques i estàndards.



UNIVERSITAT^{DE}
BARCELONA

Càtedra UB en Intel·ligència
Artificial i Mitjans