

3cat

Assistència Intel·ligent a la Descripció Contextualitzada de Fotos per Notícies



CIDAI-PAI-2025-03

Drets reservats. Aquest treball està disponible sota la llicència Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International ([CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)).

Segons els termes d'aquesta llicència, podeu copiar, redistribuir i adaptar l'obra amb fins no comercials, sempre que l'obra sigui citada adequadament, tal com s'indica a continuació.

En qualsevol ús d'aquest treball, no s'ha de suggerir que el CIDAI doni suport a cap organització, producte o servei específic. No es permet l'ús del logotip CIDAI.

Si adapteu l'obra, heu de llicenciar-la amb la mateixa llicència Creative Commons o equivalent.

Si creeu una traducció d'aquest treball, heu d'afegir la següent exempció de responsabilitat juntament amb la cita suggerida: "Aquesta traducció no la va crear el Centre of Innovation for Data tech and Artificial Intelligence (CIDAI). CIDAI no es fa responsable del contingut ni de l'exactitud d'aquesta traducció. L'edició original en català serà l'edició autèntica i vinculant".

Qualsevol mediació relacionada amb disputes derivades de la llicència es durà a terme d'acord amb les normes de mediació de la [World Intellectual Property Organization](https://www.wipo.int/mediation/).

Cita suggerida: CIDAI-PAI-2025-05 // Asistencia intel·ligent a la descripció contextualitzada de fotos per notícies per Rafael Bermúdez – 3Cat, Rafael Redondo Tejedor- Eurecat, CIDAI, 2025.

Llicència: [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Índex

1.	INTRODUCCIÓ	4
2.	DESCRIPCIÓ CONTEXTUALITZADA DE FOTOS PER NOTÍCIES	5
3.	LLIBRE D'ESTIL	9
4.	MODELS GRANS DE LLENGUATGE (LLMS)	12
5.	EXPERIMENTACIÓ	17
6.	GENERACIÓ DE DESCRIPCIÓ CONTEXTUALITZADA: FLUX FINAL	21
7.	INTERFÍCIE D'USUARI	25
8.	AVALUACIÓ FINAL D'USUARIS	31
9.	CONCLUSIONS	46
10.	ANNEXOS	48



Introducció

La descripció de fotos en mitjans d'informació, inclosa els peus de fotos, és sovint una tasca rutinària, lluny de les activitats més intel·lectuals que requereix la redacció d'una notícia o reportatge periodístic. Si bé els peus de foto responen moltes vegades a criteris editorials, la descripció de fotografies en termes purament visuals, embegudes en camps no visibles de les pàgines web, és una informació sovint no considerada per diferents factors en els fluxos de treball actuals. Això fa que persones amb diferents limitacions d'accessibilitat puguin informar-se de la mateixa manera.



En aquest context, el desenvolupament d'eines basades en models d'aprenentatge profund i intel·ligència artificial generativa, que, tot i que no automatitzin completament donin suport al professional a generar text informatiu de qualitat en la descripció de fotos, podria impactar de manera significativa en l'efectivitat de la creació de peus de fotos i la seva descripció associada en termes d'accessibilitat.

Descripció contextualitzada de fotos per notícies



La generació de subtítols d'imatges és el procés de generació de descripcions textuals per a imatges. En el context dels mitjans de comunicació, implica crear subtítols que descriguin de manera precisa i succinta el contingut, el context i la importància de les fotografies que acompanyen les notícies.

Aquesta tasca és crucial perquè els subtítols no només proporcionen als lectors una ràpida comprensió de la imatge, sinó que també milloren la narració de la notícia, afegint profunditat i context. Més enllà, un bon subtítol afegeix un context que connecta la imatge amb la narrativa més àmplia de l'article. Ha d'explicar per què la imatge és important i com es relaciona amb la història.

Els peus de foto han de reflectir amb precisió el contingut de la imatge. Això inclou identificar persones, llocs, objectes, esdeveniments i qualsevol detall rellevant que sigui significatiu per a la notícia. Aquest funció sovint requereix una comprensió de la història més àmplia, més enllà del moment concret capturat a la foto.

Els subtítols han de ser clars i concisos, evitant ambigüitats. Han de comunicar la informació essencial de manera ràpida i eficaç. A més, els subtítols efectius poden atraure els lectors, animant-los a llegir l'article complet. Aquí els models grans de llenguatge natural poden aportar la seva capacitat per reformular text i aportar ràpidament variacions en funció de, per exemple, el to editorial o emocional de la descripció o llargada en nombre de caràcters.

Exemple il·lustratiu de flux de processament per generació de subtítols de fotografies.



Si bé la identificació de, per exemple, persones amb models d'IA és una característica que no escaula amb facilitat, havent d'entrenar models per reconèixer cada persona identificable en particular, és una informació que sí està disponible en el text de la notícia que acompanya la fotografia. Aquesta tasca però, implica nous reptes com la identificació d'elements nominals al text i la seva correspondència visual a la imatge.

Més enllà, quan les fotos inclouen persones, un repte important és capturar l'essència de les seves expressions, emocions i accions. Això implica comprendre els aspectes personals dels individus de la foto i transmetre el seu significat a la història. Per exemple, descriure l'expressió triomfal d'un polític en el seu discurs de victòria en unes eleccions, o la mirada o gest decidits d'una persona que participa en una manifestació.

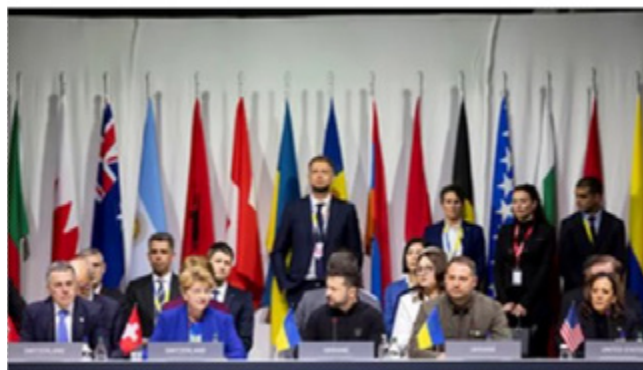
Part d'aquesta descripció implica destacar el que fan els individus i les seves interaccions amb els altres. Això afegeix dinamisme al subtítol i ajuda els lectors a entendre les accions que s'estan esdevenint.

La generació de subtítols d'imatges als mitjans de comunicació és una tasca crítica que requereix una combinació de precisió, context, claredat i implicació. S'enfronta a diversos reptes, especialment en la comprensió contextual i la identificació de detalls, així com a equilibrar el detall amb la brevetat. La descripció de personatges individuals afegeix una capa addicional de complexitat, que requereix una atenció acurada a la identificació, el to emocional, les accions i la importància. Dominar aquests elements garanteix que els subtítols de les imatges millorin la comprensió i el compromís del lector amb la notícia.

2.1. Peu de foto

Té com a objectiu oferir al públic general una explicació contextual que ajudi a entendre per què s'ha seleccionat la imatge en qüestió. Aporta detalls addicionals, com el lloc, les persones, els fets o el significat de la imatge, que són rellevants per al contingut general o el missatge que es vol transmetre.

Exemple de fotografia i peu de foto associat visible al llarg de la notícia.



Zelenski busca suport per aïllar Putin a la cimera per la pau a Ucraïna

2.2. Text alternatiu

L'objectiu del text alternatiu és donar a una persona que utilitza un lector de pantalla una descripció clara dels elements clau de la imatge. Aquesta descripció ha de ser suficient per entendre el contingut o la funció de la imatge dins del context de la notícia. Els elements descrits poden variar segons la rellevància que tinguin en aquest context específic. A nivell més tècnic, aquesta informació està incrustada en la pàgina web i habitualment no és directament visible.

<figure>

<figcaption>Zelenski busca suport per aïllar Putin a la cimera per la pau a Ucraïna.</figcaption>

</figure>

2.3. Fonts informació

Les fonts d'informació per a la descripció contextual de fotos es basen tant en el cos de la notícia com en la descripció visual de la imatge. El cos de la notícia proporciona el context necessari i ofereix detalls sobre l'esdeveniment, les persones involucrades i l'entorn. D'altra banda, la descripció visual de la foto permet identificar elements específics, com ara colors, expressions facials i accions, que complementen i enriqueixen la informació presentada al text. Combinant ambdues fonts, es pot aconseguir una descripció precisa i completa que facilita la comprensió del lector.

L'aplicació podria tenir en compte la meta-informació de la foto, a vegades present en el format EXIF, perquè pot contenir informació de context rellevant, com la font de la foto i/o una versió del peu escrita per la mateixa agència. Finalment, aquesta opció es va descartar perquè la informació en aquest camp no sempre està present. No obstant, podria ser una font interessant a considerar en una solució futura amb més funcionalitats.

En aquest projecte es va descartar també l'addició d'informació textual d'ajuda per part de l'usuari per generar el peu de foto i el text alternatiu, per tal de reduir la necessitat d'intervenció.

2.4. Funcionalitats i Interacció

L'objectiu és que l'aplicació a desenvolupar ha de poder donar un resultat sense una intervenció directa de l'usuari. Tota la informació l'ha d'obtenir del context, el qual estarà format pel text de la notícia i la fotografia, sempre seguint els criteris editorials de l'organització.

De manera opcional, l'usuari podria intervenir per confirmar la presència d'algunes persones a la fotografia per tal de suplir les limitacions dels models actuals a l'hora d'identificar persones a escala.

El sistema té en compte el llibre d'estil de 3CAT de manera integral i, per tant, els criteris editorials no es poden seleccionar de forma manual. Els límits establerts per a la generació de text són els següents:

- **Peu de foto:** Màxim de 90-110 caràcters.
- **Text alternatiu:** Al voltant de 140 caràcters, però no és un límit rígid.

Una possibilitat que es va contemplar inicialment va ser que el sistema generés diferents versions de subtítols i que el professional pogués triar posteriorment. Aquesta opció es va descartar amb l'objectiu de què l'eina s'enfoqués a la generació d'un resultat final òptim per tal d'estalviar la intervenció del professional.

Llibre d'estil



El llibre d'estil d'un mitjà de comunicació és una guia essencial per a professionals de la comunicació per a reflectir els valors corporatius en la forma com es transmet la informació per part dels seus professionals. L'estil no només defineix la manera com es presenta la informació, sinó que també influeix en la percepció i la comprensió del públic. El professional pot consultar-hi com l'ús adequat del llenguatge, el to, la claredat i la coherència poden millorar la qualitat de la comunicació. És important mantenir un estil consistent per construir credibilitat i confiança amb l'audiència.

Per tal de comprendre l'impacte del llibre d'estil en la implantació de l'aplicació objecte d'aquest treball, descriuen a continuació els punts més rellevants del llibre d'estil de 3CAT que interpel·len al professional a l'hora d'escriure descripcions de peus de foto així com el text alternatiu que afegeix una descripció visual a la foto en el seu format web.

3.1. Denominacions

Des de la perspectiva dels mitjans del 3cat, termes com *país*, *nació*, *nacional*, *govern* i *Parlament*, entre d'altres, fan referència a Catalunya si no s'indica una altra cosa.

Els mitjans del 3CAT fan servir indistintament *Espanya* i *estat espanyol* per referir-se a una mateixa realitat geopolítica. Quan la referència és de caràcter geogràfic o físic, evitem l'expressió *estat espanyol*. Fan servir també l'expressió *Països Catalans* per referir-se al conjunt de territoris de parla catalana des del punt de vista històric, geogràfic, lingüístic o cultural.

3.2. Paritat i equilibri de gènere

Es vetlla per la igualtat entre dones i homes, i es reflecteix en els continguts i la projecció pública de 3CAT amb una presència equilibrada dels dos sexes. S'evita els estereotips de gènere i el llenguatge sexista.

3.3. Imparcialitat, neutralitat i pluralisme

S'actua amb imparcialitat, neutralitat i respecte pel pluralisme. En un conflicte, es manté l'equilibri raonable a l'hora de donar presència a les parts, sense reduir la informació a la simple cessió acrítica d'espai a aquestes parts.

No es pren partit, es contextualitzen les notícies i s'ofereixen elements de judici a la ciutadania.

3.4. Fets violents

No es difonen continguts que incitin a la violència, física, sexual o psicològica o que la justifiquin. Això no va en detriment del dret a la informació o de la llibertat de creació.

En el cas que es consideri imprescindible mostrar continguts de violència extrema pel seu interès informatiu, s'adverteix prèviament que poden ferir la sensibilitat i es senyalitza de manera convenient.

No es converteix en espectacle la informació sobre fets o actes violents, ni es reproduïen estereotips discriminatoris cap a persones ni col·lectius. Es defuig el llenguatge agressiu, la morbositat i l'alarmisme. A les informacions sobre els possibles autors de fets violents es preserva la presumpció d'innocència.

Es respecta la intimitat i el dret a l'anonimat de les víctimes d'actes de violència, especialment si és sexual, o si es tracta de menors.

No es fa apologia dels maltractaments contra els animals i només es mostren si ho justifica l'interès informatiu.

3.5. Violència de gènere i violència domèstica

Per al 3CAT, la violència contra les dones és una violació dels drets humans i un atemptat contra la llibertat i la dignitat de les persones, tant si es tracta de violència física o sexual com psicològica.

Els mitjans del 3CAT no reproduïen estereotips i tòpics que contribueixen a frivoltzar, banalitzar o justificar la violència o la discriminació contra les dones, ni se'n fa espectacle.

En les informacions sobre violència sexista es respecta el dret a l'anonimat i la intimitat de les dones agredides, i no s'identifiquen sense consentiment previ. Igualment s'eviten imatges i expressions verbals sensacionalistes o morboses, així com declaracions que justifiquin l'agressió.

El tractament informatiu sobre la violència en l'àmbit familiar, que afecta sobretot menors d'edat i persones grans, segueix els mateixos criteris bàsics que la informació sobre la violència contra les dones.

No es fa servir l'expressió *violència domèstica* com a equivalent de violència masclista.

3.6. Conflictes bèl·lics i catàstrofes

En situacions de conflicte bèl·lic, s'ofereixen els diversos punts de vista i es proporcionen anàlisis imparcials i informació contextualitzada. Quan les fonts són parcials o tendeixen a instrumentalitzar la informació, es cita sempre la procedència de les dades per mantenir la neutralitat. En cas que algun dels bàndols censuri dades o imposi restriccions informatives, es comunica a l'audiència.

En la informació sobre catàstrofes, s'evita l'alarmisme i es fa l'esforç per difondre des del primer moment dades contrastades i obtingudes de fonts fidedignes.

3.7. Suïcidi

Es tracten els suïcidis, els intents de suïcidi i les autolesions amb extrema prudència, i només se'n fa ressò quan hi ha interès informatiu per la notorietat de la persona o per les circumstàncies del fet.

No es qualifica de suïcidi una mort si no se'n té la confirmació. Si es difon la identitat de la persona, es cita la causa de la mort sense fer servir eufemismes, a excepció que els seus familiars o persones properes es pronunciïn en sentit contrari.

S'evita produir i difondre continguts que incitin al suïcidi o que incloguin explicacions i detalls sobre procediments de suïcidi.

3.8. Immigració

Es valoren els fets per l'interès informatiu i no en funció de l'origen de les persones que els protagonitzen. Per evitar contribuir a l'estigmatització de minories ètniques o col·lectius d'immigrants, no es fan servir tòpics ni estereotips i no s'explicita l'origen de les persones si no és imprescindible per comprendre la informació.

Es parla de la immigració amb neutralitat i no com un problema.

No es converteixen les situacions personals en qualificatius ni en etiquetes simplificadores de persones o col·lectius, i s'evita fer servir els termes *il·legal*, *indocumentat*, *irregular* o *sense papers* per referir-se a algú. Se'n pot fer referència com a *treballador*, *immigrant* o *persona en situació irregular*.

3.9. Les persones grans

Es tracta l'envelliment com un procés vital natural i es presenten les persones grans com a part activa de la societat. Es mostren les seves aportacions en els diferents àmbits en què intervenen i no se n'infravalorem les capacitats ni es fan aparèixer només en continguts relacionats amb la dependència o la malaltia.

3.10. Crisi climàtica

Es divulga la recerca científica sobre la crisi climàtica. En les seves informacions 3CAT es basa en la veracitat i en els coneixements científics.

No es dona veu a postures acientífiques ni a empreses o corporacions condemnades per delictes contra el medi ambient. Es reconeix que una aclaparadora majoria de la comunitat científica avala la gravetat, les causes i les conseqüències de l'escalfament global.

Models Grans de Llenguatge (LLMs)



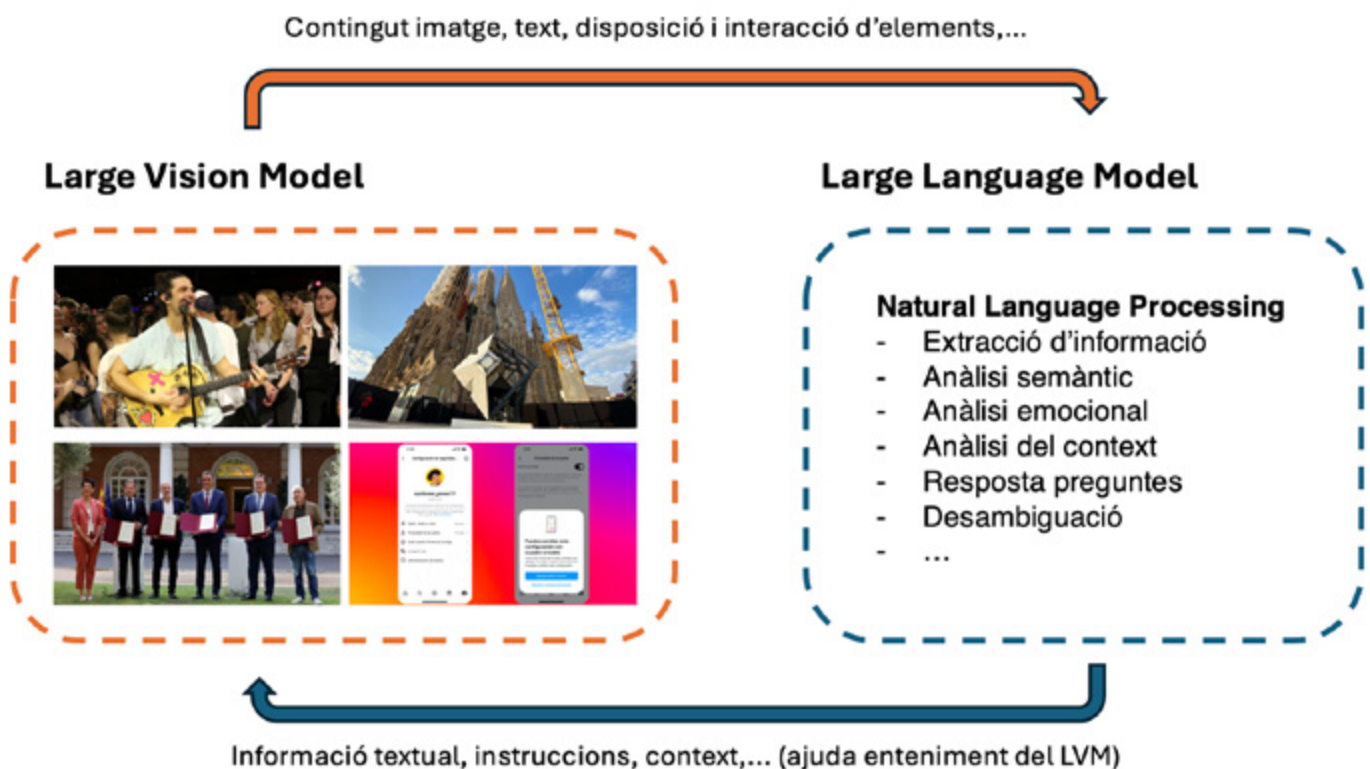
Els grans models de llenguatge (LLM) són potents sistemes d'intel·ligència artificial entrenats amb enormes quantitats de dades de text per entendre i generar un llenguatge semblant al dels humans. La seva capacitat d'adaptar-se a una gran varietat de tipus de contingut els fa especialment útils en tasques com el resum de texts i la generació de subtítols per a imatges, totes dues essencials en àmbits com el dels mitjans de comunicació.

El resum de contingut és una de les principals fortaleses dels LLM, especialment per a articles de notícies. Aquest resum es pot fer de dues maneres clau: resum extractiu i resum abstractiu. El resum extractiu implica seleccionar i extreure les frases o les parts més importants del text original. Els LLM, gràcies a la seva profunda comprensió contextual, puntuen (valoren) cada frase basant-se en factors com la rellevància i la coherència. Així, el model destil·la les parts essencials de l'article de notícies per generar resums concisos que capturen els punts clau.

El resum abstractiu, en canvi, permet al model generar noves frases que transmeten el significat general del text d'entrada. Això és especialment valuós en contextos de notícies, on sovint es necessita reformular els resums per obtenir més claredat o brevetat. Els LLM utilitzen la seva capacitat contextual per parafrasejar el contingut, oferint una major flexibilitat a l'hora de condensar articles de notícies en peces clares i digestibles.

Al mateix temps, la generació de subtítols per a imatges juga un paper crucial en els mitjans, on sovint les imatges acompanyen les notícies. Aquesta tasca implica generar text descriptiu que expliqui què succeeix en una imatge. Els LLM i els models multimodals combinen el reconeixement visual—impulsat per xarxes neuronals convolucionals (CNN)—amb tècniques de generació de llenguatge, com els models basats en transformadors, per crear subtítols que siguin coherents i rellevants per al contingut. El model no només reconeix objectes i accions dins de la imatge, sinó que també pot rela-

Un gran model de llenguatge multimodal obtingut combinant un model de gran visió i un gran model de llenguatge.



cionar-los amb el context de l'article, generant subtítols significatius que complementen el text que l'acompanya.

En ambdues tasques, els models multimodals, que integren tant dades visuals com textuais, són fonamentals. Aquests models poden gestionar tasques de processament tant d'imatges com de llenguatge, creant descripcions o resums coherents a partir de diferents tipus d'entrades. Alguns sistemes utilitzen instruccions (prompts) com preguntes específiques o el context més ampli d'un article de notícies, per guiar el procés de subtitulat o de resum, millorant-ne la rellevància. Altres generen automàticament subtítols o resums basant-se només en les dades d'entrada, una funció útil per a la cobertura en temps real dels mitjans.

Els LLM es poden ajustar per treballar en dominis específics, com el de les notícies, cosa que els fa adaptables a diferents tipus de contingut i estils, i els permet gestionar una àmplia gamma de llengües i disciplines. Tanmateix, els biaixos presents en les dades d'entrenament poden influir en la qualitat dels resums o subtítols generats, especialment quan el corpus d'entrenament no és prou divers o representatiu.

Tot i que els LLM han fet avenços impressionants tant en el resum de contingut com en la generació de subtítols per a imatges, assegurar la coherència i la claredat en articles i imatges diversos segueix sent un repte. El perfeccionament continuat d'aquests models millorarà encara més la qualitat i la fiabilitat dels resultats, ajudant els LLM a esdevenir eines efectives per processar continguts mediàtics complexos.

En resum, els LLM ofereixen un gran potencial per automatitzar i millorar la generació de resums i subtítols per a articles de notícies. La seva capacitat per processar entrades complexes i generar representacions clares i concises tant de text com d'imatges els fa valuosos en l'àmbit dinàmic dels mitjans de comunicació.

4.1. Models multimodals: descripció d'imatges

La tasca requereix una de dues aproximacions: combinar diversos models o utilitzar un únic model multimodal. Tot i que la primera és un enfocament tradicional, implica una investigació extensa, ajustaments i pot ser més tediós d'implementar.

En canvi, els models multimodals han avançat significativament en els darrers anys i estan guanyant popularitat. A més, treballar amb diversos models de codi obert simultàniament augmenta exponencialment la complexitat. Per aquests motius, l'exploració comença amb models multimodals per avaluar-ne el rendiment. Si un únic model resulta insuficient, es pot explorar i integrar un model especialitzat per a la tasca que no compleixi. Aquest enfocament redueix la complexitat global de la tasca, alhora que permet explorar les capacitats actuals dels LLM multimodals.

Els models que s'han explorat s'han seleccionat a partir dels seus rànquings a les plataformes de PapersWithCode¹ i Hugging Face². La taula següent presenta els models avaluats juntament amb les seves especificacions. Per provar ràpidament el rendiment dels models, se'ls ha donat instruccions clares del que s'espera, una imatge i el text complet d'un article de notícies de 3CAT.

El model seleccionat ha estat Pixtral, doncs és un dels models multimodals més avançats actualment, combinant visió i LLM amb un rendiment destacat en la interpretació d'imatges i text. A més, és un model relativament compacte, fet que facilita tant la seva implementació com l'ajust fi en comparació amb alternatives més grans i costoses.

Tot i que no està plenament optimitzat per al català, el fet que sigui desenvolupat per Mistral, una empresa especialitzada en llengües europees, li dona un avantatge respecte a altres models anglo-cèntrics. Aquest enfocament permet explorar-ne el potencial en el nostre cas d'ús i, si cal, complementar-lo amb models més especialitzats en llengua.

¹ <https://paperswithcode.com>

² <https://huggingface.com>

4.2. Prompt engineering, agentic AI i raonament en passos

MODEL	Variant	Altres Limitacions	Llicència	URL
BLIP	-	BLIP de LAVIS, però fins i tot amb el model base, CUDA demana molta memòria.	-	-
Pixtral	Pixtral 12B	-		https://mistral.ai/news/pixtral-12b/
Llava – One vision	lmms-lab/llava-onevision-qwen2-0.5b-si	-	Apache 2.0	https://huggingface.co/lmms-lab/llava-onevision-qwen2-72b-ov-sft
CLIP		Fa servir GPT2, que té un rendiment molt baix	MIT license	https://github.com/rmokady/CLIP_prefix_caption
qwen2	vl-7b-instruct		Apache 2.0	-
Molmo	Desconegut (web oficial) Molmo- 7B-D-0924	Aquest model és força car de fer funcionar.	Apache 2.0	-
Reka	core-20240904	-	MIT license	https://github.com/kyegomez/Reka-Torch
Reka	Flash-20240904	-	MIT license	https://github.com/kyegomez/Reka-Torch
InterntVL	InternVL – 26b	-	MIT license	https://github.com/OpenGVLab/InternVL#quick-start-with-huggingface
InternVL	InternVL – 4b	Massa sortida (incontrolable)	MIT license	https://github.com/OpenGVLab/InternVL#quick-start-with-huggingface

Prompt engineering (PE) és un aspecte crític de la interacció amb grans models de llenguatge (LLM) i agents d'IA. Implica dissenyar instruccions o preguntes el més adients possibles que guiïn aquests models per produir respostes rellevants, precises i riques en el context. Amb l'augment dels agents d'IA, el PE ha evolucionat més enllà de les simples consultes esdevenint en una tècnica específica per a dissenyar fluxos de treball estructurats que permeten els agents d'IA raonar, planificar i prendre accions de manera eficaç. Quan s'aplica a tasques

de raonament complexes, com ara la descripció d'imatges, aquests agents han d'integrar la percepció, la memòria, el coneixement i la presa de decisions per generar resultats significatius.

Per generar descripcions d'imatges precises i conscients del context, els agents d'IA han de seguir un procés estructurat de diversos passos. A dalt es mostra un procediment amb sis accions que s'alinea amb el pensament basat en agents i els principis de PE:

1. Percepció: extreure les característiques visuals clau. L'agent d'IA analitza primer la imatge per extreure característiques essencials, com ara objectes, colors, textures i relacions espacials. En aquest cas especialment rellevant són les persones i la seva disposició a la imatge. Els models avançats utilitzen tècniques de visió per computador per detectar entitats, classificar escenes i reconèixer activitats. Un exemple: *identificació d'un gat assegut a l'ampit de la finestra amb la llum del sol entrant*.

2. Comprensió del context: inferir el significat de l'escena. Més enllà de la detecció bàsica d'objectes, l'agent ha d'inferir contexts complexos que tenen a veure amb el reconeixent les interaccions o les narracions implícites. Les metadades externes (per exemple, la geolocalització, les marques de temps i la intenció de l'usuari) poden millorar la comprensió. Exemple: *si un grup de persones està al voltant d'un pastís d'aniversari, la IA pot deduir què s'està celebrant*.

3. Integració de la memòria: fent referència a coneixements passats. L'agent aprofita la memòria per connectar la imatge amb coneixements previs, preferències de l'usuari o interaccions històriques. En aquest cas, per exemple, la guia d'un llibre d'estil ètic i periòdic. Exemple: *si l'usuari pot donar indicacions per tal de modular el vocabulari inclusiu i respectuós*.

4. Recuperació de coneixements: enriquir descripcions amb dades, també externes. Els agents d'IA accedeixen a bases de dades, coneixements enciclopèdics o dades de formació prèvia per proporcionar un context addicional. Aquest context en aquest cas és directament el cos de la notícia, però potencialment podria accedir a bases de dades externes relacionades, com fer consultes a internet. Exemple: *si una imatge mostra un punt de referència com la Torre Eiffel, l'agent pot inferir la informació rellevant actual del monument o més enllà podria recuperar la importància històrica o cultural*.

5. Raonament i planificació: estructurar una descripció coherent. L'agent d'IA estructura una resposta lògica i atractiva seqüenciant els detalls de manera adequada. És important l'equilibri de la granularitat, evitant l'excés de detall alhora que es garanteixen coneixements significatius. Exemple: *per una banda podríem tenir una descripció senzilla com "Un arbre amb fulles", mentre que en un altre extrem es podria generar una resposta excessivament detallada com "Un roure alt amb fulles daurades de tardor balancejant-se a la brisa"*.

6. Iteració i perfeccionament mitjançant la retroalimentació. Els agents d'IA poden refinar les respostes mitjançant la retroalimentació dels usuaris, l'aprenentatge de reforç i el modelatge adaptatiu. L'aprenentatge continu garanteix una major precisió i una millor alineació contextual al llarg del temps. Exemple: si un usuari corregeix la descripció d'una raça de gossos feta per una IA, les respostes futures incorporen aquest aprenentatge.

A diferència dels models tradicionals que generen text de manera passiva, els agents d'IA raonen activament a través d'un problema, recuperen el coneixement rellevant i perfeccionen els resultats de manera iterativa.

Experimentació



Al llarg del projecte, es van realitzar diverses rondes d'experimentació i en cadascuna es van realitzar diverses observacions que es van incorporar de manera iterativa en cada iteració, afegint elements augmentant així la complexitat del sistema a cada iteració.

5.1. Conjunt de notícies

Es van seleccionar 14 notícies amb les seves imatges corresponents (31 en total) per a una experimentació preliminar ([Annex B](#) i [Annex C](#)). Per a l'avaluació final seran 13 les notícies diferents ([Annex D](#)).

Les notícies cobreixen una varietat de temes, i les imatges inclouen una àmplia gamma de subjectes, com ara figures polítiques, membres del públic, objectes quotidians, objectes inusuals, cossos celestes i captures de pantalla d'interfícies d'usuari.

Aquesta diversitat d'objectes i situacions està dissenyada per avaluar la capacitat dels models per comprendre i generar amb precisió textos alternatius per a persones amb discapacitat visual o escriure llegendes per a elements visuals i situacions tant comunes com inusuals.

5.2. Experimentació models de traducció

Tot i que el corpus d'entrenament en anglès és realment massiu comparat amb el català, els models més grans, com ara el ChatGPT o Gemini, mostren un rendiment alt també en català. Ara bé, quan treballem en servidors local amb models de capacitat limitada (al voltant de 10k milions de paràmetres) el rendiment en català observat en les primeres proves va ser insuficient.

Es per aquesta raó que, després d'un pas previ de traducció, tot el procés descriptiu i inferència fet pel model de llenguatge es fa en anglès i posteriorment es torna a traduir al català.

Tot i no ser part principal de l'objectiu del projecte, el pas de traducció va a comportar nous reptes que descrivim a continuació. Cal esmentar abans, que en fases properes a producció, la consulta a serveis de pagament de tercers més eficients son perfectament compatibles amb aquesta solució per tal de treballar directament en una sola llengua.

5.2.1. Traducció de la notícia

Durant les proves es va descobrir que alguns dels problemes es podien mitigar amb un text més ben traduït. Per aquest motiu, es van explorar diferents models de traducció.

Com a part del projecte AINA¹, una iniciativa centrada en fomentar l'ús del català en contextos tecnològics i acadèmics, es van explorar els recursos lingüístics de codi obert dedicats a la llengua catalana desenvolupats per aquest projecte. En aquest marc, es van provar diversos models, incloent-hi FLOR-6.3B, una de les opcions més populars. No obstant això, malgrat la seva reputació, es va veure que aquest model basat en instruccions va oferir un rendiment insatisfactori pel que fa a la qualitat de la traducció.

A més, dins del mateix marc, també es van avaluar models específics per a la traducció, com ara *aina-translator-en-ca*, *aina-translator-ca-en* i *diverses configuracions del model Plume* (32k, 128k i 256k). Entre aquests, el Plume 256k va demostrar la millor qualitat de traducció, oferint una solució unificada sense necessitat de canviar de model segons la direcció de la traducció. Tot i aquest avantatge, es va observar una degradació del rendiment en traduccions de textos més llargs.

Per abordar aquest problema, es van explorar tècniques de processament del llenguatge natural per dividir els textos automàticament en fragments més o menys lògics. Tot i això, en provar eines com NLTK² per a aquesta tasca, es va comprovar que no sempre generen divisions adequades, fet que en alguns casos provoca la pèrdua de context general durant la traducció, amb l'impacte negatiu consegüent en la qualitat global del resultat.

Finalment, es va provar el model Llama 3.1 8b (model en codi obert desenvolupat per Meta), altament valorat per al català a Hugging Face La-Leaderboard³. Tot i que de vegades introduïa barbarismes, va demostrar un rendiment sòlid en la traducció de català a anglès, fet que el fa una opció adequada. Considerant que el model Pixtral-12B ofereix un rendiment suficientment bo en la traducció d'anglès a català, es va decidir alternar entre tots dos models durant la inferència.

¹ <https://projecteaina.cat/> - <https://huggingface.co/projecte-aina>

² <https://www.nltk.org/>

³ <https://huggingface.co/spaces/la-leaderboard/la-leaderboard>

5.2.2. Traducció del llibre d'estil

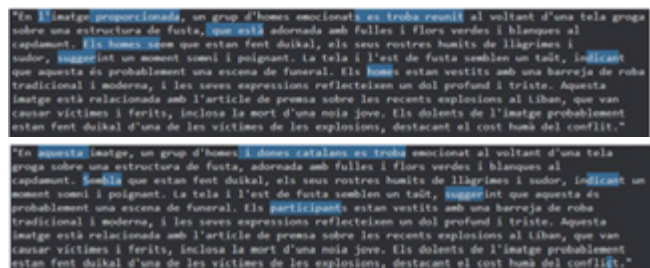
Durant l'experimentació i implementació del llibre d'estil, es va notar un comportament de biaix inusual del model. El llibre d'estil va ser presentat com una font del model tant en català com en anglès. Quan se li va proporcionar la versió en anglès (traduïda), el model estava molt més inclinat a alterar l'ús de termes que defineixen el gènere o a incloure (inadequadament) adjectius i adverbis relacionats amb Catalunya.

Per exemple, en la comparació de textos a continuació, es pot observar que després d'introduir el llibre d'estil, el model canvia la seva "opinió" sobre el gènere i la nacionalitat de les persones presents en la imatge. A més, es pot veure que canvia la paraula "homes" per "participants". També es va observar una inclinació a utilitzar el terme "Països Catalans" de manera més freqüent de l'habitual o fora de context.

Aquestes observacions són preocupants ja que indiquen que la integració del llibre d'estil pot estar introduint involuntàriament biaixos en el resultat del model. Per tant, l'estructura del llibre d'estil va ser alterada amb la intenció de mitigar aquests biaixos.

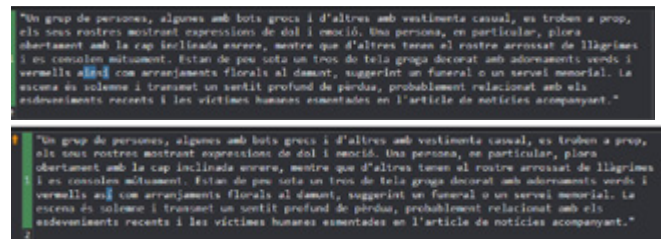
A tal efecte es va reestructurar el llibre d'estil, organitzant-lo en seccions temàtiques per facilitar-ne la interpretació pel model. A més, es va reduir la seva extensió per evitar problemes de contextualització dins de la finestra de context del model, ja que, quan la quantitat d'informació és massa gran, la qualitat del resultat tendeix a degradar-se. Aquesta reducció també assegura que les instruccions es compreguin correctament i no generin interpretacions errònies, millorant així la precisió i la coherència del model sense introduir biaixos no desitjats.

Demostració de comportament esbiaixat en introduir el llibre d'estil en anglès. Esquerra – text generat abans de sol·licitar al model que utilitzi les regles del llibre d'estil. Dreta: text alterat amb les regles del llibre d'estil.



D'altra banda, aquestes alteracions no van ocórrer quan es va passar el llibre d'estil al model en català. De fet, per a la inferència realitzada en aquesta mateixa imatge, hi va haver molt poca diferència.

Demostració de comportament esbiaixat en introduir el llibre d'estil en català. Esquerra: text generat abans de sol·licitar al model que utilitzi les regles del llibre d'estil. Dreta: text alterat amb les regles del llibre d'estil.



Hi ha múltiples explicacions per a això. Una d'elles és que les dades d'entrenament subjacents del model en ambdós idiomes poden tenir biaixos existents relacionats amb el gènere i la nacionalitat, que el llibre d'estil podria estar exacerbant en comptes de mitigar. Una altra opció és que la traducció i interpretació de certs termes entre el català i l'anglès pot portar biaixos inherents que s'estan amplificant quan s'aplica el llibre d'estil en diferents idiomes.

En qualsevol cas, quan es va descobrir aquest biaix, es va monitoritzar i tenir en compte mitjançant l'ajust de certes instruccions, frases en el llibre d'estil i mitjançant l'ús de la versió en català (sent aquesta la que porta menys variació).

5.3. Observacions experimentals: instrucció, raonament i estil

La taula de l'Annex A mostra els resultats d'utilitzar el model multimodal Pixtral 12B per generar llegendes i textos alternatius per a la imatge i article de notícies respectius. Aquest llistat d'articles i imatges són una selecció inicial reduïda i estreta de 3CAT amb la intenció de fer proves preliminars. A la columna més a la dreta de la taula es poden observar les notes d'avaluació. Entre d'altres no compleix el llibre d'estil i dona excés d'informació.

Tenint en compte els comentaris proporcionats a la taula anterior, es van implementar millores en la configuració i es va generar una nova ronda de resultats. Degut a la seva extensió, aquests resultats es mostren a l'[Annex B](#).

Per a aquesta iteració, es va utilitzar el conjunt d'articles de notícies mencionat anteriorment per a l'avaluació. Es van introduir canvis en l'estratègia de redacció dels *prompts*, fent-los més detallats i dividint-los en segments lògics que es van introduir al model de manera seqüencial. A més, es van incorporar les instruccions del llibre d'estil al procés.

Durant les proves, es van observar les següents qüestions:

- **Llibre d'estil:** El llibre d'estil inclou criteris específics per a Catalunya. No obstant això, la versió en anglès introdueix biaixos addicionals.
- **Es va afegir una nova norma al llibre d'estil:** "Quan es faci referència al lloc d'origen, donar preferència als termes de l'àrea cultural o lingüística en lloc del nom de l'estat."
- **Chain of Thought (CoT):** Es va provar el CoT, una tècnica que consisteix a dividir el procés de raonament en passos intermedis per millorar la capacitat del model d'entendre i resoldre problemes complexos. Tot i això, no va oferir bons resultats amb el model, ja que causava al·lucinacions i la generació d'informació addicional no suportada.

Després d'aquesta ronda de comentaris, es van identificar els principals punts per treballar:

1. El model ha d'evitar suggeriments, explicacions i justificacions.
2. El model ha de poder identificar persones com a pas previ, amb la possibilitat que l'usuari validi aquestes identifications.
3. L'estructura del text alternatiu ha de prioritzar primer la descripció de la imatge i després la informació contextual.

Tenint en compte els punts esmentats anteriorment, es va dur a terme una altra ronda de millores. Es va construir un flux de treball més lògic ja enfocada cap a la creació d'una interfície d'usuari (UI). Els prompts es van ajustar amb l'objectiu d'aconseguir que el model fos equilibrat, generant un text concís i precís sense suggerir ni sobre explicar.

Els resultats es poden observar a l'[Annex C](#), juntament amb els punts d'avaluació a la columna més a la dreta. Aquesta ronda de comentaris es va centrar a reduir les especulacions, fer que el model fos més precís i directe, i que utilitzés un llenguatge més gramaticalment correcte.

Generació de descripció contextualitzada: flux final



Basant-se en els comentaris de diverses iteracions, es va desenvolupar un flux de processament i conversació amb el model LLM per tal d'automatitzar al màxim la tasca, complint els requisits de rendiment i incorporant la intervenció/comentaris de l'usuari.

El flux de processament s'il·lustra al diagrama següent, on les caixes blaves representen accions realitzades per l'usuari, i les caixes vermelles indiquen accions gestionades pels models d'IA. Tenint en compte les exigències computacionals dels models, la seqüència actual d'accions d'aquest flux està optimitzat per equilibrar els requisits de rendiment amb el temps d'inferència.

Al principi del flux, es carrega el model Llama per traduir les notícies a l'anglès i extreure noms del text. La traducció es realitza perquè el model Pixtral va demostrar un millor rendiment quan se li proporciona context en anglès.

L'usuari interactua amb la interfície per confirmar la presència de persones a la imatge, una tasca facilitada pel fet que el model extreu noms directament del text. Aquests noms es poden utilitzar per crear una llista ordenada de persones, permetent que l'usuari especifiqui individus rellevants per a la imatge.

Un cop el context necessari està preparat, es generen el text alternatiu i el peu de foto a partir del primer utilitzant el model Pixtral. Ambdós es produeixen inicialment en anglès, on el model aplica les normes d'estil abans de traduir els resultats al català. Tot i que aquesta tasca de traducció podria beneficiar-se potencialment del model Llama, el model Pixtral ofereix un rendiment adequat, evitant la càrrega i descàrrega entre models.

6.1. Instrucció dels models generatius: pasos de raonament

A continuació es descriu explícitament tot el flux de la conversa amb els models gran de llenguatge LLMs. El pasos de raonament (Chain-of-Thought COT prompting), permet capacitats de raonament complexes mitjançant passos de raonament intermedis.

1. Obtenir noms del text (Llama 3.1 8b). Passem la notícia al model:

I'm going to provide you with a text. Your task is to identify and extract the names of people that are explicitly shown in the text. Do not translate the names. Do not explain yourself. Just list the names from the text. Don't numerate or anything similar.

{News}

2. Traduir text de català a anglès (Llama 3.1 8b). Passem la notícia al model:

Translate all texts from Catalan to English without changing any meaning. Also don't translate names.

{News}

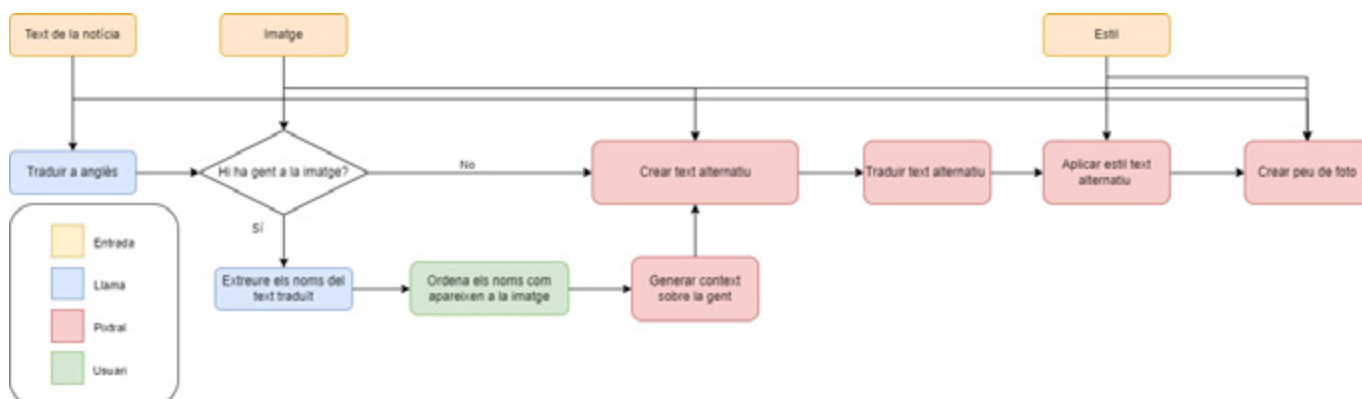
3. Assignar noms a persones (Pixtral). Passem la foto i els noms ordenats al model:

I'm going to provide you with an image. Your task is to identify the number of people or groups in the image and determine whether each represents an individual or a collective. Use visual context for inference, but do not explain yourself.

Draw a conclusion!

{Image}

Il·lustració flux de processament del sistema: conversa LLM i interacció amb l'usuari.



Please tell me how many people are in the image for each of the following names.

Important clarification: Treat any group with two or more people as a collective, even if it consists of only two individuals. Do not infer based on context—if there are two or more people together, it must be classified as a collective. An individual refers strictly to a single person.

The names correspond to positions in the image, following a left-to-right, top-to-bottom order, like reading a book. The pattern is: Top row, left to right. Next row, left to right, and so on.

The names may refer to a single person or a group. Name(s) and positions:

{Names ordered}

For each name, specify the number of people and whether it is an individual or a collective. Draw a conclusion based strictly on the rule above, without interpretation.

4. Construint el text alternatiu sense noms (Pixtral).
Passem la notícia i la imatge

User:

I'm going to provide you with a piece of news and an accompanying image. Your task is to create alt text suitable for blind individuals, describing the image in detail while integrating relevant context from the news. The goal is not to explain the news itself, but to use its information to provide a better understanding of the image. Focus on describing the visual elements, enriched by any relevant context from the news to give a fuller sense of the scene.

Assistant:

Understood! I'll wait for the news and image before generating the alt text. Once received, I'll describe the image in detail and combine it with the relevant context from the news, without explaining the news itself. Also, there are people or monuments I will mention them.

User:

Create a concise alt text for blind people that explains the image and how it connects to the text below. Be as detailed as possible. Take into account that it's very important, if there are people state who are they. Only write the alt text. Nothing more, nothing else." + "\n" + {news_text_en}

Assistant:

Waiting for you to send the image.

User:

{Image}

5. Construint el text alternatiu amb noms (Pixtral).
Passem la notícia, la imatge i els noms ordenats:

User:

I'm going to provide you with a piece of news and an accompanying image. Your task is to create alt text suitable for blind individuals, describing the image in detail while integrating relevant context from the news. The goal is not to explain the news itself, but to use its information to provide a better understanding of the image. Focus on describing the visual elements, enriched by any relevant context from the news to give a fuller sense of the scene.

Assistant:

Understood! I'll wait for the news and image before generating the alt text. Once received, I'll describe the image in detail and combine it with the relevant context from the news and the people context - {people_context}, without explaining the news itself. Also, if there are people or monuments, I will mention them.

User:

Create a concise alt text for blind people that explains the image and how it connects to the text below. Be as detailed as possible. Take into account that it's very important, if there are people state who are they. Only write the alt text. Nothing more, nothing else.

{News}

Assistant:

Waiting for you to send the image.

User:

{Image}

6. Traduir text alternatiu d'anglès a català (Pixtral).
Passem el text alternatiu en anglès:

Assistant:

{alt_text_en}

User:

Tradueix el alt text al català. Utilitza els noms propis tal com són, sense traduir-los.

7. Aplicar estil (Pixtral). Passem el text alternatiu en català i l'estil:

User:

Tradueix el alt text al català. És extremadament necessari que només escriguis el text alternatiu. Res més.

Assistant:

{alt_text_cat}

User:

Ara vull que milloris el text alternatiu en català i
fent ús del llibre d'estil que t'enviaré: {style_guide}

8. Obtenir peu de foto (Pixtral). Passem notícia en català, la imatge, la guia d'estil i el text en català amb l'estil aplicat:

User:

Aquí tens el text de la notícia: {news_text_cat}

User:

Aquí tens la imatge.

{Image}

User:

Aquí tens el llibre d'estil: {style_guide}

Assistant:

{alt_text_catalan_wth_style}

User:

Genera un títol en català per a la imatge seguint el llibre d'estil, amb un màxim de 120 caràcters i sense utilitzar signes de dos punts (:). Has de tenir en compte la descripció de la imatge i el context de la notícia.

6.2. Aspectes tècnics dels models i rendiment computacional

Pixtral té un pes aproximat de 25 GB i actualment consumeix uns 17 GB de memòria a la GPU i 25 GB a la CPU. No obstant això, la suma d'aquests valors no equival directament al pes total del sistema, ja que hi ha un *overhead* associat a la gestió de recursos compartits entre la GPU i la CPU. A més, una part significativa de la memòria es reserva per a la finestra de context, que augmenta proporcionalment a la mida màxima de context configurada. Aquesta finestra de context es duplica tant a la GPU com a la CPU, contribuint també a l'ús total de memòria.

El sistema pot gestionar fins a 8.000 *tokens* (aproximadament 6.000 paraules) i una imatge per *prompt* com a màxim. La velocitat de processament del model és de 16 *tokens* per segon, cosa que determina la rapidesa amb què es responen les peticions.

Llama 3.1 8b té un pes aproximat de 16 GB i actualment consumeix uns 17 GB de memòria a la GPU. El sistema pot gestionar fins a 10.000 *tokens* com a màxim, i la seva velocitat de processament és de 37 *tokens* per segon, fet que determina la rapidesa amb què es responen les peticions.

De mitjana, executar tot el flux de processament amb una notícia i una imatge triga al voltant de 3-4 minuts en una targeta Nvidia L40S. Tot i que el rendiment podria arribar a 1 minut amb targetes de capacitat computacional més elevada.

La primera part (abans de la intervenció de l'usuari) triga al voltant d'un minut. Una de les raons és que el model Llama és més lleuger que el model Pixtral i necessita considerablement menys temps per carregar-se a la GPU.

La segona raó és la quantitat de context que es proporciona. Al final del flux de processament, el model Pixtral rep una quantitat immensa d'informació, la qual cosa requereix temps per processar-la i utilitzar-la.

Interfície d'usuari



La interfície d'usuari es va construir amb la llibreria de software Gradio. Això permet un desplegament fàcil dels models d'aprenentatge profund, però a la vegada limita la personalització dels elements en comparació amb les pràctiques més tradicionals de desenvolupament web.

Tota el flux de processament es pot executar amb tan sols 2 clics, considerant que no hi ha persones i sense comptar els clics necessaris per enganxar el text i la imatge:

1. Fent clic als botons "Obtenir noms del text" i "Generar text alternatiu i peu de foto".
2. Si hi ha persones a la imatge, després d'executar el primer botó, l'usuari haurà d'interactuar amb un menú desplegable. Aquest s'utilitza per indicar el nombre de persones que l'usuari no reconeix o no vol incloure en el model, fet que donarà lloc al mateix nombre de marcadors de posició anomenats 'Desconegudes' al menú desplegable. A més, l'usuari pot afegir manualment un nom mitjançant un camp d'entrada situat sota el menú desplegable.
3. Després d'escollir els noms, juntament amb els marcadors de posició buits, l'usuari pot fer clic al botó inferior i obtenir els textos.

Il·lustració de la interfície d'usuari final: amb el camp d'inserció de la notícia, la fotografia, la selecció de persones i els 2 camps de generació amb el text alternatiu i el peu de foto.

Generador de text alternatiu i peu de foto

Text de notícies

Enganxa el teu text de notícies...

Obtenir noms del text

Arrastra i deixa anar una imatge

Coloque la imagen aqui

Haga click para cargar

Nombre de personas 'Desconegudas' a la foto

Selecciona quantes entrades 'Desconegudas' vols afegir al menú desplegable de noms

0 3 10

Tria els noms tal com apareixen a la imatge, de dalt a baix i d'esquerra a dreta

Selecciona un o més noms del menú desplegable.

Afegex noms manualment

Escriu els noms separats per comes

Afegir nom

Generar text alternatiu i peu de foto

Text alternatiu

Text alternatiu generat...

Peu de foto

Subtítol generat...

Començar amb una nova notícia

A continuació, detallem pas a pas l'ús de l'aplicació.

0. Accedeix a la pàgina i tria la notícia i la/s fotografia/es que vol utilitzar.

1. Enganxa l'article de notícies complet al primer quadre de text anomenat "Text de notícies".

Nota: Tot i que, en principi, és millor netejar l'article d'elements extres (com enllaços externs, informació sobre l'autor, etc.), l'aplicació funciona bé fins i tot si aquests elements hi són presents.

2. Enganxa la imatge de la qual vols obtenir el text alternatiu i el peu de foto.

Generador de text alternatiu i peu de foto

Text de notícies

Concerts de la Mercè 2024: horaris, escenaris i cartell d'artistes

Consulta, dia a dia, tots els artistes que passen pels escenaris de la festa major de Barcelona

Redacció

18/09/2024 - 19:20Actualitzat08/10/2024 - 11:47

compartir a Facebook

compartir a X

Si vols gaudir de tots els concerts de la Mercè 2024, et donem els detalls sobre els horaris, els escenaris i els artistes que protagonitzaran les actuacions musicals durant la festa major de Barcelona.

Els concerts tornaran a ser un dels plats forts de la Mercè i al llarg dels cinc dies hi haurà actuacions d'artistes com Figa Flawes, Mama Dousha, Oques Grasses, Maria Jaume, Triquell, Miki Núñez, La Fúrniga o Els Catarres.

Obtenir noms del text

Arrossega i deixa anar una imatge



Nombre de persones "Desconegudes" a la foto

Selecciona quantes entrades "Desconegudes" vols afegir al menú desplegable de noms

0

10

Tria els noms tal com apareixen a la imatge, de dalt a baix i d'esquerra a dreta

Selecciona un o més noms del menú desplegable.

Afegeix noms manualment

Escriu els noms separats per comes

Afegir nom

Generar text alternatiu i peu de foto

Text alternatiu

Text alternatiu generat...

Peu de foto

Subtítol generat...

Començar amb una nova notícia

3. Clica el botó etiquetat com "Obtenir noms del text" per obtenir els noms del text. Si la foto no conté persones que es volen mencionar, pots saltar directament al pas 6.

4. Tria quantes "Persones desconegudes" hi ha a la foto: persones o grups que no pots reconèixer però que consideres importants per al context de les notícies o fotografies. Ajusta el control lliscant situat a la banda dreta.

Generador de text alternatiu i peu de foto

Text de notícies

Concerts de la Mercè 2024: horaris, escenaris i cartell d'artistes

Consulta, dia a dia, tots els artistes que passen pels escenaris de la festa major de Barcelona

Redacció

18/09/2024 - 19:25 Actualitzat 08/10/2024 - 11:47

compartir a Facebook

compartir a X

Si vols gaudir de tots els concerts de la Mercè 2024, et donem els detalls sobre els horaris, els escenaris i els artistes que protagonitzaran les actuacions musicals durant la festa major de Barcelona.

Els concerts tornaran a ser un dels plats forts de la Mercè i al llarg dels cinc dies hi haurà actuacions d'artistes com Figa Flawes, Mama Dousha, Oques Grasses, Maria Jaume, Triquell, Miki Núñez, La Fúmiga o Els Catanes.

Obtenir noms del text

Arrastra i deixa anar una imatge

Nombre de persones "Desconegudes" a la foto

Selecciona quantes entrades "Desconegudes" vols afegir al menú desplegable de noms

0 10

Tria els noms tal com apareixen a la imatge, de dalt a baix i d'esquerra a dreta

Selecciona un o més noms del menú desplegable.

- JDRAN
- Jdrian
- Bkito
- Abdullah Miniawy
- Albaharo
- Ale Hop
- Alejandro y María Laura
- Alfred García
- Amantra
- Arrels de Gràcia
- Augusto Bracho
- Aymée Nuviola Quintet Feat. Kermuel Roig
- Alta Mon Amour

Generar text alternatiu

Text alternatiu

Text alternatiu generat...

Començar amb una nova imatge

5. Afegeix els noms i els desconeguts en ordre. Els noms s'han de triar de manera que indiquin la seva presència a la imatge en ordre de dalt a baix i d'esquerra a dreta (com en llegir un llibre).

En cas que els noms no apareguin a la llista desplegable, pots afegir-los utilitzant el camp "Afegeix noms manualment". Has d'escriure els noms, ja sigui un per un o separant-los amb comes, i fer clic al botó "Afegeix nom" per afegir-los.

Nota: El model detecta i reconeix si els noms afegits corresponen a grups o a persones, així que no cal posar el nom de cada persona en cas que no es vulgui.

Tria els noms tal com apareixen a la imatge, de dalt a baix i d'esquerra a dreta

Selecciona un o més noms del menú desplegable.

Chill Mafia

Figa Flawas

Afegir nom

6. Fes clic al botó taronja anomenat “Generar text alternatiu i peu de foto” per crear automàticament el text alternatiu i el peu de foto.

Generador de text alternatiu i peu de foto

Text de notícies

Concerts de la Mercè 2024: horaris, escenaris i cartell d'artistes
Consulta, dia a dia, tots els artistes que passen pels escenaris de la festa major de Barcelona
Redacció
18/09/2024 - 19:20 Actualitzat 08/10/2024 - 11:47
compartir a Facebook
compartir a X

Si vols gaudir de tots els concerts de la Mercè 2024, et donem els detalls sobre els horaris, els escenaris i els artistes que protagonitzaran les actuacions musicals durant la festa major de Barcelona.

Els concerts tornaran a ser un dels plats forts de la Mercè i al llarg dels cinc dies hi haurà actuacions d'artistes com Figa Flawas, Marna Dousha, Oques Grasses, Maria Jaume, Triquell, Miki Núñez, La Fúmbiga o Els Catarnes.

Obtenir noms del text

Arrossega i deixa anar una imatge

Nombre de persones 'Desconegudes' a la foto
Selecciona quantes entrades 'Desconegudes' vols afegir al menú desplegable de noms

0 10

Tria els noms tal com apareixen a la imatge, de dalt a baix i d'esquerra a dreta

Selecciona un o més noms del menú desplegable.

Figa Flawas

Afegir noms manualment

Afegir nom

Generar text alternatiu i peu de foto

Text alternatiu

Text alternatiu generat...

Peu de foto

Subtítol generat...

Començar amb una nova notícia

Generador de text alternatiu i peu de foto

Text de notícies

venresday
Carmen y María 23.45 h

Placa Major de Nou Barris
venresday
Simfonia Vozes i Maria Hinojosa 19.30 h

dimarts 24 de setembre
venresday
Jardins del Doctor Pla i Armengol
Taller of Power 12.00 h
Taller de Musica Latin Ensemble 13.00 h
Xamfrà 17.00 h

Obtenir noms del text

Antecessa i deixa anar una imatge



Nombre de persones "Desconegudes" a la foto :
Selecciona quantes entrades "Desconegudes" vols afegir al menú desplegable de noms

0 10

Tria els noms tal com apareixen a la imatge, de dalt a baix i d'esquerra a dreta
Selecciona un o més noms del menú desplegable.

Figa Flawas

Afegix noms manualment
Escriu els noms separats per comes

Afegir nom

Generar text alternatiu i peu de foto

Text alternatiu

Descripció de l'imatge: Dos homes es troben a l'openair davant d'un paisatge de tramuntana, posant per a una fotografia. L'home de l'esquerra porta un pañuelo blanc i ulleres de sol, un collaret superposat i una jaqueta beix. Ell té un gran ram de flors, que inclou florets de lila, rosa i blanc, juntament amb verdura. L'home de la dreta també porta un pañuelo blanc i ulleres de sol, i té bigotí. Ambdues persones miren cap a la càmera, amb la tramuntana que esdevé una llum calorosa damunt seu. Aquesta imatge està relacionada amb l'article de notícies sobre els concerts de la Mercè 2024, específicament destacant l'actuació de Figa Flawas.

Peu de foto

Homes amb flors a l'openair de la Mercè 2024

Començar amb una nova notícia

7. Si tens més fotografies per a la mateixa notícia, pots eliminar la foto actual i continuar des del pas 2. Si no, pots esborrar tots els camps i començar amb una nova notícia i una nova foto fent clic al botó situat a la part inferior de la pàgina, **"Començar amb nova notícia"**.

Nota: Es guarda la traducció a l'anglès de la notícia, fet que, si s'utilitza la mateixa notícia, fa que el procés sigui molt més ràpid.

Avaluació Final d'Usuaris



Per tal avaluar la metodologia desenvolupada i l'eficàcia del sistema s'han dut a terme entrevistes i qüestionaris amb diferents perfils professionals, tan sobre els texts alternatius com els peus de foto.

El focus aquí és la informació generada i l'adequació al context de la notícia i el llibre d'estil. Es van observar alguns errors de traducció, que no formen part central de l'objectiu aquest estudi i per tant seva rellevància aquí es va relativitzar. Veure més detalls a la secció 5.3.1.

Els resultats es mostren en forma de mitjana (línia vertical de punts vermell) i distribució de la desviació al voltant d'aquesta (barres horitzontals del colors en quartils). Els resultats se agrupen d'una banda en àrees d'empresa i també per notícia per tal d'observar la distribució en funció del perfil professional i del tipus de contingut.

Degut als reptes introduïts pel procés de traducció entre el català i l'anglès, amb la generació de barbarismes o errors gramaticals, els textos generats van passar per un procés de correcció supervisada per tal de no incloure baixos innecessaris a les valoracions més enllà de la qualitat semàntica del contingut generat.

Exemple generat correctament de text alternatiu i peu de foto en una notícia amb bon tractament del llibre d'estil. També s'observen un parell de errors gramaticals i de traducció.

	TEXT ALTERNATIU Un grup de menors emigrants sense acompanyament s'asseuen en fila, vestits amb manta vermella i expressions de preocupació. Un jove nen, vestit amb una samarreta esportiva de verd, és assistit per un voluntari que du guants i una camisa blanca amb creus vermelles. El voluntari es troba agenollat al costat del nen, examinant-li les mans. A l'esquena, es veuen més voluntaris i menors, indicant una escena busy i sobrepoblada. La escena reflecteix el context de l'article de notícies, mostrant el moment immediat després de l'arribada de menors emigrants als Països Catalans.
PEU DE FOTO Menors emigrants rebent assistència després de l'arribada.	

Exemple generat incorrectament de text alternatiu i peu de foto en una notícia esbiaixat principalment pel context de la notícia. També s'observen una imprecisió en la disposició de les persones a la imatge.

	TEXT ALTERNATIU Aquesta imatge mostra un grup de quatre persones en una habitació que sembla una instal·lació sanitària o una oficina. A l'esquerra, hi ha una persona de peu, probablement un professional sanitari o un representant, portant una vesta negra amb una corda al coll, el que suggereix el seu rol oficial. A banda, hi ha tres dones velles. Dues d'elles estan assegudes en dues cadires de fusta, i una d'elles està asseguda en una butaca de rodes. La dona de la butaca de rodes té cabell gris curt i està vestida amb un vestit fosc, mentre la dona asseguda a la cadira més a l'esquerra té cabell curt i frisat i està vestida amb una blusa amb patronat. La dona del mig, també asseguda en una cadira de fusta, té cabell gris curt i està vestida amb una vesta fosca. Les tres dones seem que estan participant en una discussió o escoltant atentament la persona de peu. L'atmosfera global suggereix una reunió formal o una consulta, probablement relacionada amb l'acord recent de reforma de les pensions als Països Catalans, on els representants aborden les necessitats i els drets dels ciutadans ancians.
PEU DE FOTO Reunió sobre Reforma de Pensions amb Ancians.	

8.1. Importància del context

El primer pas és la validació de la metodologia desenvolupada, en particular la influència del text de la notícia i la identificació de persones en la generació del text alternatiu.

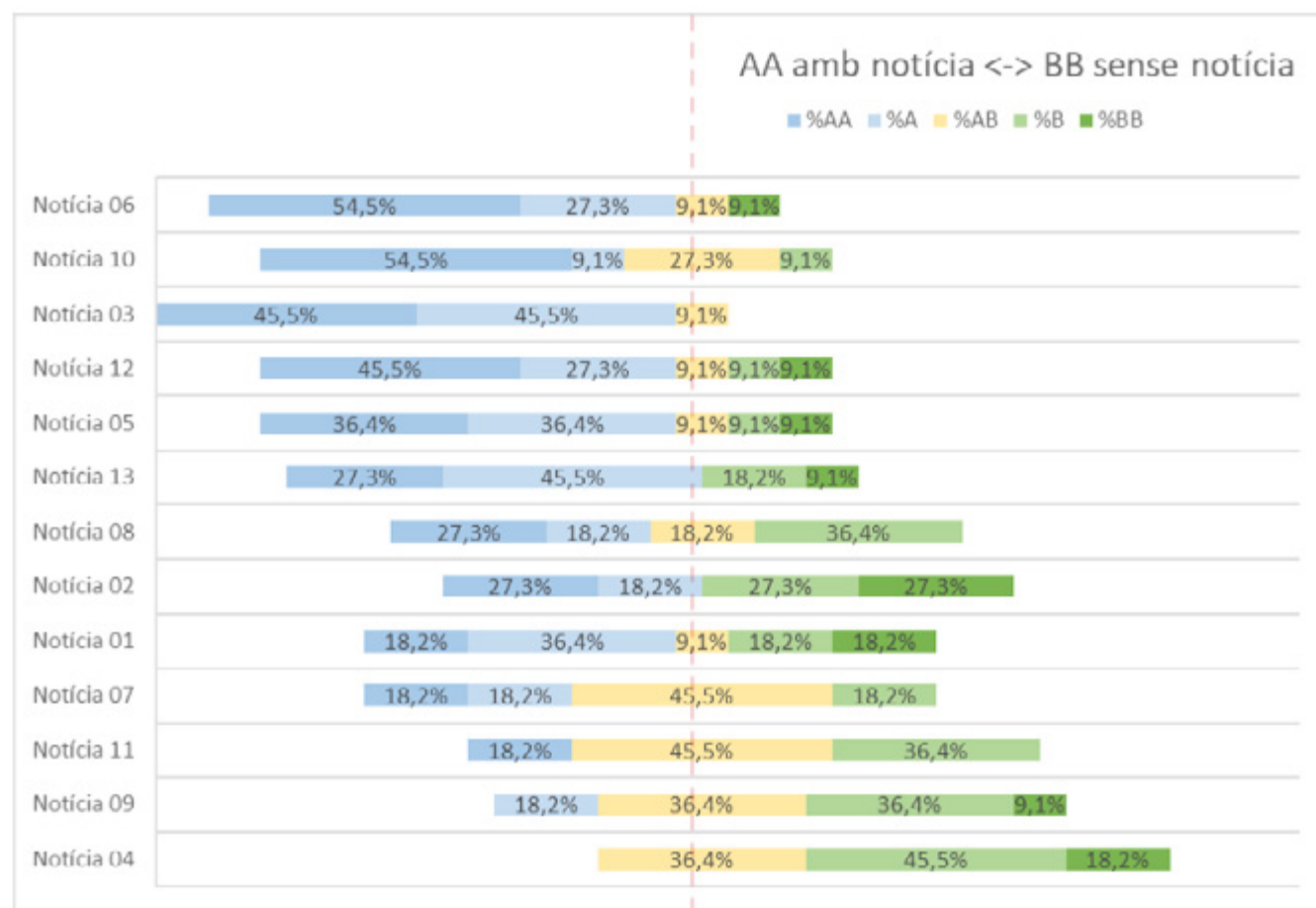
Per això, es va dissenyar una enquesta en què es presenten diverses notícies amb una imatge associada i dues versions del text generat: amb i sense informar el text de la notícia, es a dir, en el segon cas el camp d'inserció de la notícia en la interfície queda buit i per tant el model generatiu no més es basa en el que veu a la imatge.

El perfils professionals per aquesta part de l'avaluació van ser: 8 redactors i editors de notícies, 1 persona responsable d'accessibilitat i 2 usuaris final.

Els participants han comparat les dues opcions i han valorat el grau de preferència seguint l'escala següent:

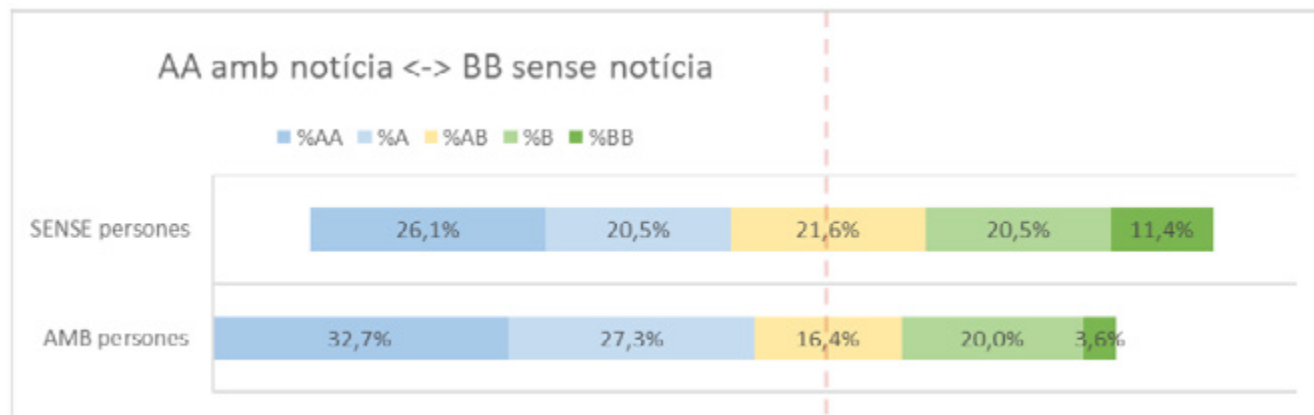
- **AA** > El text A és molt més adequat que B
- **A** > El text A és una mica més adequat que B
- **A-B** > Els textos A i B són igual d'adequats
- **B** > El text B és una mica més adequat que A
- **BB** > El text B és molt més adequat que A

S'ha de considerar que les imatges es troben a l'encapçalament de cada notícia i els textos generats poden contenir errors. L'objectiu de l'enquesta és comparar quina de les dues opcions és més adequada.

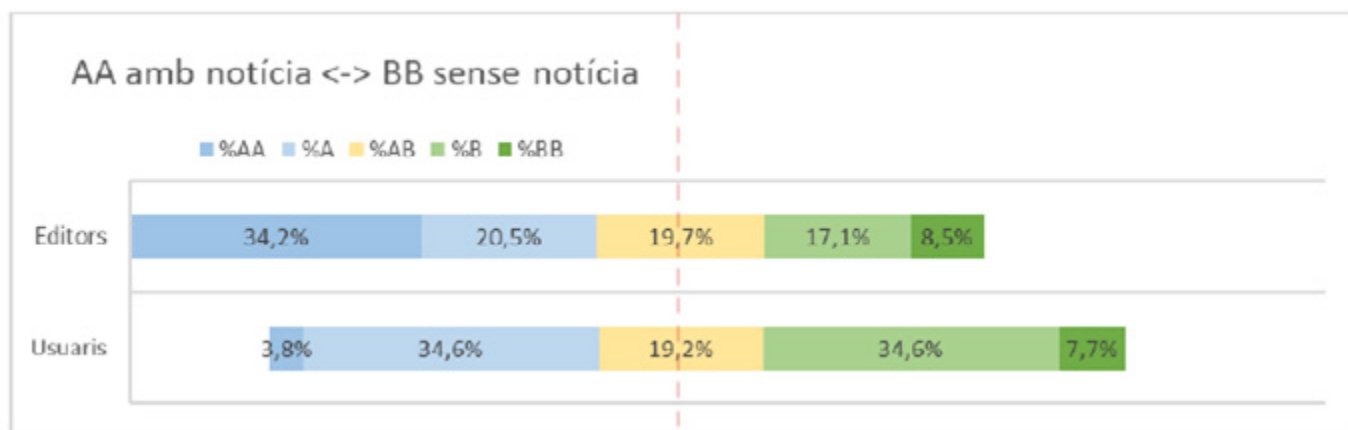


En la gràfica a dalt podem veure que hi ha una lleugera preferència pels texts alternatius generats tenint en

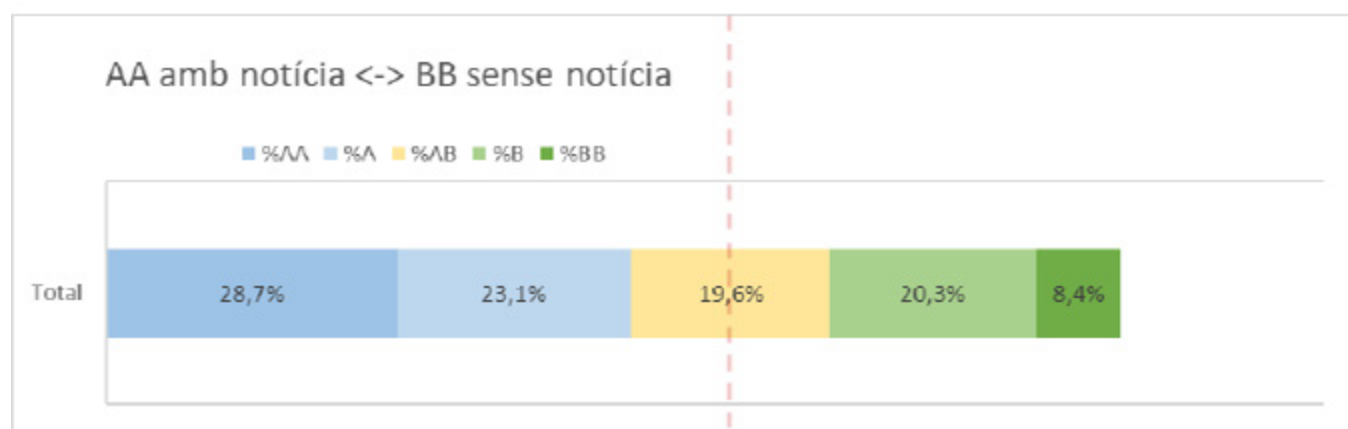
compte el text de la notícia.



Per les imatges que no contenen persones hi ha preferència a utilitzar el text de la notícia. Respecte les imatges que contenen persones, podem veure que hi ha una predilecció clara pels texts alternatius generats amb la notícia informada.



Agrupats per perfils professionals, els editors tenen una clara preferència pels textos alternatius generats amb el context de la notícia informada, mentre que pels usuaris finals no es veu cap preferència clara per cap de les dues alternatives.



De forma generalitzada, veiem que hi ha una lleugera preferència dels text alternatius generats tenint en compte la notícia. A més, la presència de persones a les imatges incrementa la preferència cap a imatges amb notícia informada. Finalment, els Editors prefereixen la notícia informada mentre que pels usuaris finals no hi ha una clara preferència.

La preferència general pels textos generats amb el context de la notícia no és d'estranyar, ja que, en molts casos, com més context tingui un model, millor coneixement tindrà de l'essència de la tasca que li permetria ser més concret i actuar d'una manera més adequada.

En alguns casos, només el context informatiu exacte de la notícia ja pot ser útil per generar una descripció significativa (per exemple, notícies que mencionen persones específiques que el model no pot reconèixer immediatament). Aquest és un detall fàcil de detectar tant pels usuaris com pels editors.

En altres situacions, el context de la notícia proporciona al model els detalls necessaris per enfocar l'atenció cap a un element concret i donar-li més importància en el text generat. En aquest cas, el grup d'editors és més propens a detectar aquests detalls, ja que tenen una millor comprensió de la quantitat necessària de context i, potencialment, també de com s'ha de presentar.

Aquest resultat ja d'entrada validen en gran mesura la metodologia multimodal seguida en aquest projecte.

8.2. Avaluació de la utilitat

Per tal d'avaluar la utilitat de l'eina desenvolupada, s'han dut a terme entrevistes personalitzades on s'han presentat els resultats i s'ha recollit puntuacions amb una escala Likert ¹ sobre cinc característiques clau: *funcionalitat*, *fiabilitat*, *facilitat d'ús*, *personalització* i *rendiment*.

- Per mesurar FUNCIONALITAT es van mostrar exemples concrets amb una fotografia, la notícia associada i el text alternatiu generat. Els participants van avaluar si "El text alternatiu/peu de foto generat descriu la part més important de la imatge tenint en compte el context".
- De manera similar, per mesurar la FIABILITAT, els entrevistats van avaluar si "El text alternatiu/peu de foto es podria publicar sense intervenció humana prèvia".
- Pel que fa a la PERSONALITZACIÓ, els usuaris van analitzar la interfície i van valorar "L'eina proveeix de totes les funcionalitats per poder aconseguir un resultat totalment adaptat a les meves necessitats".
- La FACILITAT d'ús es va mesurar mitjançant l'experiència directa amb la interfície sobre la informació "L'eina té una interfície intuïtiva i fàcil d'utilitzar".
- Per avaluar el RENDIMENT es va analitzar el temps de generació del text alternatiu i peu de foto, considerant si "la generació del text alternatiu i peu de foto en 1 minut és un temps totalment acceptable", sent 1 min un temps de referència possiblement obtingut amb un sistema més òptim i millor capacitat computacional.

Per aquesta part de l'avaluació es van escollir 4 persones amb 4 perfils professionals: *Informatius*, *Accessibilitat*, *RSC* i *Documentació*.

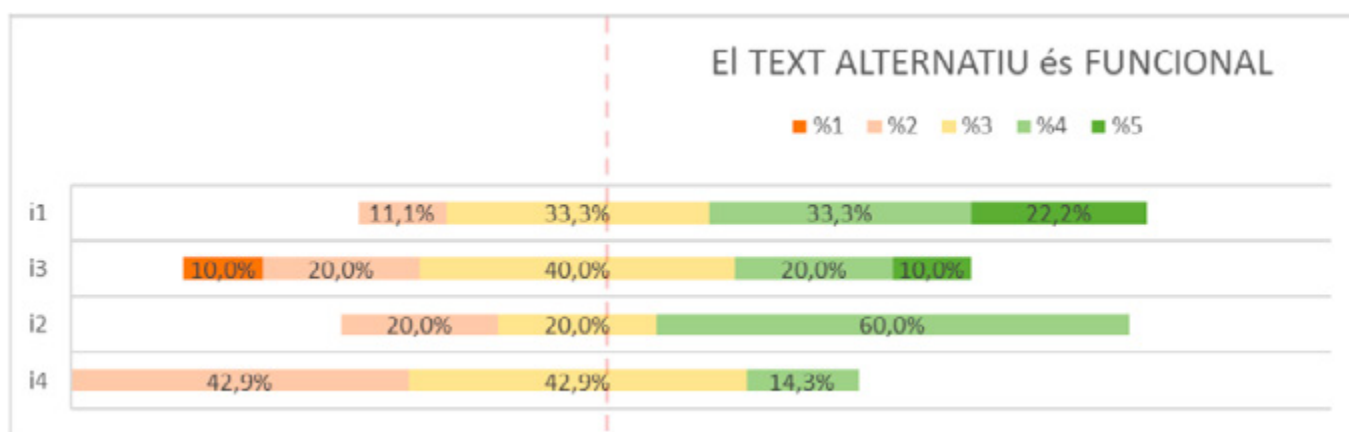
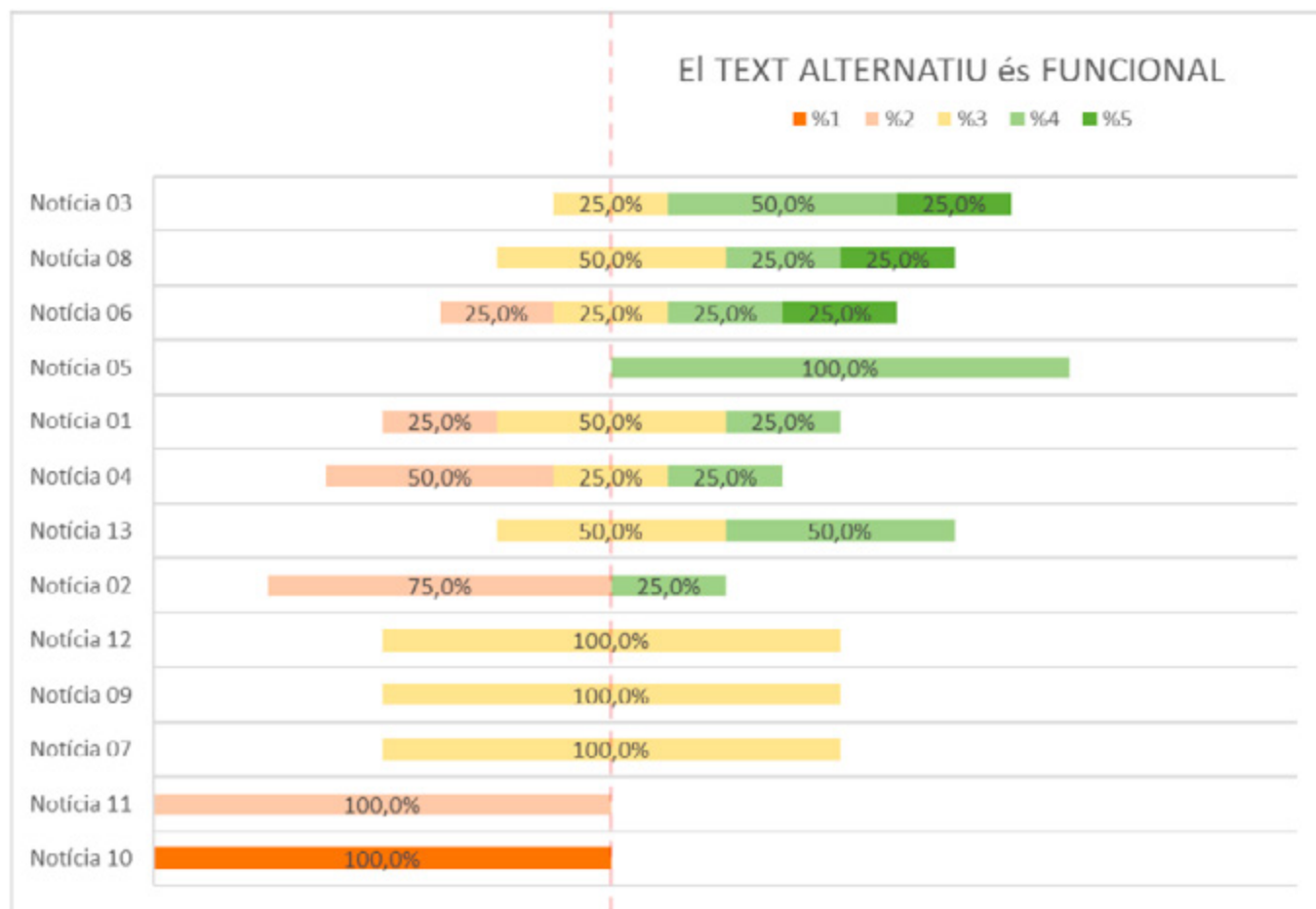
8.2.1. Rellevància de la supervisió humana

Es important notar que en base a la política interna de 3Cat, per garantir el màxim nivell de qualitat, els diferents perfils involucrats han de validar el text escrit abans de ser publicats, tant dels textos alternatius com els peus de foto.

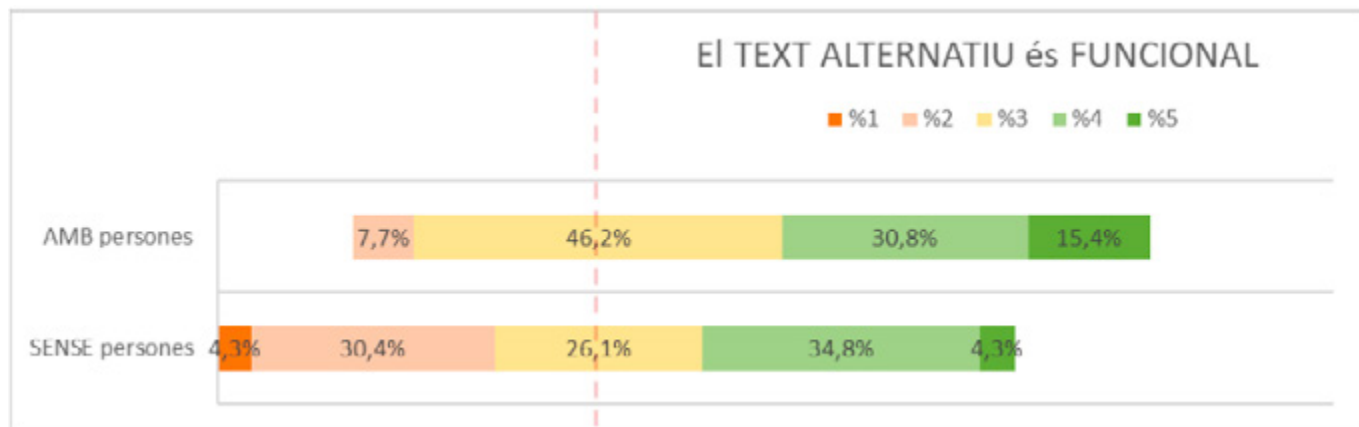
La intervenció humana és considera imprescindible independent de la eficàcia del sistema. En aquest sentit, l'objectiu és reduir el temps de supervisió humana al mínim per tal de que el professional pugui optimitzar la seva dedicació a tasques de major valor afegit.

¹ https://ca.wikipedia.org/wiki/Escala_Likert

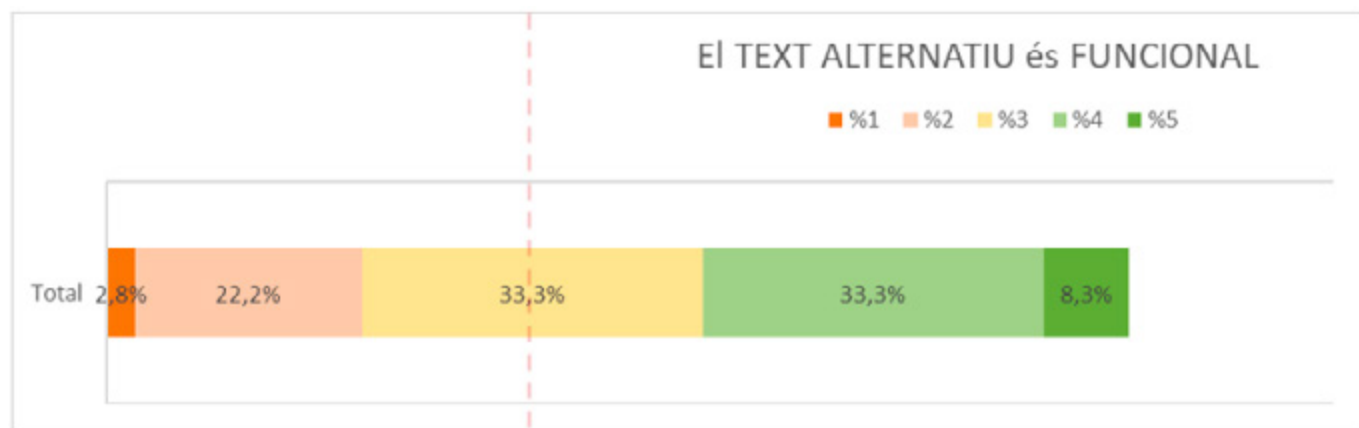
8.2.2. Funcionalitat del text alternatiu



Un 40% dels text alternatius generats son considerats funcionals, un 23% son clarament no funcionals, i la resta 37% no esta clara del tot si descriuen correctament la imatge. Entre els perfils professionals, hi ha 2 que observen en general una funcionalitat del text alternatius, 1 és neutral i un quart clarament no troba funcionals els text generats.

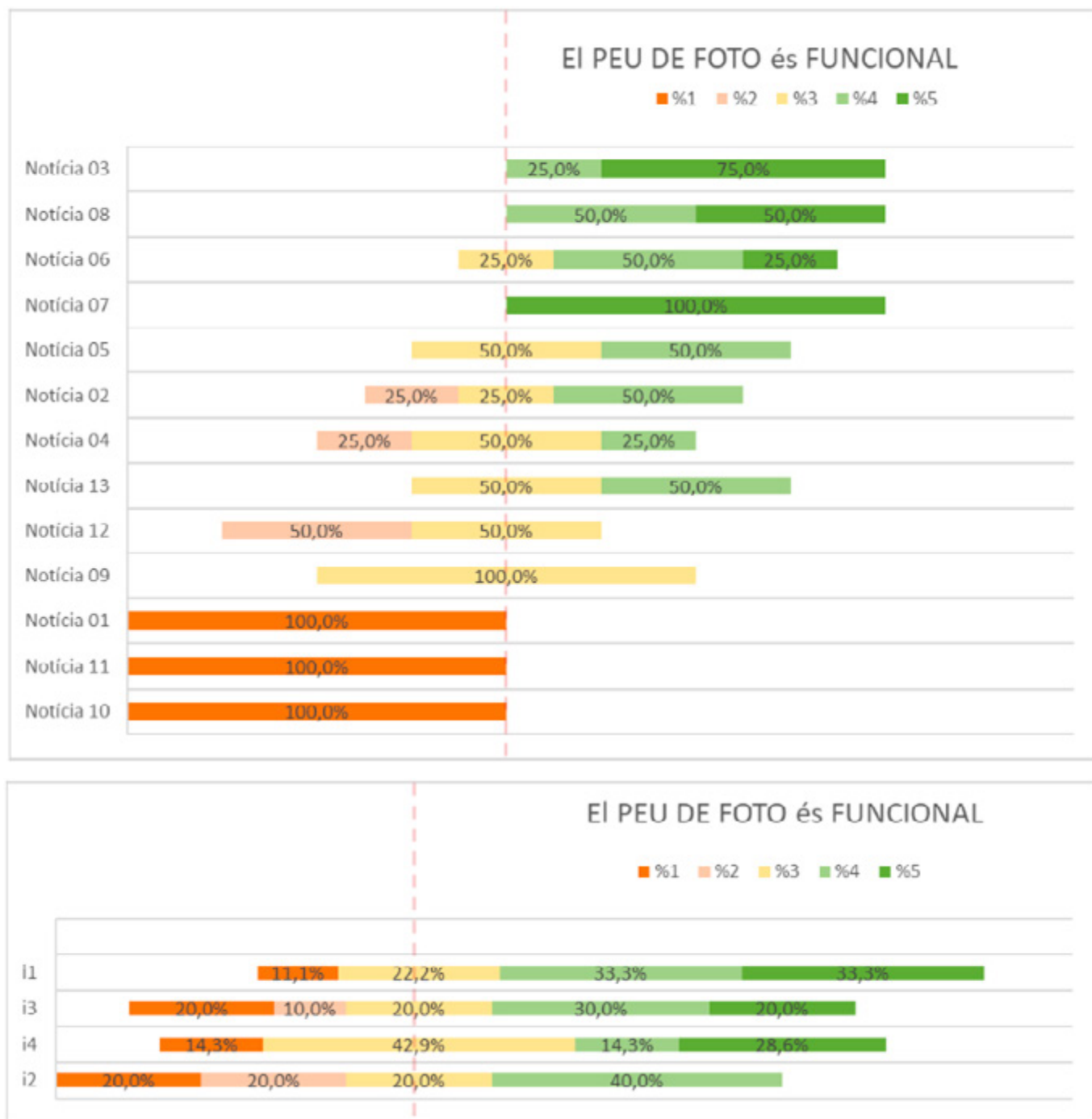


La generació de text alternatiu és valorat més funcional si les imatges contenen persones.



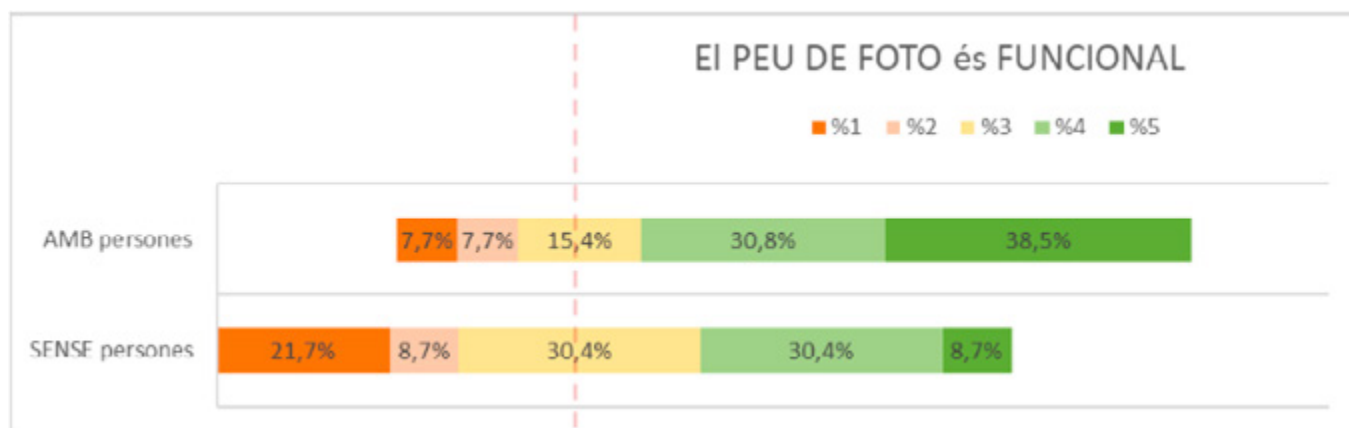
En termes globals, les respostes son més favorables a la funcionalitat del text alternatiu. Cal tenir en compte, però, que hi ha un cert grau de polarització entre els resultats. Alguns considerats poc o gens funcionals i d'altres considerats molt funcionals.

8.2.3. Funcionalitat del peu de foto

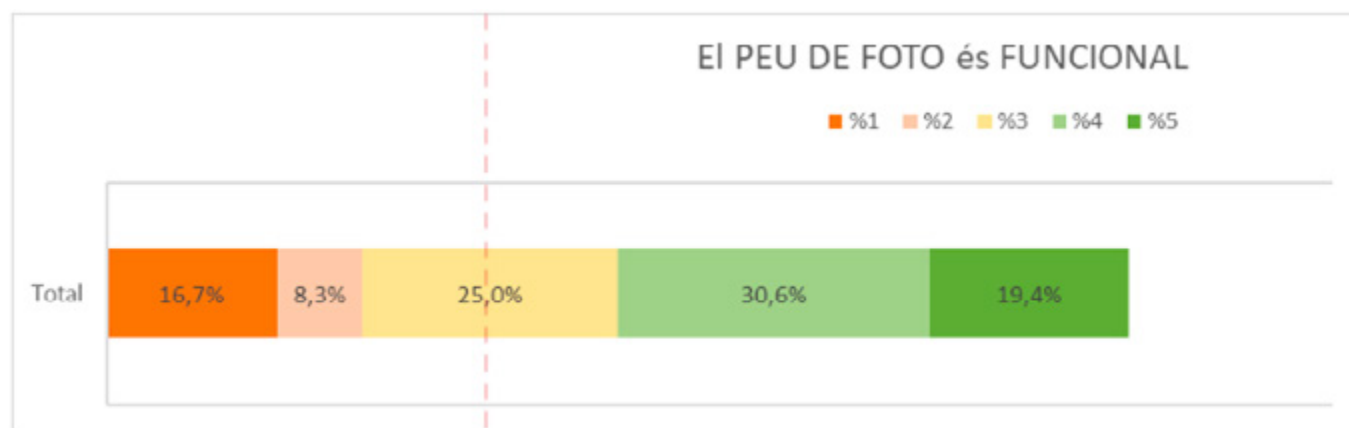


Un 54% dels peus de foto generats son considerats funcionals, un 30% son clarament no funcionals, i la resta 16% no esta clara del tot si descriuen correctament la imatge. Es podria dir que les opinions estan més polaritzades respecte dels text alternatius.

Entre els perfils professionals en general observen a major o menor intensitat la funcionalitat dels peus de foto generats de manera positiva.

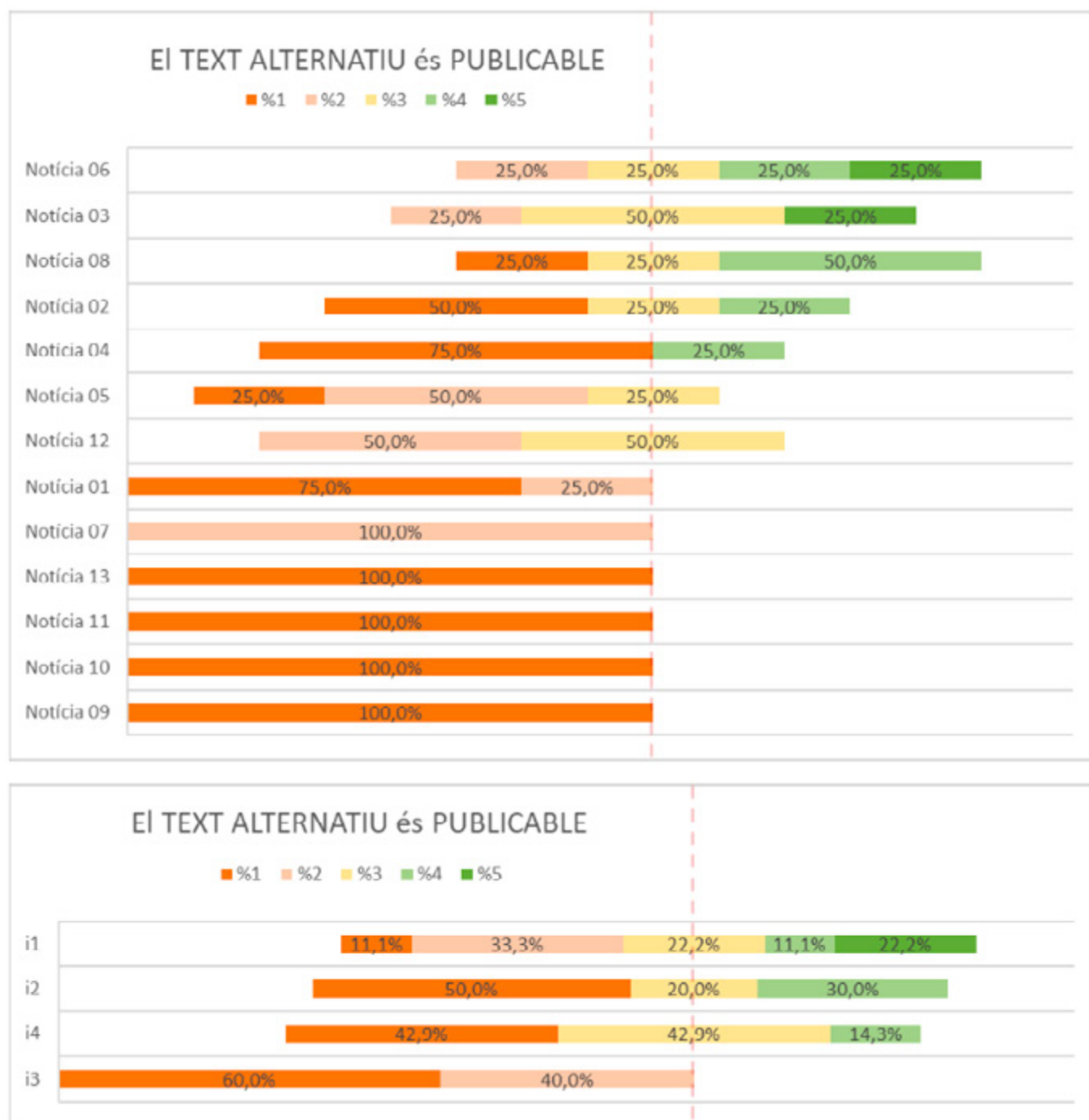


Quan es té en consideració si la imatge conté o no persones, es detecta percepció significativa de millora de la funcionalitat dels peus de foto, especialment en el cas de persones presents a la imatge.

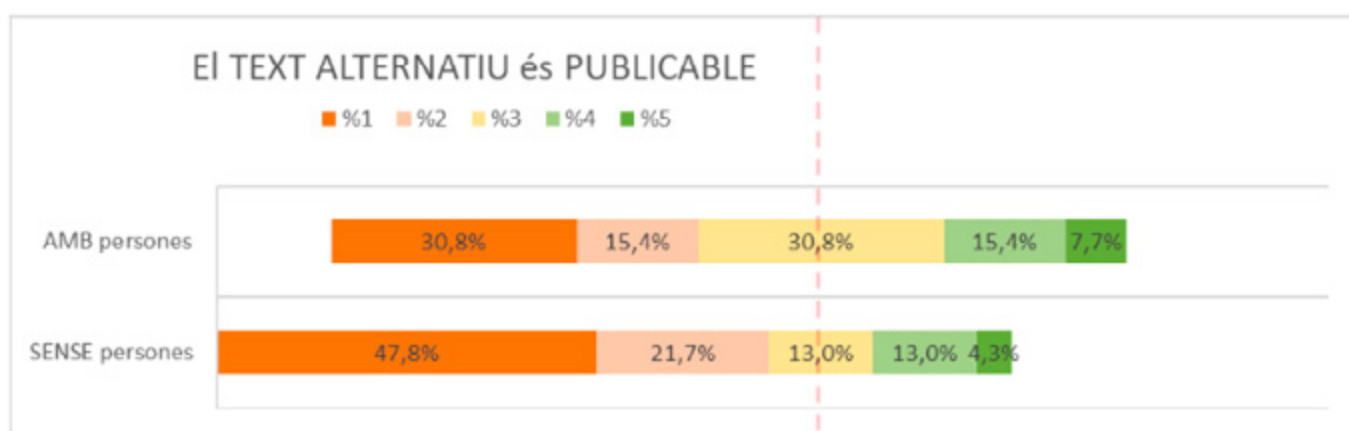
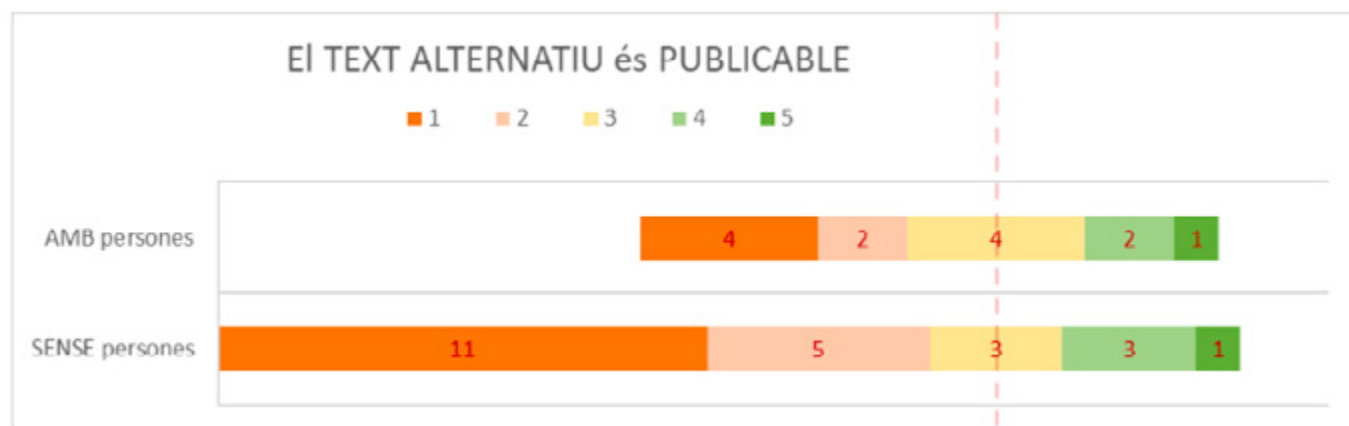


Globalment, els peus de foto generats es consideren majoritàriament funcionals. La consideració de funcional millora quan les imatges de la notícia contenen persones.

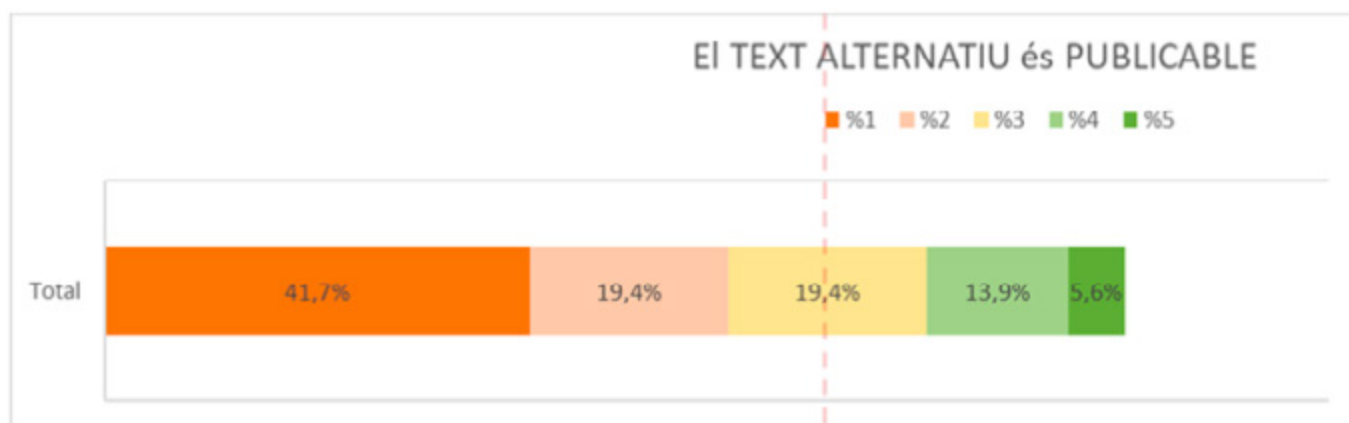
8.2.4. Fiabilitat del text alternatiu



Només un 15% dels peus de foto generats son considerats fiables, en el sentit de poder ser publicats amb mínima intervenció humana. El 77% son clarament no fiables. Entre els perfils professionals 3 observen baixa fiabilitat, especialment un d'ells.

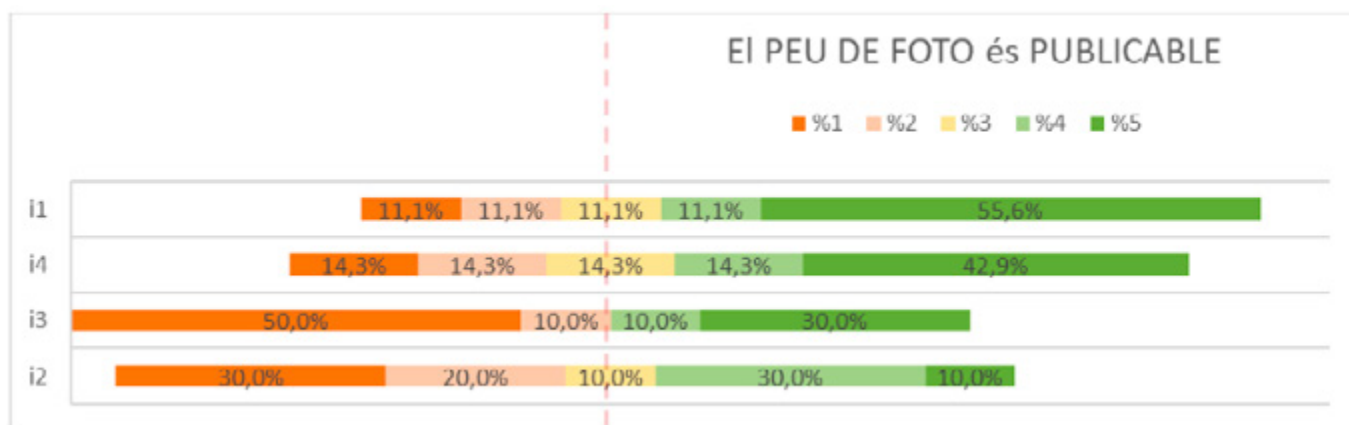
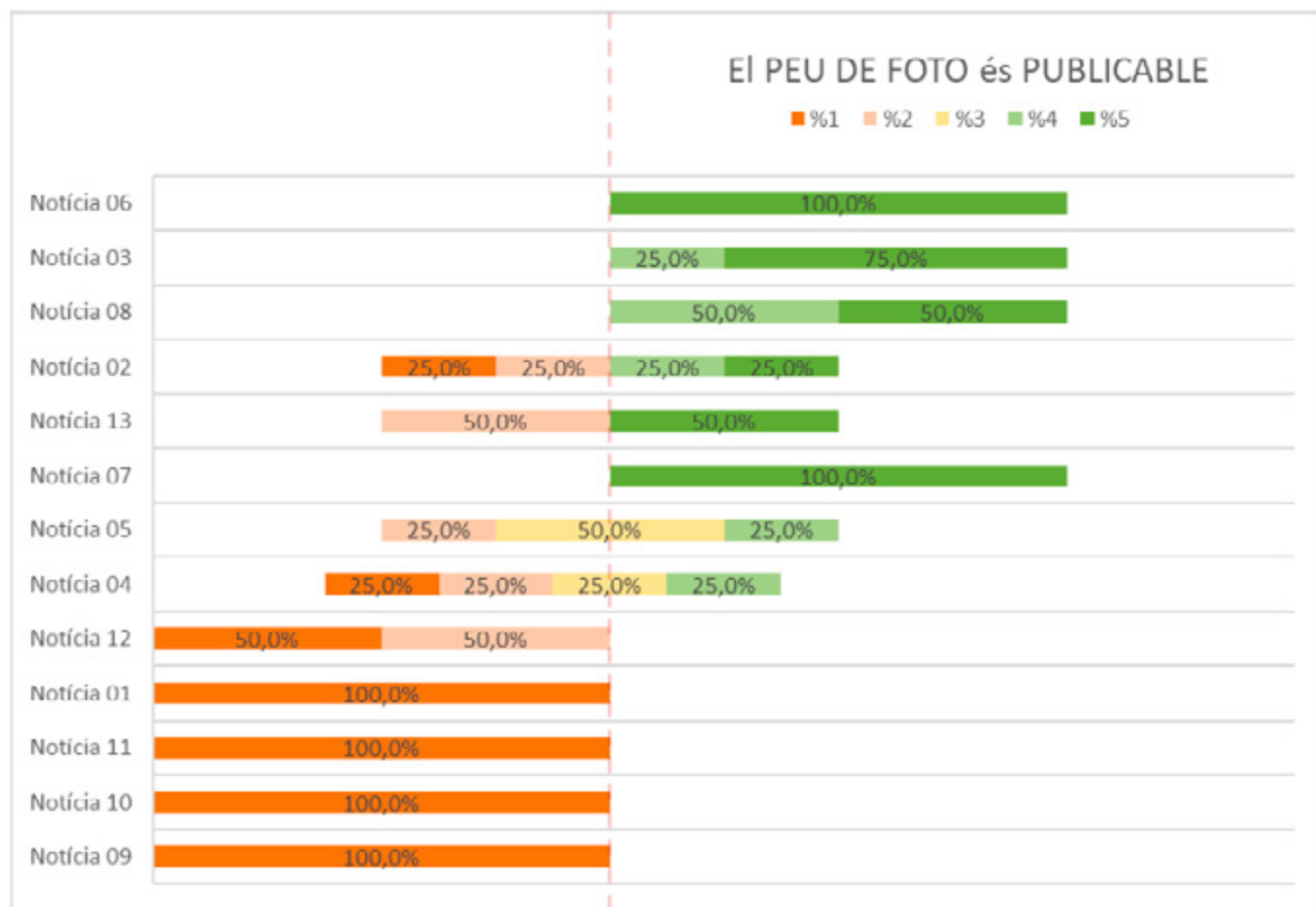


El resultats són homogenis, independentment que la generació sigui en base a imatges amb o sense persones. Es dos casos mostren majoritàriament que el text alternatiu no és publicable amb mínima intervenció humana, especialment en el cas sense persones.



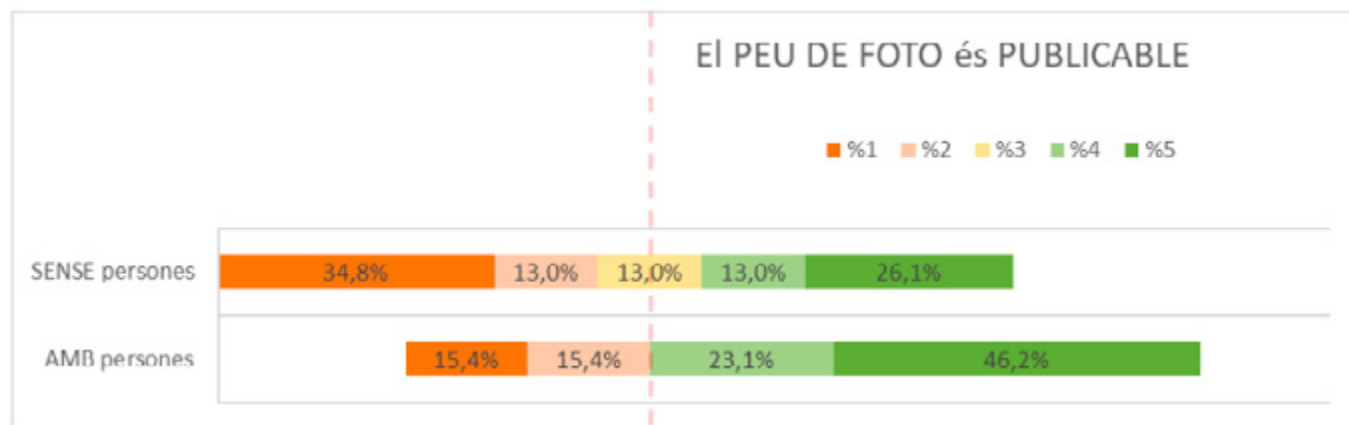
De manera general, el text alternatiu es considera, no publicable amb mínima intervenció humana i el fet que hi hagi persones a la imatge no afecta gaire el resultat.

8.2.5. Fiabilitat dels peus de foto

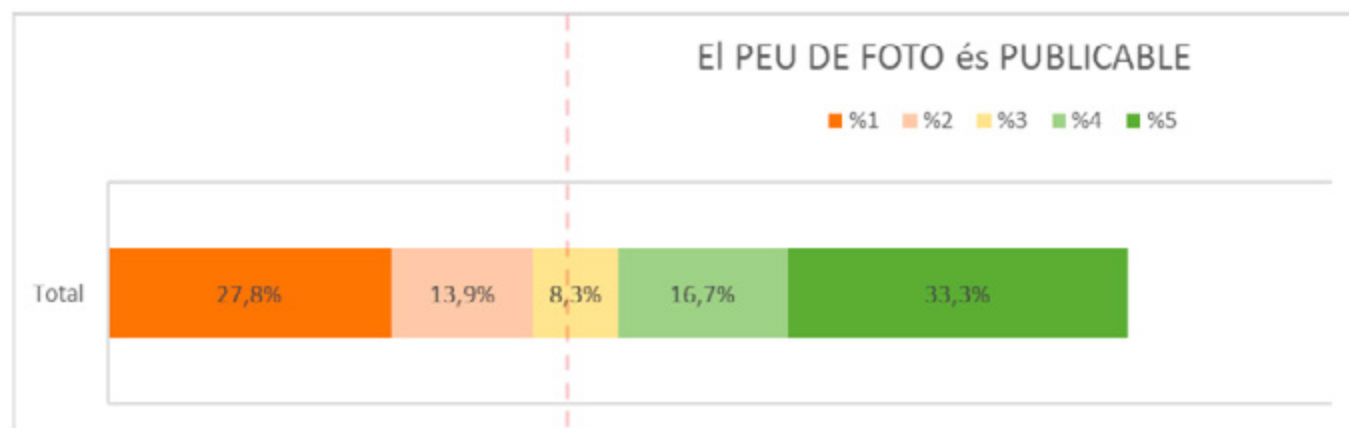


Un 30% dels peus de foto generats son considerats fiables de publicar amb mínima intervenció humana, un 46% son clarament no fiables, i la resta 23% (3 casos) no esta clara del tot si descriuen correctament la imatge.

Entre els perfils professionals, 2 observen bona fiabilitat en general, mentre que 1 d'ells regular i l'altre lleugerament negatiu.



El peu de foto generat es considera més publicable quan les imatges contenen persones.



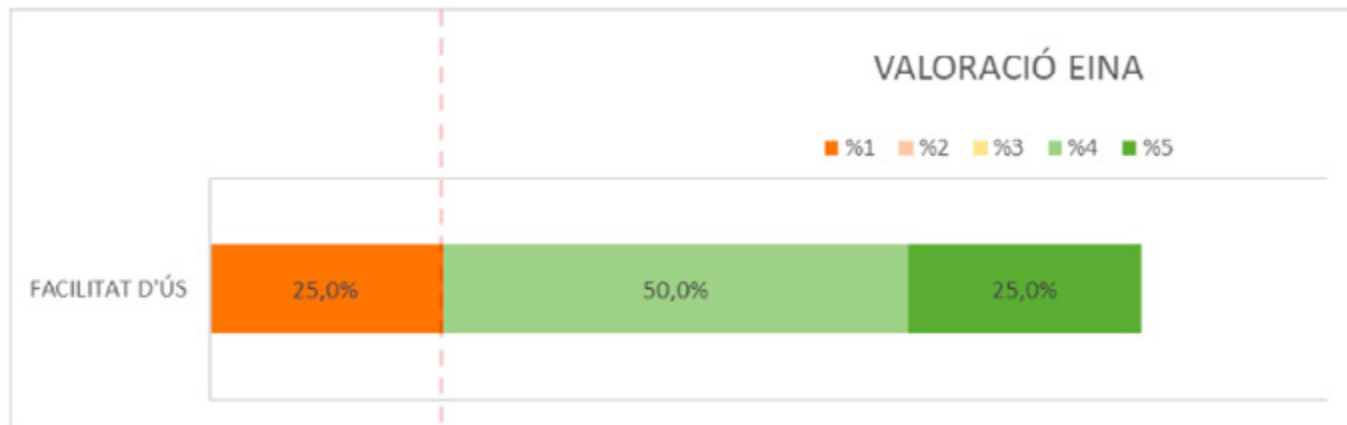
En general, s'observa una lleugera inclinació a considerar els peu de foto publicable amb mínima intervenció humana, tot i que no de manera significativa ni suficient. Hi ha un bloc de notícies amb peus de foto que es consideren clarament publicables, mentre que n'hi ha un altre que es considerat totalment al contrari. De nou, quan les imatges contenen persones les respostes tornen a millorar.

Aquí podem observar, si comparem resultats dels apartats anteriors, que en general els peus foto tenen un tendència a ser valorats més positivament que els text alternatius per ser publicats amb mínima supervisió humana. Possiblement degut a que la informació necessària per generar un peu de foto es més reduïda i és més probable cometre algun error. Tot i així, les opinions mostren que els text generats en general estan lluny de la qualitat necessària per ser directament publicats.

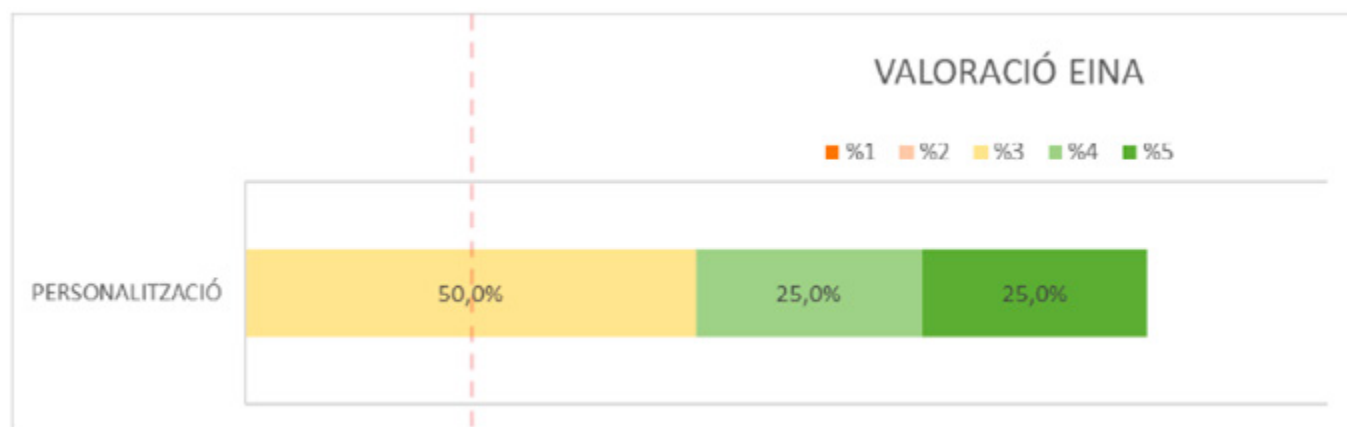
Cal esmentar que al llarg de les 13 notícies no és va observar una tendència clara en el rendiment en funció de la temàtica ni el tipus de contingut de la notícia.

8.2.6. Facilitat d'ús

L'opinió general és que la interfície es pot utilitzar amb 3 de 4 participants positius al respecte.

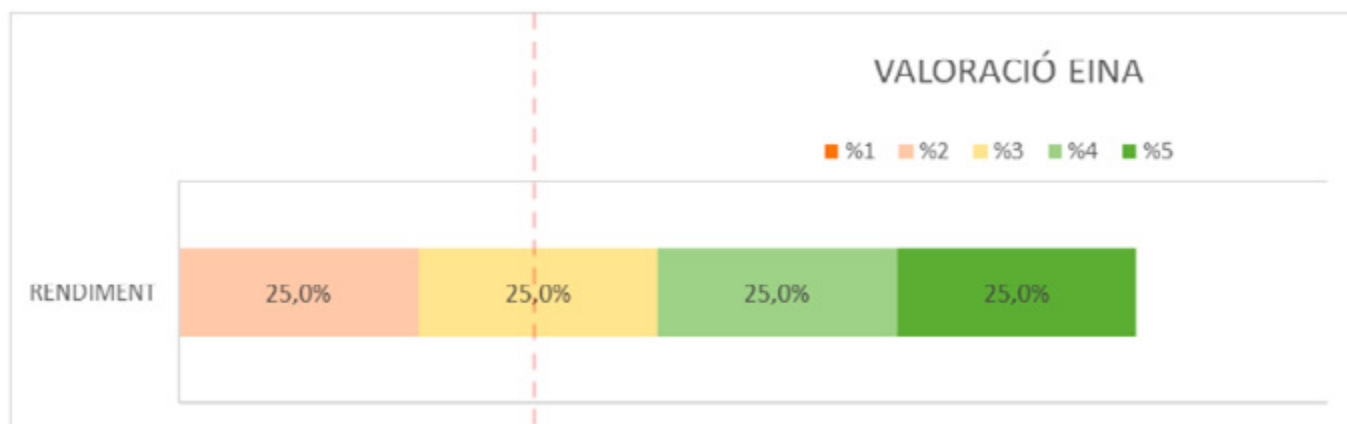


8.2.7. Personalització



Mentre cap dels participants va expressar opinions negatives, la meitat es va mantenir neutral pel que fa a la personalització de l'eina. Durant les sessions de pluja d'idees, van sorgir diverses propostes per millorar les pràctiques de disseny i les solucions. No obstant això, moltes d'elles van haver de ser significativament simplifiades o completament reformulades a causa de les limitacions de Gradio en el disseny d'elements d'interfície.

8.2.8. Rendiment



La meitat va considerar positivament el rendiment de l'eina. L'altre meitat va tenir una opinió neutral. Una possible raó són els diferents orígens dels entrevistats i els seus requisits individuals per a aquesta eina. Un altre factor que contribueix podria ser la presència d'eines comercials com ChatGPT, que pot dificultar als usuaris avaluar el rendiment d'una eina de codi obert i comprendre el temps d'espera esperat per a la generació de text.

8.3. Avaluació global per perfils professionals

8.3.1. Redactors

El general, els redactors han coincidit en aquests comentaris:

- Els textos generats són més llargs de lo necessari: son freqüents els detalls superflus que, més que millorar la descripció de la imatge, poden dur a confusió.
- Caldria centrar el focus de la descripció: utilitzar la notícia per a establir el context adequat, i descriure les parts importants de la imatge en funció d'aquest context. Però no utilitzar la notícia per a ampliar la descripció amb suposicions, que poden arribar a ser errònies.
- L'aplicació del llibre d'estil sembla que porta a interpretacions o suposicions incorrectes, especialment en l'àmbit geogràfic: a molts textos es fa referència al Països Catalans quan a la imatge no hi ha res que hi tingui relació.
- Es destaca la possible utilitat d'enriquir imatges i desar-ne el resultat, per tal de facilitar la feina a l'hora de fer cerques al fons documental de cara a la il·lustració de futures notícies. En aquest cas, es manté la importància de saber detectar el focus de la imatge, com ara la detecció de l'acció que s'està mostrant o el lloc on es produeix, més enllà de multitud de detalls no significatius pel context.

8.3.2. Usuaris finals

Pels usuaris finals, els comentaris s'han mantingut en una línia similar a la dels redactors:

- Les descripcions haurien de limitar-se a la descripció del que es veu a la imatge: més curtes i sense repetir el que ja s'explica a la notícia.
- Hi ha parts de text que resulten de deduccions errònies i que porten a confusió, especialment pel que fa a les localitzacions.

Conclusions



En els últims anys hem assistit a una revolució dels models de llenguatge natural, i en particular el models multimodals, que són capaços de resumir un text i a la vegada descriure imatges amb un gran nivell de complexitat i detall.

La fusió de *prompt engineering* i el raonament d'agents d'IA desbloqueja nous potencials en la descripció de la imatge. En estructurar el procés al voltant de la percepció, la inferència del context, la memòria, la recuperació de coneixement, el raonament i la iteració, els sistemes poden generar descripcions riques i significatives que van més enllà del simple reconeixement d'objectes.

En aquest context, la generació de text alternatiu no només ha de ser vista com una tasca tècnica, sinó com una part integral del procés periodístic. Tant el text alternatiu com els peus de foto proporcionen context addicional i enriqueixen l'experiència de lectura, també a aquelles persones amb algunes de les seves capacitats limitades.

En el cas particular dels textos alternatius, sovint aquesta informació és una còpia exacta del peu de foto o directament es deixa buit, afectant l'accessibilitat i la comprensió de les imatges per part dels usuaris.

El sistema aquí presentat mostra un 77% de texts alternatius i un 70% de peus de fotos qualificats amb bona o mitjana qualitat. Mentre que si considerem si són directament publicables amb mínima intervenció humana el percentatge baixa a 23% i 54%.

La interfície va rebre valoracions positives de 3 dels 4 participants en termes de facilitat d'ús. No va rebre cap valoració negativa en termes de personalització, encara que la meitat de participants es va mantenir neutrals. La meitat dels avaluadors va valorar positivament el rendiment del sistema, mentre que l'altra meitat va mantenir una opinió neutral.

Els models de llenguatge oberts en català no estan madurs i els traductors no són precisos, encara. Considerem que és qüestió de temps que models petits cada vegada siguin més eficients, degut a millores arquitecturals i a l'increment del corpus d'entrenament, en comparació amb el rendiment actual dels models massius actualment comercialitzats a través de serveis de pagament.

L'ús del llibre d'estil aporta claredat, coherència i ètica per corregir l'ús incorrecte i indegut de certs termes en alguns contextos. D'altra banda el llibre d'estil introdueix biaixos innecessaris, la gestió dels quals necessita un millor control en la instrucció del model en segons quins contextos.

Un context de la notícia massa ampli pot derivar en suposicions errònies sobre la imatge i de vegades amb detalls superflus. Tot i que s'ha observat que l'impacte del context de la notícia va millorar de mitjana la qualitat del text generat, en alguns casos podria aportar excessiva informació. D'una banda, la generació de diferents versions, amb i sense context, amb i sense identificació de persones, podria aportar fàcilment variacions de ràpida supervisió per part del professional. D'altra banda, aquest comportament podria perfeccionar-se de nou instruint millor al model. Així, la generació més estructurada del resum de la notícia utilitzat per generar la descripció de la imatge (p. ex. en què, qui, com, quan, on i perquè), podria millorar la percepció de la qualitat del text.

La metodologia semi-supervisada d'identificació de persones aquí desenvolupada, mitjançant la cerca en el cos de la notícia però també la seva disposició a la imatge, ha mostrat una opinió clarament positiva dels professionals a l'hora de generar textos de major qualitat. Per tant, la intervenció humana en la detecció de persones en fotografies pot millorar significativament la qualitat del resultat final i la percepció de les notícies. Aquesta intervenció és especialment crucial en la generació de text alternatiu.

No obstant, la metodologia de supervisió a l'hora d'indicar l'ordre de les persones podria no ser evident en segons quines disposicions dels personatges a la imatge, per exemple en segons quines perspectives i distàncies a la càmera. En aquest sentit, agents que puguin fer consultes a models específics per reconeixement de cares, com models de similaritat facial en el cas de tenir disponible un conjunt d'imatges de persones rellevants, podrien ajudar amb major eficiència i alleugerir el procés de supervisió.

Les polítiques internes de les organitzacions periodístiques han de garantir que, tot i que es pugui assolir un alt nivell de qualitat amb eines automàtiques, la supervisió humana continua sent necessària per assegurar la precisió i la rellevància de les descripcions en tots els casos. Això és fonamental per mantenir la credibilitat i la confiança dels lectors.

Finalment, l'objectiu principal de la generació de descripcions de fotografies no és tant l'estalvi de temps, sinó la millora de la qualitat de les descripcions de les notícies, allà on actualment no hi és, convertint aquest sistema com part d'un assistent periodístic més integral.

A mesura que els agents d'IA continuen evolucionant, la seva capacitat d'entendre, descriure i interactuar de manera autònoma amb el contingut visual redefinirà el compromís amb la tecnologia tant en les aplicacions quotidianes com en els dominis professionals.

Annexos



10.1. Annex A – Primera Avaluació¹

Noticia	Fotografia	Text Alternatiu	Peu Foto	Avaluació
<u>Una ballarina de 109 anys ensenya l'art de la dansa a una nena de 7 (ccma.cat)</u>		Text alternatiu: Eileen Kramer, una ballarina i coreògrafa de 109 anys, es veu donant una classe magistral de ballet a una nena de 7 anys anomenada Violeta. Kramer, asseguda a la seva cama, demostra un moviment de mans, que Violeta, vestida amb un tutú violet, segueix atentament. La escena captura la bellesa i la sensibilitat de la seva interacció, destacant la passió endurera de Kramer per la dansa i la seva influència continuada com a artista.	Eileen Kramer, 109 anys, ensenya ballet a Violeta, 7 anys. Una classe magistral de bellesa i passió!	Acceptable
<u>“Pagarà per l’error”: Netanyahu amenaça l’Iran, que avisa que respondrà si Israel contraataca (3cat.cat)</u>		Text alternatiu: Aquesta imatge mostra Benjamí Netanyahu, el Primer Ministre d'Israel, parlant des d'un podi. Darrere d'ell, hi ha un fons amb la bandera d'Israel i text en hebreu i anglès que diu “ESTAT D'ISRAEL PRIMER MINISTRE”. Netanyahu va vestit amb un traje i corbata i sembla que s'adreça a un públic, probablement discutint el recent conflicte amb l'Iran.	Netanyahu amenaça l'Iran després d'un atac amb més de 100 míssils.	Acceptable Però no compleix llibre d'estil: Tradueix noms propis
<u>Un cotxe cent per cent elèctric és per primer cop el més venut en un mes a Catalunya (3cat.cat)</u>		Text alternatiu: Aquesta imatge mostra un garatge on hi ha diversos vehicles elèctrics carregant-se. Hi ha tres vehicles visibles: un vehicle blanc a l'esquerra, un vehicle negre al centre i un vehicle gris a la dreta. Els vehicles estan connectats a cables blaus que els enllacen a les fonts de càrrega. No hi ha persones visibles a la imatge	Vehicles elèctrics carregant en un garatge, destacant les vendes rècord del Tesla Model 3 a Catalunya.	Acceptable Però especula i excés d'informació que no descriu la imatge

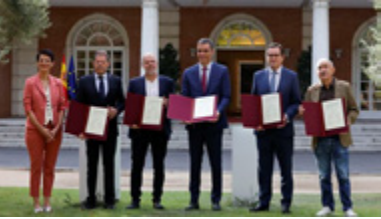
¹ Per als Text Alternatius reproduïts en aquesta i les següents taules s'ha respectat el text generat pel model de llenguatge sense corregir errors ortogràfics ni de terminologia

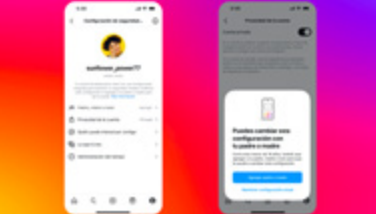
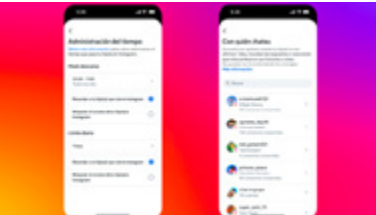
10.2. Annex B – Segona Avaluació

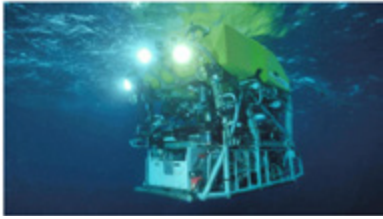
Avaluació del conjunt de notícies amb el text alternatiu i peu de foto generats pel sistema. En vermell i groc es ressalten alguns elements no correctament generats. Veure detalls sobre traducció i la influència del llibre d'estil a la secció 5.3.1.

Notícia	Explosions al Líban: què en treu Israel i què podem esperar de la “nova fase” de la guerra (3cat.cat)		
Fotografia	Text Alternatiu	Peu Foto	Avaluació
	Una persona està manejant un dispositiu electrònic petit amb una pantalla, possiblement un telèfon mòbil o un dispositiu similar. L'individu utilitza ambdós mans, una per agafar el dispositiu i l'altra per interactuar amb la pantalla. La persona porta una pulsera a una mà i un anell a un dit. L'entorn mostra diferents articles electrònics i embalatges, suggestint que el lloc podria ser una botiga o una àrea d'exposició de dispositius electrònics.	Una persona revisa un dispositiu electrònic en una botiga.	Acceptable Però errors gramaticals i especula.
	Un grup de persones, algunes portant barretes, es troben a prop unes de les altres sota una estructura decorada amb flors i teixit verd. Semblen celebrar o participar en un esdeveniment festiu, amb algunes persones sorrint i d'altres que semblen implicades en una conversa. La escena és animada i comunal.	Grup celebrant sota una estructura decorada.	Inacceptable Mala interpretació de les emocions. No ha tingut en compte la notícia, possiblement perquè la notícia no parla directament d'aquest esdeveniment.

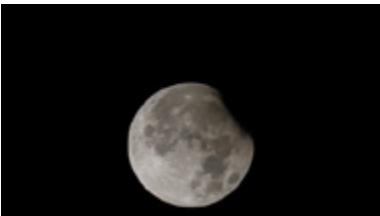
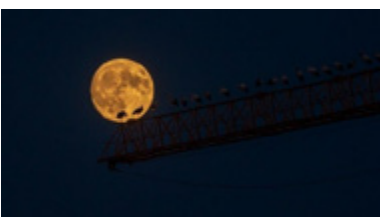
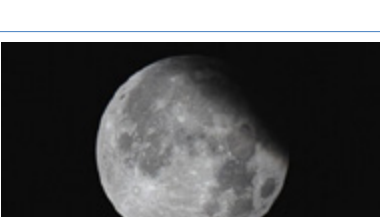
Noticia	<u>La Sagrada Família mostra un detall de la creu que coronarà la basílica i que serà un mirador (3cat.cat)</u>		
	La basílica de la Sagrada Família es mostra amb una grúa groga gran a la dreta i una vista detallada d'un fragment de la creu a l'avantall. El fragment de la creu és part de la torre de Jesús, que està en construcció i serà la més alta dels sis torres centrals. La creu té quatre braços, cadascun mesurant 17 metres d'alt i 13,5 metres d'ample, i servirà com a mirador una vegada completada. La façana intrincada de la basílica, que presenta nombroses agulles i escultures, és visible a l'esquerra sota un cel parcialment nubós.	La Sagrada Família mostra la creu del Jesús amb una grúa.	Acceptable Errors gramaticals i ortogràfics. Excés d'informació que no es descriptiva de la imatge (quant mesura)
	Una vista detallada d'una estructura de creu feta de vidre i ceràmica vidriada blanca, part de la basílica de la Sagrada Família. La creu és un fragment de la torre de Jesús, que serà la més alta dels sis towers centrals i servirà com a mirador. la imatge mostra el disseny intricat i els materials utilitzats, inclosos els trossos de ceràmica corbats i piràmids i diverses capes de vidre corbat i laminat. La creu es construeix amb diverses tècniques i materials especialment dissenyats per a ella.	Detall de la creu del Jesús Tower a la Sagrada Família, feta de vidre i ceràmica.	Acceptable Excés d'informació que no es descriptiva de la imatge.
Noticia	<u>Detingut el suposat conductor de la narcollanxa que va matar dos guàrdies civils a Cadis (3c</u>		
	Una vista aèria d'una persecució a alta velocitat a l'oceà. Un vaixell verd i blanc amb el logotip i el text "Guardia Civil" i "M Interior" està perseguint un vaixell més petit i gris. L'aigua està pertorbada pel moviment dels vaixells, creant onades. La situació suggereix una operació policial, probablement involucrant l'intercepció d'un vaixell sospitos de tràfic de drogues.	La Guardia Civil persegueix una embarcació sospitosa en alta mar.	Acceptable Ha de recordar que està parlant d'una imatge, no una vista aèria. Mala interpretació d'institucions ja que és un vaixell de la guarda civil no un logotip. Excés de detall en l'aigua. Especula.

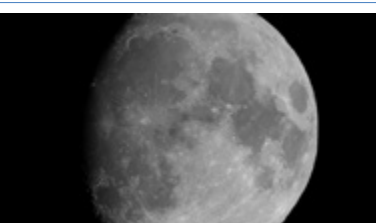
Noticia	Les claus de la reforma de les pensions firmada per govern espanyol, sindicats i patronals (3cat.cat)		
	Un grup de sis persones està davant d'un edifici amb columnes i finestres arcuades. Estan agafant carteres vermelles que contenen documents. d'esquerra a dreta, les persones són: 1. Elma Saiz, Ministra d'Inclusió, Seguretat Social i Migracions, vestida amb una jaqueta rosa i pantalons blancs. 2. Pedro Sánchez, el Primer Ministre espanyol, vestit amb un traje fosc i corbata. 3. Antonio Garamendi, President de CEOE (Confederació Espanyola d'Empreses), amb traje fosc i corbata. 4. Ángel Asensio, President de Cepyme (Confederació d'Empreses Petites i Mitjanes), també amb traje fosc i corbata. 5. Unai Sordo, Secretari General de CCOO (Confederació Sindical de Comissions Obreres), vestit amb traje fosc i corbata. 6. Pepe Reyes, Secretari General d'UGT (Union General de Treballadors), vestit amb traje fosc, corbata i jaqueta marró. El grup ha signat recentment un acord de reforma de pensions. L'edifici darrere d'ells és La Moncloa, la residència oficial del Primer Ministre espanyol. L'acord pretén proporcionar més flexibilitat per rebre pensions mentre es continua treballant, introdueix contractes de transició i inclou altres reformes relacionades amb la jubilació anticipada, les companyies d'assegurança mútua i els treballadors estacionals.	El govern espanyol, empreses i sindicats signen una reforma de pensions per a més flexibilitat laboral.	Inacceptable No interpreta un lloc conegut (La Moncloa) ni els actors (els 6 signants de l'acord). El peu de foto és un segon titular. Problema amb la identificació: president del Govern Espanyol i no Primer Ministre. Excés d'informació no descriptiva de la imatge (Últim paràgraf)
	Tres dones velles es troben en una habitació. dues d'elles estan sentades en cadires, i una està en una butaca de rodes. Semblen estar involucrades en una conversa. Una de les dones a la butaca de rodes està sorrient i mirant cap a la càmera. Una altra dona, de pie a l'esquerra, porta una falda amb patronatge i un passe de coll, possiblement indicant que és una cuidadora o personal del centre. La habitació té una atmosfera calorosa i amable.	Tres dones velles conversant a una habitació. Una a butaca de rodes, dues a cadires. Amb una atmosfera calorosa.	Inacceptable No compleix llibre d'estil: Tractament irrespectuós de les persones grans. Especula Excés de detall.

Notícia	Instagram anuncia un nou compte més segur per a adolescents, però sense verificar l'edat, per Toni Noguera Martínez (3cat.cat)		
	Instagram introdueix "Comptes per a adolescents" amb proteccions integrades per als usuaris més joves. La imatge presenta un fons amb gradació de colors i el logotip d'Instagram al centre. Al voltant del logotip hi ha diferents icones que representen característiques de privacitat i seguretat. El text "Comptes per a adolescents" es troba escrit en un patró circular al voltant del logotip, i una franja blava a la part inferior diu "Proteccions integrades."	Instagram presenta comptes per a adolescents amb proteccions integrades per a una experiència més segura.	Acceptable Excés d'informació que no descriu la imatge en la primera frase.
	Instagram anuncia una nova compte més segura per a adolescents sense verificar l'edat. Les comptes seran privades per defecte i inclouran el mode de son, tot i que aquesta funcionalitat arribarà més tard a Europa. Les noves comptes per a adolescents estan dissenyades per a adolescents i s'activarà automàticament per a tots els usuaris de menys de 17 anys. Aquestes comptes inclouen proteccions com comptes privades per defecte, missatges restringits, contingut sensible limitat, interaccions limitades, recordatoris de temps límit i mode de son habilitat. Els pares tindran noves funcions de control, incloent veure qui els seus fills parlen, establir límits d'ús, i bloquejar l'aplicació en horaris específics. Instagram està treballant en eines de verificació d'edat però no té un mètode establert per verificar l'edat dels usuaris.	Instagram presenta comptes més segurs per a adolescents: privades per defecte i mode de son inclòs. Protecció addicional per a menors.	Inacceptable Excés d'informació que no descriu la imatge, ja que sembla que està fent un resum de la notícia.
	Dos captures de pantalla de les noves característiques d'Instagram per a comptes de joves. La captura de pantalla de l'esquerra mostra la secció 'Administració del compte' amb opcions per al mode de son, recuperació de la compte i límit de temps diari. La captura de pantalla de la dreta mostra la secció 'Xat amb', llistant noms d'usuari i el nombre de seguidors de les persones amb qui l'usuari pot xatejar. L'escenari té un gradació de colors que transita del violet al rosa.	📱 Nova funcionalitat d'Instagram per a joves: comptes privats per defecte, mode son i més! 🗝️😊	Acceptable Tot i que incompleix llibre d'estil ja que els emojis no estan permesos.

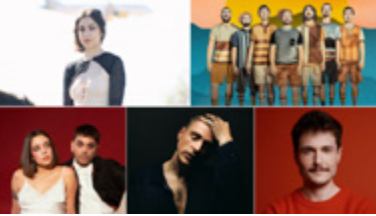
Noticia	El PSOE augmenta la distància sobre el PP, segons el CIS del setembre		
	L'imatge mostra dos homes en un ambient formal, probablement un debat polític o una discussió. A l'esquerra, un home amb cabell fosc, vestint un traje blau i una corbata vermella, sembla parlar o gesticular. A la dreta, un altre home amb cabell gris, vestint un traje fosc i una corbata verda, també parla o gesticula. Els dos homes es troben darrere de púlpits, i hi ha altres persones assegudes a l'esquena, observant la discussió. L'ambient suggereix un lloc formal i oficial, possiblement una cambra governamental o parlamentària.	"Debat polític entre dos líders en una cambra oficial."	Inacceptable Especula i problema amb la identificació: Pedro Sanchez i Alberto Nuñez Feijóo.
Noticia	Incendis al Priorat: controlat el foc de Cabacés (3cat.cat)		
	Un pal elèctric de fusta alt està situat en una zona boscosa sota un cel blau clar. El pal té múltiples traves amb cables elèctrics que s'estenen des d'ell. Al voltant del pal hi ha diferents tipus d'arbres i folre verd dens. La escena és tranquil·la, sense signes visibles de perturbacions o activitat humana.	"Pola elèctrica enmig del bosc sota un cel blau."	Acceptable Problema amb el focus: No ha tingut en compte que el pal té un precinte policial que és el més rellevant
Noticia	Revelen imatges i els últims missatges del submergible Titan: "Tot bé per aquí" (3cat.cat)		
	Imatge submarina d'un submergible robòtic amb llums brillants il·luminant l'oceà fosc. El submergible està equipat amb diversos braços mecànics i instruments, i sembla estar explorant o realitzant recerca. La imatge probablement és presa des d'una càmera adossada al submergible, mostrant l'aigua profunda i blava que l'envolta.	Imatge del submergible robòtic il·luminant l'oceà fosc amb instruments i braços mecànics.	Acceptable

Notícia	La Fiscalia demana aturar el protocol de les Canàries sobre menors migrants per “desempament” (3cat.cat)		
	Un grup de nois joves, tots vestits amb roba taronja brillant, estan sentats al terra en una fila contra una paret de ciment. Semblen estar en un ambient casual, possiblement a l'aire lliure, amb alguns d'ells mirant en direccions diferents. Els nois semblen ser d'edat similar i probablement estan implicats en una activitat relaxada o informal.	Nois vestits amb roba taronja sentats al terra contra una paret de ciment.	Inacceptable Mala interpretació: els nois són menors migrants. Especula amb l'ambient i l'activitat.
	Un grup de persones, incloses migrants i treballadors humanitaris, es troben en un entorn al aire lliure. Molts dels migrants duen mantes vermelles, mentre que alguns treballadors humanitaris duen camises blanques i pantalons naranjas. Un treballador humanitari, portant guants, està assistint un jove migrat que està sentat i sembla rebre atenció mèdica. Hi ha botelles d'aigua i altres subministraments propers. La escena suggereix un esforç humanitari, possiblement després d'una operació de rescat o en un camp de refugiats.	Trabajadores humanitarios atienden a migrantes en operació de rescate.	Acceptable Especula i te varius errors gramaticals, ortogràfics i castellanismes.
Notícia	Llarena acusa els Mossos de no fer prou per detenir Puigdemont i vol que un jutge ho investigui (3cat.cat)		
	Un grup de persones, inclòs un home vestit amb traje i corbata, es troben a prop uns de les altres. L'home amb el traje sembla ser la figura central, amb diverses persones al seu voltant, algunes portant barretes. El grup sembla estar involucrat en una conversa o interacció amb l'home del traje. L'ambient sembla ser a l'aire lliure, possiblement en un esdeveniment o reunió pública.	El magistrat Llarena acusa els Mossos de no haver fet prou per arrestar Puigdemont.	Inacceptable Especula i problema amb la identificació: és el president Carles Puigdemont.
	Un grup d'homes vestits de tàctiques està junts, amb un home al centre que apunta el dit cap amunt. L'home al centre és Carles Puigdemont, l'expresident català. A la seva esquerra hi ha Josep Rull, un antic ministre català. Altres homes del grup inclouen figures polítiques i oficials. Estan de peu a l'aire lliure, amb edificis i arbres visibles al fons. L'atmosfera sembla formal i seriosa.	Carles Puigdemont i Josep Rull amb altres oficials, en un acte formal.	Inacceptable Problema amb la identificació: Parla d'un home i és Carles Puigdemont, tot i que després ho anomena. Les persones que es citen a la notícia i surten a la foto, s'haurien de citar al text alt. I al caption Especula amb l'atmosfera.

Noticia	Així s'ha vist l'eclipsi parcial de superlluna - El temps (3CAT .cat)		
	Aquesta imatge captura una vista d'un eclipse parcial de superlluna. La lluna apareix gran i brillant, amb una part d'ella ombrada per l'ombra de la Terra, creant una àrea fosca distinta a la seva superfície. El cel al voltant de la lluna és fosc, amb somees lleugeres de nius. L'escena general és serena i mostra la bellesa d'aquest esdeveniment astronòmic.	Visió de l'eclipse parcial de superlluna entre nius i cels clars.	Acceptable Tot i que utilitza massa retòrica.
	Aquesta imatge mostra la lluna al cel nocturn durant una eclipsi de superlluna parcial. La lluna apareix com un cercle gran, brillant i de color grisós amb algunes tints més fosques, indicant que l'ombra de la Terra cobreix parcialment la lluna. El fons és completament fosc, destacant la brillantor de la lluna i els detalls de l'eclipsi.	Superlluna parcial: Una magnífica vista de l'eclipsi matinal a Catalunya.	Acceptable
	Aquesta imatge captura una escena nocturna serena que inclou una eclipsi parcial de superlluna. El cel és fosc, creant un contrast marcat amb la lluna brillant, parcialment eclipsada. La lluna apareix com una gran esfera lluminosa amb una part ombrada, indicant l'eclipsi. Al primer pla, hi ha el silueta d'una estructura metàl·lica, possiblement un pont o una torre, amb diversos ocells posats sobre ella. Els ocells estan espaiats de manera uniforme a l'estructura, afegint un sentit de simetria a l'escena. L'atmosfera general és calmada i tranquil·la, destacant la bellesa de l'esdeveniment celeste.	Eclipsi parcial de superlluna amb ocells al llit d'una estructura metàl·lica.	Acceptable Tot i que utilitza massa retòrica i poètic.
	Aquesta imatge mostra una vista parcial de la superlluna durant un eclipsi. La lluna sembla lleugerament més gran del normal a causa de la seva proximitat a la Terra, i una ombra està projectada sobre una part de la seva superfície, indicant l'eclipsi. La lluna està principalment il·luminada, amb només una part en ombra. L'entorn és fosc, destacant la brillantor i els detalls de la lluna.	Superlluna parcial durant un eclipsi, visible des de Catalunya.	Acceptable

Noticia	Eclipsi parcial de superlluna la matinada de dimecres - El temps (3cat.cat)		
	Aquesta imatge mostra una seqüència de sis fotografies de la Lluna davant un fons negre, il·lustrant la progressió d'una eclípsi lunar parcial. Cada lluna apareix lleugerament diferent, demostrant els canvis d'il·luminació durant l'eclípsi. Les llunes es disposen en una línia diagonal des de l'esquerra superior fins al dret inferior, amb la primera lluna sent la més il·luminada i l'última sent la menys il·luminada. La imatge captura l'efecte gradual d'ombra de la Terra sobre la Lluna durant l'eclípsi.	Eclipsi lunar parcial: seqüència de la fase de la Lluna durant l'eclípsi	Acceptable
	La Lluna plena és gran i brillant, parcialment ocultada per l'ombra de la Terra, creant una eclípsi lunar parcial. La Lluna està alta al cel nocturn, envolta de núvols. Per sota de la Lluna, hi ha siluetes d'arbres i una casa. La escena està ambientada de nit, amb el cel il·luminat per la Lluna com a principal font de llum.		Acceptable
	Un cos celeste gran, gris i texturat, la Lluna, es mostra de manera prominente contra un fons negre. La Lluna mostra diferents tons de gris, indicant diferents elevacions i cràters. Aquesta imatge captura la Lluna durant un eclipse lunar parcial, on l'ombra de la Terra cobreix parcialment la superfície de la Lluna.	Lluna parcialment eclipsada el dimecres matí.	Acceptable
	L'imatge mostra una eclípsi lunar parcial, també coneguda com a eclípsi de superlluna parcial. Aquest esdeveniment celeste es produeix quan la Terra es col·loca entre el Sol i la Lluna, fent que la Lluna passi per l'ombra de la Terra. La Lluna apareix parcialment escurida, amb una tonalitat vermella a la part superior, indicant l'ombra parcial que projecta la Terra. Aquesta eclípsi és una de les quatre superllunes de l'any i és visible des de la nostra regió, tot i que les condicions meteorològiques poden afectar-ne la visibilitat.	Eclipsi lunar parcial: la Lluna passa per l'ombra de la Terra, mostrant una tonalitat vermella	Acceptable

Notícia	Les xifres de la regeneració d'aigua: quanta se'n reutilitza i per a què?, per Carla Riverola Brutau																	
	Dos matraces d'aigua col·locades en un rebord davant d'una instal·lació de tractament d'aigües. Una matracca conté un líquid groguenc, mentre que l'altra conté aigua clara.	Mostrant el procés de regeneració d'aigües a una estació de tractament.	Acceptable Amb alguns errors gramaticals.															
	Una gran instal·lació rectangular de regeneració d'aigua amb una barana de metall al llarg del cantó superior. la instal·lació està envolta d'una superfície de ciment amb patrons visibles d'fluix d'aigua. Al fons, hi ha diversos edificis, inclosa una estructura alta i moderna amb una part superior daurada. El cel està cobert de núvols, suggerint un dia nubós.	Estació de regeneració d'aigua amb flux visible i edificis al fons.	Acceptable Amb alguns errors gramaticals.															
 ORIGEN DE L'AIGUA QUE CONSUMIM Font: Aigües de Barcelona <table><thead><tr><th>Any</th><th>Superficial</th><th>Desalinitzada</th><th>Regenerada</th><th>Residuals</th></tr></thead><tbody><tr><td>2021</td><td>63%</td><td>3%</td><td>0%</td><td>34%</td></tr><tr><td>2023</td><td>23%</td><td>33%</td><td>25%</td><td>19%</td></tr></tbody></table>	Any	Superficial	Desalinitzada	Regenerada	Residuals	2021	63%	3%	0%	34%	2023	23%	33%	25%	19%	L'imatge mostra dos diagrames circulars comparant els orígens de l'aigua consumida el 2021 i el 2023. Les categories de fonts d'aigua inclouen superficial, desalinitzada, regenerada i residuals. El 2021, el 63% de l'aigua consumida era superficial, el 34% eren residuals i el 3% era desalinitzada. El 2023, la composició canvia significativament: el 23% és superficial, el 33% és desalinitzada, el 19% són residuals i el 25% és regenerada. la imatge subratlla l'augment en l'ús d'aigua regenerada al llarg dels anys.	L'origen de l'aigua que consumim canvia: més aigua regenerada el 2023.	Acceptable
Any	Superficial	Desalinitzada	Regenerada	Residuals														
2021	63%	3%	0%	34%														
2023	23%	33%	25%	19%														

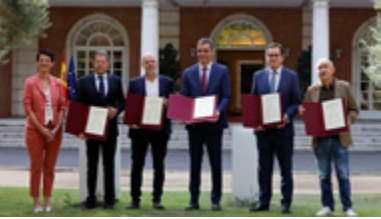
Notícia	Concerts de la Mercè 2024: horaris, escenaris i cartell d'artistes (3cat.cat)		
	Concerts de La Mercè 2024: horaris, escenaris i programa d'artistes. Aquesta imatge inclou fotografies promocionals de diversos artistes que actuaran al festival de La Mercè de Barcelona. La primera fila, de l'esquerra a la dreta, inclou imatges de Figa Flawas, Mama Dousha, Oques Grasses, Maria Jaume, Triquell, Miki Núñez, La Fúmiga i Els Catarres. La segona fila, de l'esquerra a la dreta, inclou imatges de Lua de Santana, Nicole Torrealba, San Isidro, Pinan 450f, Vic Mirallas, Aïta Mon Amour, Banda Municipal de Barcelona i Valentina & The Electric Post. Aquests artistes són part del programa diversificat del festival, que inclou tant artistes consolidades com emergents. El festival presenta una varietat de gèneres musicals i actuacions, mostrant la riquesa cultural de Barcelona i la seva ciutat convidada, Casablanca.	Concerts de La Mercè 2024: horaris, escenaris i artistes destacats. Descobreix el programa musical del festival de Barcelona.	Acceptable
	A la imatge, hi ha dos individus de peus al costat, posant per a la càmera. La persona de l'esquerra és un home amb barba, portant ulleres de sol i una jaqueta de color clar. La persona de la dreta és una dona portant un pañuelo blanc, ulleres de sol i una jaqueta fosca. Els dos porten un ram de flors de diferents colors, inclosos el rosa, el violet i el blanc. L'entorn mostra un paisatge urbà durant el capvespre o l'alba, amb edificis i un cel nubós.	Dos individus posant amb flors i paisatge urbà al fons.	Inacceptable Problema amb la identificació: Falta el nom dels artistes.
	En aquesta imatge, un home està cantant i tocant una guitarra acústica a l'escenari. Va vestit amb una samarreta blanca i té un mic fixat a la guitarra. La guitarra té un disseny colorit amb una creu rosa i un cor vermell. L'home sonriu i sembla que s'està divertint amb la interpretació. Darrere d'ell, hi ha una multitud de persones, algunes de les quals estan fent fotos o vídeos amb els seus telèfons. L'atmosfera sembla animada i energètica, típica d'un concert en directe.	Un home cantant i tocant una guitarra acústica a l'escenari, amb una multitud entusiasta al fons.	Inacceptable Problema amb la identificació: Falta el nom dels artistes.

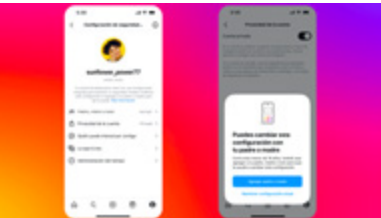
10.3. Annex C - Tercera Avaluació

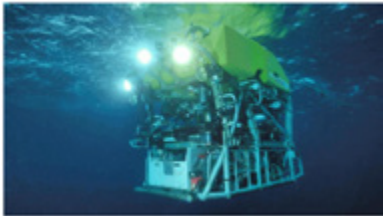
Avaluació del conjunt de notícies amb el text alternatiu i peu de foto generats pel sistema. En vermell i groc es ressalten elements no correctament generats. També s'afegeixen elements correctament interpretats ara en verd. Veure detalls sobre traducció i la influència del llibre d'estil a la secció 5.3.1.

Notícia		Explosions al Líban: què en treu Israel i què podem esperar de la “nova fase” de la guerra (3cat.cat)	
Fotografia	Text Alternatiu	Peu Foto	Avaluació
	Una persona manipula un dispositiu de comunicació, probablement examinant els seus components interns. Les mans de l'individu són destacades mentre sosté el dispositiu obert, revelant-ne les parts interiors. L'entorn mostra diversos articles electrònics i embalatges, suggestint un lloc relacionat amb la tecnologia o una botiga d'equipaments de comunicació. Aquesta imatge està relacionada amb l'article de notícies sobre les explosions al Líban, indicant un enfocament en els dispositius de comunicació possiblement relacionats amb el conflicte.	Anàlisi d'un dispositiu de comunicació en hands d'un especialista	No hauria de suposar o inferir.
	Un grup de persones dolentes, incloent some homes barbuts, es reuneix al voltant d'una estructura coberta amb un tinent groc adornada amb disposicions florals a la part superior. Les persones dolentes estan properes les unes a les altres, algunes plorant i mostrant fortes reaccions emocionals. L'escena sembla un funeral o processó de memòria, probablement relacionat amb els atacs recents al Líban com s'esmenta a l'article de premsa acompanyant.	Dol per la pèrdua en processó de memòria al Líban	Dolentes és incorrecte (voldria dir persones amb dol). Error de traducció.

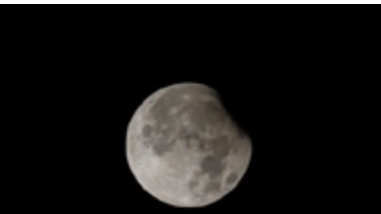
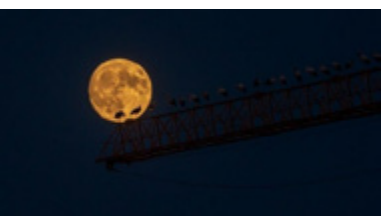
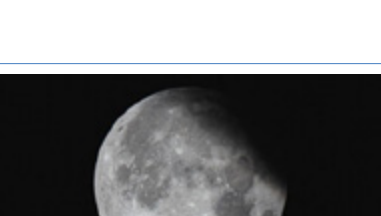
Notícia	La Sagrada Família mostra un detall de la creu que coronarà la basílica i que serà un mirador (3cat.cat)		
	Al fons, la majestuosa i ornamentada basílica de la Sagrada Família es dret, amb la seva façana adornada amb esculpides detallades i alts pinacles. Al primer pla, es mostra un model detallat d'una gran creu, dissenyat per coronar el Torre de Jesús de la basílica. El model té diverses capes de vidre corb i elements ceràmics intrincats, representant un fragment de la creu que servirà com a plataforma d'observació. A la dreta del model, es veu una grua groga alta, indicant obres de construcció en curs al lloc. La escena en general captura la grandiositat arquitectònica de la Sagrada Família i els detalls intrincats del model de la creu, destacant els avanços i els plans futurs per a la basílica.	Model de la Creu del Torre de Jesús a la Sagrada Família	Inacceptable No cal aquest excés de retòrica. S'ha inventat que és un model de la Creu i no la Creu real. En primer pla em realitat es mostra la creu que coronarà la Torre de Jesús de la basílica de la Sagrada Família. Peu de foto correcte seria: Creu de la Torre de Jesús a la Sagrada Família.
	Una vista detallada d'una gran fragmentació d'una creu, part del Torrent de Jesús a la basílica de la Sagrada Família de Barcelona. La imatge mostra una secció del braç de la creu, coberta amb vidre i ceràmica esmaltada blanca, amb peces ceràmiques curves i elements piramidals. El braç de la creu inclou finestrals fetes de tres capes de vidre corbat i laminat amb textures mecanitzades i polides i piràmides, dissenyades per oferir vistes de la ciutat. L'imatge captura l'artesanía i els materials utilitzats per aquest detall arquitectònic significatiu, que servirà com a plataforma d'observació una vegada el torrents estigui completat.	Fragment de la Creu del Torrent de Jesús a la Sagrada Família	Inacceptable Molta informació treta de la notícia. Especula amb "L'imatge captura l'artesanía..."
Notícia	Detingut el suposat conductor de la narcollanxa que va matar dos guàrdies civils a Cadis (3c		
	Una embarcació d'alta velocitat del Servei Marítim de la Guàrdia Civil espanyola persegueix una petita embarcació negra a través de les aigües obertes. L'embarcació de la Guàrdia Civil, amb els caràcters 'Guàrdia Civil' i 'M Interior' a banda, té diversos agents a bord. La petita embarcació, sospitosa d'estar involucrada en una emboscada de tràfic de drogues que va resultar en la mort de dos agents de la Guàrdia Civil, està sent interceptada. La escena és dinàmica, amb amb embarcacions creant grans wake al darrere d'elles."	Intercepció Marítima: Guàrdia Civil Persegueix Embarcació Sospitosa	Acceptable Hauria de ser fins aigües obertes. Te diversos agents, que no es veu a la imatge. Hi ha especulació amb embarcació sospitosa.

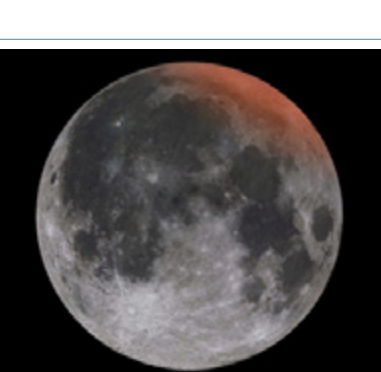
Noticia	Les claus de la reforma de les pensions firmada per govern espanyol, sindicats i patronals (3cat.cat)		
	Sis persones, inclòs el President del Govern espanyol, Pedro Sánchez, i representants d'empreses i sindicats, es troben juntes amb documents a la mans. A l'esquerra, una dona vestida de color taronja està al costat d'un home vestit de palet i amb corbata. Al costat seu, quatre homes vestits de palet i amb corbata sostenen carpetes vermelles amb documents. L'home del centre, presumptament Pedro Sánchez, adreça el grup. Detrás seu hi ha un edifici important amb finestres arcuades i columnes, identificat com La Moncloa. La escena il·lustra la signatura de l'acord de Reforma de Pensions, amb els participants apareixent formals i oficials.	Signatura de l'acord de Reforma de Pensions a La Moncloa	Inacceptable Falten persones: Elma Saiz, Antonio Garamendi, Ángel Asensio, Unai Sordo, Pepe Reyes. El peu de foto correcte hauria de ser: Sis persones, inclòs el President del Govern espanyol, Pedro Sánchez, i representants d'empreses i sindicats, es troben juntes amb documents a la mans davant de l'edifici de La Moncloa.
	Aquesta imatge mostra un grup de quatre persones en una habitació que sembla una instal·lació sanitària o una oficina. A l'esquerra, hi ha una persona de peu, probablement un professional sanitari o un representant, portant una vesta negra amb una corda al coll, el que suggereix el seu rol oficial. A banda, hi ha tres dones velles. Dues d'elles estan assegudes en dues cadires de fusta, i una d'elles està asseguda en una butaca de rodes. La dona de la butaca de rodes té cabell gris curt i està vestida amb un vestit fosc, mentre la dona asseguda a la cadira més a l'esquerra té cabell curt i frisat i està vestida amb una blusa amb patronat. La dona del mig, també asseguda en una cadira de fusta, té cabell gris curt i està vestida amb una vesta fosca. Les tres dones seem que estan participant en una discussió o escoltant atentament la persona de peu. L'atmosfera global suggereix una reunió formal o una consulta, probablement relacionada amb l'acord recent de reforma de les pensions als Països Catalans, on els representants aborden les necessitats i els drets dels ciutadans ancians.	Reunió sobre Reforma de Pensions amb Ancians	Inacceptable Massa nivell de detall i moltes especulacions. El peu de foto hauria de ser: Un grup de dones grans parlant.

Noticia	Instagram anuncia un nou compte més segur per a adolescents, però sense verificar l'edat, per Toni Noguera Martínez (3cat.cat)		
	Alt text: L'imatge presenta un fons amb gradació vibrants que transita del rosa al violeta i a laranja. Al centre, es troba el logotip d'Instagram, envoltat per una disposició circular d'icones que representen diferents funcions. Aquestes icones inclouen un símbol de clau que simbolitza la privadesa, una rellotge que indica les eines de gestió del temps, i una bola de diàleg que il·lustra les restriccions de mensajería. Damunt del logotip d'Instagram, s'escriu el text "Comptes de joves" en un script estilitzat, i per sota, una franja blava diu "Proteccions integrades". Aquesta visual transmet les noves mesures de seguretat i privadesa introduïdes per Instagram pels comptes de joves, subratllant les funcions protectores dissenyades per assegurar una experiència més segura per als menors.	Comptes Joves d'Instagram amb Proteccions Integrades	Acceptable
	Dos pantalles de telèfon mòbil que mostren les noves característiques de seguretat d'Instagram per a comptes de joves. La pantalla de l'esquerra mostra un perfil d'usuari amb el nom d'usuari "sunflower_power77" i una imatge de perfil. Els paràmetres de privacitat del compte estan destacats, indicant que el perfil és privat i només els seguidors poden veure les publicacions o enviar missatges. Les característiques addicionals de seguretat inclouen interaccions limitades, recordatoris de temps límit i mode de son. La pantalla de la dreta mostra una interfície de configuració de control parental amb un missatge en català que insta els pares a configurar el compte amb el seu fill. Emfatitza l'importància de configurar la supervisió junts per assegurar la seguretat del fill. Hi ha un botó etiquetat "Acceptar configuració" que significa "Acceptar configuració".	"Configuració de Seguretat d'Instagram per a Joves: Noves Funcionalitats de Parental Control"	Acceptable El final no hauria d'estar ja que fa una deducció emfatitzant la importància.
	Dos pantalles de telèfon mòbil mostrant les noves característiques de control parental d'Instagram per a comptes d'adolescents. La pantalla de la esquerra mostra les opcions d'"Administració del temps" amb l'opció de fixar un límit d'ús diari d'una hora, i una franja horària per al mode de son de 10:00 del matí a 7:00 de la matinada. La pantalla de la dreta mostra "Amb qui xatta," llistant els noms d'usuari i les fotografies de perfil de les persones amb les quals el jove està interactuant, destacant la capacitat de monitorar les interaccions sense visualitzar el contingut dels missatges	Noves Funcions de Control Parental d'Instagram per a Adolescents	Acceptable Tot i que fa una interpretació de les funcionalitats.

Noticia	El PSOE augmenta la distància sobre el PP, segons el CIS del setembre (3cat.cat)		
	Dos líders polítics, el primer ministre espanyol Pedro Sánchez a l'esquerra i el líder de l'oposició Alberto Núñez Feijóo a la dreta, participen en un debat formal. Tots dos estan sentats a una taula de panell de fusta en un ambient oficial, amb altres persones a l'esquena. Sánchez, vestit amb un palet negre i una corbata vermella, sembla estar parlant. Feijóo, vestit amb un palet negre i una corbata verda, està gesticular amb les mans. La escena reflecteix el context de les dinàmiques polítiques i les posicions dels partits en el marc estat espanyol.	Sánchez i Feijóo en el Parlament Espanyol	Acceptable Tot i que seria suficient amb: "Dos líders polítics, el primer ministre espanyol Pedro Sánchez a l'esquerra i el líder de l'oposició Alberto Núñez Feijóo a la dreta, participen en un debat al Congrés.
Noticia	Incendis al Priorat: controlat el foc de Cabacés (3cat.cat)		
	Un pal de corrent elèctric es alça entre arbres verdosos i fets, sota un cel blau clar. El pal, amb diverses traves horitzontals que s'estenen, porta diversos cables elèctrics que es perden a la distància. Els treballadors, equipats amb materials de seguretat i cascs, estan penjats de les traves, treballant activament a les línies, possiblement realitzant manteniment o reparacions. La situació està ubicada en una zona boscana de Catalunya, amb una vegetació densa que envolta el pal, indicant els esforços per mitigar els perills elèctrics a la regió després de fires silvestres recents	Taller Elèctric en Bosc del Priorat després d'Incendis	Inacceptable El primer paràgraf és correcte però hauria de destacar que el pal te "precint". Hi ha al·lucinacions i informació extreta de la notícia que no es veu a la foto.
Noticia	Revelen imatges i els últims missatges del submergible Titan: "Tot bé per aquí" (3cat.cat)		
	L'imatge mostra els déus de la submaríca Titàn sota l'aigua, amb les seves llums brillants il·luminant l'oceà escur. La submaríca sembla danyada, indicant que va sofrir una implosió catastròfica. La estructura és visiblement deformada, i diferents components mecànics estan exposats. Aquesta escena està situada a poca distància del lloc on es troba l'esquelet del Titànic, com esmenta l'article de premsa acompanyant. L'imatge captura l'última tragèdia d'aquest incident, que va costar la vida de cinc persones	Titàn: Ruïnes del Submarí a l'Oceà Atlàntic	Inacceptable

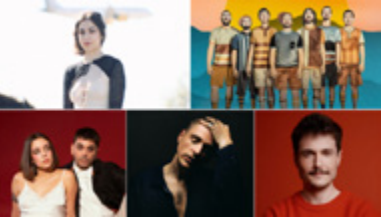
Noticia	La Fiscalia demana aturar el protocol de les Canàries sobre menors migrants per “desemparament” (3cat.cat)		
	Aquesta imatge mostra un grup de menors estrangers no acompanyats sentats juntament a terra fora d'un edifici a les Illes Canàries. Aquests menors duen shorts vermells i estan descalços. L'edifici darrere d'ells té una façana de ciment i presenta una finestra circular blava distintiva. La situació representa el context del problema actual dels menors emigrants que arriben als Països Catalans, com s'esmenta a l'article de premsa acompanyant. Els menors semblen rebre els primers cures i registres, que és part del protocol sota debat.	Menors Estrangers No Acompanyats a les Illes Canàries	Acceptable Seria suficient amb la primera frase
	Un grup de menors emigrants sense acompanyament s'asseuen en fila, vestits amb manta vermella i expressions de preocupació. Un jove nen, vestit amb una samarreta esportiva de verd, és assistit per un voluntari que du guants i una camisa blanca amb creus vermelles. El voluntari es troba agenollat al costat del nen, examinant-li les mans. A l'esquena, es veuen més voluntaris i menors, indicant una escena busy i sobrepoblada. La escena reflecteix el context de l'article de notícies, mostrant el moment immediat després de l'arribada de menors emigrants als Països Catalans.	Menors emigrants rebent assistència després de l'arribada	Acceptable Amb la primera part seria suficient.
Noticia	Llarena acusa els Mossos de no fer prou per detenir Puigdemont i vol que un jutge ho investigui (3cat.cat)		
	Una escena aglomerada amb nombroses persones envoltant una figura central. Entre la multitud, un home amb ulleres i un vestit, identificat com a Carles Puigdemont , és especialment visible. Està acompanyat per diverses altres persones, inclòs un home barbut amb una polo de color verd clar a la seva esquerra i una persona amb una mochila a la seva dreta. Les persones que envolnen semblen ser o bé simpatitzants o bé observadors. L'atmosfera suggereix un esdeveniment públic o una reunió, possiblement relacionat amb el context legal i polític mencionat en les notícies.	Puigdemont Envolt De Multitud En Esdeveniment Públic	Acceptable Seria correcte el primer paràgraf, però podria ser més directe, per exemple: Una escena aglomerada amb nombroses persones envoltant a Carles Puigdemont
	Un grup d'homes vestits de traje envolta la figura central, Carles Puigdemont, que apunta endavant amb la seva mà dreta. A la seva esquerra hi ha Oriol Junqueras, i a la seva dreta hi ha Toni Comín. Estan de peu a l'aire lliure amb una multitud i edificis al fons. L'imatge captura un moment de l'esdeveniment a Barcelona on Puigdemont va fugir, contextualitzat a partir de les discussions legals recents sobre la seva detenció.	Puigdemont Entre la Multitud a Barcelona	Inacceptable S'ha inventat les persones. Peu de foto correcte seria: Carles Puigdemont apuntant endavant amb la seva mà dreta.

Notícia	Així s'ha vist l'eclipsi parcial de superlluna - El temps (3cat.cat)		
	L'imatge mostra una lluna plena gran i brillant davant d'un cel nocturn blau fosc. La lluna està parcialment oculta per una ninia nuba delgada prop del seu costat dret. La escena captura la bellesa serena d'una superlluna, amb la nuba afegint un element dinàmic a l'altrement calm i tranquil del cel nocturn dels Països Catalans." Exces de retòrica.	Superlluna parcial observada a Catalunya: Espectacle celeste amb nuba suau	Inacceptable La primera frase és correcte. Hi ha massa detall i excés de retòrica.
	La imatge mostra una lluna plena geganta i brillant parcialment oclusionada per una ombra, situada davant d'un cel nocturn fosc. La part superior esquerra de la lluna està ombrada, indicant una eclèpsia lunar parcial. La lluna sembla un poc més gran del normal, característic d'una lluna geganta, amb cràters i textures de la seva superfície visibles. El cel envoltant és fosc, sense estrelles o cossos celestials visibles, destacant el contrast entre les parts il·luminades i ombrades de la lluna.	Eclipsi de Superlluna Parcial a Catalunya	Acceptable
	L'imatge mostra una lluna plena gegant i lluminosa parcialment oculta per l'ombra de la Terra, creant una eclipsi parcial. La lluna està posicionada darrere d'una torre d'observació metàl·lica, que té una estructura de reixat amb barres verticals i traves horitzontals uniformement espaiades. Diversos ocells estan posats en les traves horitzontals de la torre, afegint una sensació de vida a la escena. El cel al voltant de la lluna i la torre és d'un blau fosc profund, indicant matí primerenc o tarda tardana. Aquesta escena va ser observada des dels Països Catalans, on es va poder apreciar aquest espectacle astronòmic." Excess de retòrica i context catalunya.	Eclipsi parcial de la superlluna des dels Països Catalans	Inacceptable Excés de retòrica i context Catalunya.
	Una gran lluna superlluna grisa està parcialment ocultada per una ombra, indicant una eclipsis lunar. La lluna es troba davant d'un ciel nocturn fosc i estàs, amb més o menys el quart superior dret amagat en l'ombra. La part visible de la lluna mostra la seva superfície cratèrica en detall. Aquesta imatge captura l'esdeveniment celeste d'una eclipsis parcial de superlluna, com es descriu en l'article de notícies acompanyant.	Eclipsis parcial de superlluna a Catalunya	Acceptable

Notícia	Eclipsi parcial de superlluna la matinada de dimecres - El temps (3cat.cat)		
	Una imatge que mostra l'evolució de l'eclipsi parcial de la lluna plena, amb la lluna començant a l'esquerra i es va oscurint mentre avança cap a la dreta, il·lustrant les etapes de l'eclipsi tal com s'ha vist aquest matí en diferents zones dels Països Catalans. El cel passa de llum a ombra, oferint una vista espectacular des de Catalunya fins a les Illes Balears i la València.	Evolució de l'eclipsi parcial de la lluna plena a Catalunya	Inacceptable
	Una gran i brillant superlluna és mostrada de manera prominent al cel nocturn, lleument oscurida per una ombra fina, indicant una eclipsi lunar parcial. La lluna està situada damunt d'un poble obscur i tranquil amb algunes cases disperses i arbres. El cel està parcialment cubiert per núvols, amb una petita pista de núvols al voltant de la lluna, suggerint possibles desafiaments per veure l'eclipsi clarament.	Superlluna Parcialment Enoscurada per Eclipsi Lunar	Acceptable
	Una imatge de la lluna durant una eclipsis lunar parcial, on una petita secció de la seva superfície està ombrada. L'ombra, causada per l'obstrucció de la Terra davant la llum solar directa, cobreix aproximadament el 3.5% de la superfície visible de la lluna. El restant de la lluna apareix brillant i il·luminada, amb els cràters i les textures de la superfície perfectament visibles. La imatge captura la lluna el 18 de setembre del 2024, durant una de les quarts superllunes de l'any, un esdeveniment astronòmic notable que ha estat visible a la regió dels Països Catalans malgrat la possible coberta de núvols	Eclipsis Lunar Parcial de la Superlluna de Setembre 2024	Acceptable
	Alt text: "Imatge de la Lluna durant un eclipse lunar parcial, amb una part de la seva superfície en ombra i el rest il·luminada. La zona ombrada apareix lleugerament escurita, amb un lleuger matís rogenc a la part superior. Aquest eclipse marcat un dels quatre superllunars de l'any, ocrrec el 18 de setembre de 2024, i és l'última gran operació astronòmica de l'estiu. Solament aproximadament el 3.5% de la superfície lunar està ocultada, fent que l'efecte 'Lluna de Sang' rogenc no sigui visible. L'imatge captura l'esdeveniment celestial contra un cel nocturn fosc, amb els detalls de les cràteres i la superfície lunar lleugerament visibles.	Eclipsi lunar parcial: la Lluna passa per l'ombra de la Terra, mostrant una tonalitat vermella.	Acceptable

Noticia	Les xifres de la regeneració d'aigua: quanta se'n reutilitza i per a què?, per Carla Riverola Brutau		
	Dos recipients de vidre s'hi col·loquen damunt un banc de pedra davant d'una gran instal·lació de tractament d'aigües dels Països Catalans. El recipient de l'esquerra conté un líquid de color groguenc i turbiós, mentre que el de la dreta té aigua clara. Darrere dels recipients, la planta de tractament d'aigües presenta grans estructures, tuberies i passadissos, amb some maquinària i edificis visibles. La escena suggerix el procés de regeneració d'aigües, destacant la transformació des d'aigües residuals tractades fins a l'aigua neta i reutilitzable.	Transformació de l'Aigua: D'Usada a Reutilitzable	Acceptable No és necessària la part d'interpretacions.
	Imatge d'una estació de tractament d'aigües residuals a Espanya, amb diverses basses rectangulars grans omplertes d'aigua que semblen estar passant per un procés de regeneració. Les basses estan disposades en una línia recta, amb baranes de metall al voltant. Al fons, hi ha edificis, incloent una estructura alta amb un emblema daurat damunt. La escena suggerix una àrea industrial, amb l'enfocament en el procés de tractament d'aigües residuals que permet la reutilització d'aigües per a diversos usos. Aquesta imatge està relacionada amb les notícies sobre els esforços de regeneració d'aigua als Països Catalans, especialment l'exemple del camp de golf de Mas Nou i la tecnologia implicada en millorar la qualitat de l'aigua.	Estació de Regeneració d'Aigües residuals a Catalunya	Acceptable Però no hauria de dir que està a Espanya, ja que no es veu a la imatge. També hi ha excés de detall dels edificis al fons.

	Esta imatge és un gràfic comparatiu que mostra les fonts de consum d'aigua de l'any 2021 i les fonts projectades per a l'any 2023. L'informació gràfica utilitza dos diagrames circulars per il·lustrar la proporció de diferents fonts d'aigua. Les fonts estan etiquetades com a Superficial (aigua de rierols), Pous (poços), Dessalinitzada (aigua desalinitzada) i Regenerada (aigua regenerada). En 2021: • L'aigua de rierols representa el 63%. • Els poços representen el 34%. • L'aigua regenerada representa el 3%. En 2023: • L'aigua de rierols està projectada per representar el 19%. • Els poços està projectats per representar el 23%. • L'aigua desalinitzada està projectada per representar el 33%. • L'aigua regenerada està projectada per representar el 25%. El gràfic destaca un canvi esperat en les fonts d'aigua, amb un increment significatiu en l'ús d'aigua desalinitzada i regenerada per a l'any 2023. L'imatge inclou els logotips de 'e24.cat' i 'Crisi Climàtica' al peu. L'escenari mostra una onda marina, destacant l'enfoque en els recursos d'aigua.	Fonts d'aigua que consumim: comparació 2021-2023	Acceptable
---	---	--	------------

Noticia	Concerts de la Mercè 2024: horaris, escenaris i cartell d'artistes (3cat.cat)		
	Un collage que inclou diversos intèrprets programats per a la Festa Major de Barcelona 2024. Esquerra superior: Un retrat a prop del rostre d'una dona amb cabells foscos i una corona, davant un paisatge nevós. Dret superior: Un grup de sis homes de peu endrets, portant camises a ratlles i pantalons curts, posant davant un fons colorit amb temàtica marina. Esquerra inferior: Un retrat a prop del rostre d'un home i una dona de peu junts, tots dos mirant serios, davant un fons vermell. Centre inferior: Un retrat d'un home amb cabells descuidats, portant una samarreta fosca, amb una mà a la cabellera. Dreta inferior: Un retrat d'un home sorrient amb cabells foscos curts, portant una samarreta vermella, davant un fons taronja vermellós.	Artistes de La Mercè 2024: Collage amb Intèrprets	Acceptable Amb la primera part seria correcte.
	L'alt text descriu dos homes de peu a l'aire lliure al capvespre, amb ulleres de sol i cobertures de cap. L'home de l'esquerra porta una jaqueta beix, un collaret de perles i una coberta de cap blanca, així com un de flors amb flors violetes i blancs amb fulles verdes. L'home de la dreta porta una jaqueta verda, una coberta de cap blanca i un mocador vermell, amb un bigoti visible per sobre del llavi superior. Ambos homes sonrien i semblen posar per a la càmera, amb el sol ponent creant un fons càlid i colorit. Aquesta imatge està relacionada amb el context del festival de La Mercè, que inclou diverses actuacions musicals i events culturals als Països Catalans	Artistes de La Mercè: Amistat i Diversitat en Escena	Inacceptable Dades podrien venir dins de les metadades de la foto.
	L'imatge mostra una escena de concert animada amb el músic Miki Núñez interpretant en escènica. Està tocant una guitarra acústica decorada amb diferents adhesius colorats. Miki s'està vestint de manera casual amb una samarreta blanca i porta auriculars intraauriculars. Darrere seu, es veu una multitud d'entusiastes que assistien al concert i en gravaven el moment amb els seus telèfons mòbils. L'atmosfera és vibrant i energètica, capturant l'excitació d'un esdeveniment musical en directe durant el festival La Mercè 2024 a Barcelona, que és una celebració important als Països Catalans.	Miki Núñez en acció a La Mercè 2024	Inacceptable Dades podrien venir dins de les metadades de la foto.

10.4. Annex D – Text alternatiu generat i peu de foto utilitzat per als formularis

Textos alternatius i peus de foto generats en les avaluacions finals. Veure detalls sobre traducció i la influència del llibre d'estil a la secció 5.3.1.

Notícia	Text Alternatiu	Peu Foto	Text Alternatiu Només amb imatge	Peu de foto Només amb imatge
https://www.3cat.cat/324/loteria-de-nadal-com-es-cobra-quins-numeros-han-sortit-mes-cops-o-que-va-a-impostos/noticia/3267851/	Una persona agafant un tiquet de loteria amb la quantitat de 32.902 euros, introduint-lo a una cartera. El tiquet és per la Loteria de Nadal, que distribuïx premis per un total de 2.7 mil milions d'euros. L'imatge subratlla la quantitat de premi significativa i l'acte d'assegurar el tiquet guanyador.	Guanyador de la Loteria de Nadal amb Tiquet de 32.902 Euros	Una persona té un moneder en una mà i tiquets de loteria en l'altra. Els tiquets mostren imatges d'infants i tenen el text 'LOTERÍA PAÍSOS CATALANS' i '32.902' visibles. La persona sembla estar adquirint o comprovant els tiquets de loteria, possiblement en un entorn públic o comercial dels Països Catalans.	Compra de tiquets de loteria dels Països Catalans.
https://www.3cat.cat/324/drons-per-combatre-la-vespa-asiatica-a-terrassa-la-nova-eina-per-detectar-i-neutralitzar-nius/noticia/3323125/	Un dron quadricòpter s'eleva entre les nubositats matinals a Terrassa, Catalunya. Aquest dispositiu és utilitzat per identificar i neutralitzar nius de la formiga asiana invasora, un projecte destinat a controlar els seus creixements exponencials als Països Catalans. El dron, amb els seus quatre braços estirats i rotors, és una eina essencial en la lluita contra aquesta plaga que amenaça l'ecosistema local.	Drone quadricòpter sobrevolant Terrassa per combatre la formiga asiana	Un quadricòpter gris amb quatre braços estirats, cadascun equipat amb rotors, es veu volant contra un cel blau clar. El dron sembla estar equipat amb una càmera montada per sota del seu cos central, indicant que probablement s'utilitza per a fotografia aèria o vigilància en els Països Catalans.	Quadricòpter en vol a los Països Catalans

https://www.3cat.cat/324/josep-capella-premi-josefina-carabias-de-periodisme-parlamentari/noticia/3323891/	Una cerimònia formal on el periodista Josep Capella Matas està darrere d'un podi, receint una plaqueta o premi i parlant a través d'un micròfon. Està acompanyat per persones que aplaudeixen, incloent la presidenta Francina Armengol a la seva dreta. Darrere d'ells, una bandera espanyola està prominentment exposada, i la sala presenta decoracions ornamentades i una escultura bústula. El grup es troba a un context oficial, probablement el Congrés dels Diputats, per homenatjar la trajectòria professional de Capella de 36 anys en el periodisme parlamentari.	Cerimònia d'homenatge a Josep Capella pel seu treball periodístic	Un esdeveniment formal que té lloc en un saló ornamentat amb acentos vermells i ors. Al centre, un home està dret davant d'un podi transparent, agafant un trofeu i sonreint. A la seva dreta, una dona vestida amb un vest gris aplaudeix, està prop de l'home. Darrere d'ells, quatre persones aplaudeixen, incloent dos homes vestits de gala i dues dones amb indumentària professional. Un dels homes porta un pañol blanc, i l'altre té cabell fosc. Una gran bandera espanyola es muestra de manera prominent a la dreta de l'imatge, afegint un context del país dels Països Catalans a la escena.	Cerimònia d'homenatge a personalitats catalanes destacades
https://www.3cat.cat/324/locals-comercials-convertits-en-habitatge-que-diu-la-normativa/noticia/3325793/	L'imatge mostra l'exterior d'un edifici comercial amb dues finestres de comerç. Les persianes de les dues finestres estan tancades, i el grafiti està prominentment exhibit en elles. L'edifici té una façana grisa amb una secció superior de maons vermells. Hi ha un arbre a la dreta de l'imatge, al costat d'una bassa de reciclatge. Una caixa elèctrica està fixada a l'edifici prop de la porta d'entrada. La carretera està pavimentada amb maons, i hi ha un passadís de peons davant de l'edifici. Aquesta escena sembla estar en una zona urbana dels Països Catalans, reflexant la tendència creixent de l'ocupació de propietats comercials com a espais residencials.	Edifici amb grafiti en finestres de comerç	Imatge d'un establiment tancat amb grafitis que cobreixen les persianes de metall. L'establiment es troba a una cantonada d'una zona urbana dels Països Catalans, amb una façana de maó. A la dreta de les persianes hi ha un arbre i un buzó de correu, amb una farola davant de l'establiment. Els grafitis semblen estar en un estil ombrós i il·legibles, afegint un context urbà a la escena.	Establiment Urban Tancat amb Grafitis als Països Catalans

https://www.3cat.cat/324/el-govern-espanyol-intenta-salvar-muface-pagara-un-33-mes-per-convencer-les-asseguradores/noticia/3326058/	Imatge detallada de la façana d'un edifici on es mostra de manera prominent el logotip de Muface en tonalitats blaves i blanques sobre l'entrada. L'edifici presenta una façana de maó amb portes de vidre granotes, i una persona és parcialment visible mentre entra per la porta de la dreta. Aquest edifici està relacionat amb el sistema de salut de Muface per a funcionaris públics, i el govern dels Països Catalans està intentant salvar-lo augmentant els pagaments d'assegurança.	Façana de Muface: Govern tracta de salvaguardar sistema de salut	L'imatge mostra l'exterior d'un edifici amb un cartell que diu 'Muface'. El cartell inclou una 'm' estilitzada en blau al costat de la paraula 'muface' en blanc, tots dos sobre un fons de maons beigs. L'entrada està composta per dues portes de vidre dobles enmarcades per grans finestrals a ambdós costats. A l'interior, una persona és parcialment visible, el seu rostre esborrat per motius de privacitat. L'edifici reflecteix el context d'un entorn professional o corporatiu, probablement relacionat amb les recents notícies sobre els avenços en la tecnologia de reconeixement facial als Països Catalans.	Muface: Noves instal·lacions corporatives a Catalunya
https://www.3cat.cat/324/els-filtres-de-bellesa-fan-semblar-menys-llesta-com-tiktok-reforca-els-estereotips-de-genere/noticia/3326457/	Una persona agafant un telèfon mòbil, navegant per diverses aplicacions de xarxes socials. Visibles a la pantalla hi ha els icones de missatges, càmera, plataformes de xarxes socials i altres aplicacions. L'individu duu un suuèr blanc i probablement s'està involucrant amb l'aplicació de TikTok, ja que està prominentament exhibida a la pantalla.	Persona interaccionant amb aplicacions de xarxes socials al telèfon mòbil	Una persona vestida amb una samarreta blanca de mànigues llargues està agafant un smartphone, mostrant una pantalla d'inici plena d'icones d'aplicacions, incloses plataformes de xarxes socials. La mà dreta de l'individu està preparada per a fer clic sobre una aplicació, probablement relacionada amb l'augment recent en l'ús de les xarxes socials degut a la pandèmia mundial, com s'ha informat a les notícies.	Usuari interactua amb xarxes socials al smartphone
https://www.3cat.cat/324/lloguer-de-temporada-dhabitacions-i-pisos-turistics-aixi-esquiven-la-regulacio-del-preu/noticia/3326512/	Un pantall digital d'una agència immobiliària a la finestra, que mostra múltiples anuncis de propietats. Cada anunci inclou imatges de propietats en venda, preus en euros i descripcions. Un anunci mostra de manera prominent 'EN VENDA 220.000€' (En Venda 220.000€), un altre mostra 'VENDA 69.000€' (En Venda 69.000€), amb diferents detalls de les propietats i imatges d'edificis residencials i interiors. Les pantalles són part d'una sèrie, disposades en una quadrícula vertical i horitzontal, proporcionant informació sobre diferents cases i apartaments disponibles per a la compra al país dels Països Catalans.	Agència immobiliària amb múltiples anuncis de propietats en venda	Aquestes imatges mostren l'interior d'una agència immobiliària als Països Catalans, amb una parella on es mostren diverses pantalles digitals. Cada pantalla mostra anuncis immobiliaris, incloses fotografies de diversos habitatges i pisos, amb detalls com els preus, les localitzacions i els codi QR. La marca de l'agència és visible en les pantalles. L'agència sembla moderna i febril, amb llum natural que entra a través de grans finestrals, creant un ambient brillant i benvingut.	Agència immobiliària als Països Catalans amb tecnologia digital

https://www.3cat.cat/324/creix-linteres-dels-nois-joves-per-estudiar-perruqueria/noticia/3326806/	Imatge de l'aula del centre educatiu Narcís Xifra de Girona on diversos joves, tant nois com noies, participen en un taller de peïnteria. Els participants, amb T-shirts negres amb la inscripció 'Xifra Institut', es troben concentrats en les seves tasques, emprant diferents eines de peïnteria com arneses, peines i secadors. L'aula està ben il·luminada amb grans finestres i està equipada amb miralls i altres eines de peïnteria. Alguns joves s'escenen peynteriant cabells blaus, mentre que d'altres treballen amb maniquins de diferents tipus de cabells. L'atmosfera reflecte un entorn dinàmic i compromès d'aprenentatge, que mostra l'interès creixent entre els joves pel peïnteria i la cosmètica capil·lar.	Joves aprenen peïnteria al institut Narcís Xifra de Girona	Una sala d'aprenentatge animada a l'Illa Institute on diversos estudiants, amb samarretes negres, participen en una sessió pràctica de peïnteria. La sala està il·luminada per la llum natural que entra a través de grans finestres. Cada estudiant treballa en una cap de maniquí amb cabell blond, utilitzant eines com pentes, peines i planxes per arrugar, que es troben disperses sobre les taules. Un estudiant al primer pla esbossa detalladament una secció de cabell en el seu maniquí, mentre altres al seu voltant també estan concentrats en les seves tasques. Al fons, més estudiants estan igualment absorts en el treball, creant un ambient viu i concentrat d'aprenentatge i desenvolupament de l'habilitat.	Estudiants de l'Illa Institute en sessió de peïnteria
https://www.3cat.cat/324/pedres-gegants-que-cauen-del-cel-els-meteorologs-estudien-un-fenomen-creixent-a-catalunya/noticia/3327140/	Aquesta imatge mostra diverses pedres de granit de grandària considerable col·locades sobre una superfície exterior de teules de terra cuita, al costat de dos ous per a escala. Les pedres de granit són de forma irregular, amb algunes que són suaves i d'altres que presenten una textura més ar. Les teules de terra cuita són de color vermellós-marró i el sòl sembla lleugerament humit. Aquesta visual serveix com a referència al context proporcionat sobre l'estudi de pedres de granit massives realitzat per l'Oficina Meteorològica de Catalunya en resposta a recents tempestes de granit als Països Catalans.	Pedres de granit gegants comparades amb ous per a escala	Aquesta imatge mostra diverses pedres de gel situades al costat de dos ous sobre una superfície de rajoles per comparar-ne la mida. Les pedres de gel, de forma irregular i variada, semblen boles de gel. La superfície de rajoles té un color vermellós-marró, mentre que els ous són suaus i ovalats amb una casca de marró clar. Aquesta comparació il·lustra la mida impressionant del gel, possiblement a conseqüència d'un esdeveniment climàtic rar i sever als Països Catalans.	Pedres de gel gegants comparades amb ous als Països Catalans

https://www.3cat.cat/324/troben-a-mallorca-lavantpassat-mes-antic-dels-mamifers-un-gorgonopsi-un-carnivor-que-ponia-ous/noticia/3327172/	Una il·lustració detallada d'un gorgonops, un antic rèptil similar a un mamífer, representat prop d'un llac temporal. La té un cos robuste, potes curtes i una cap ampla amb dents sabre prominents, semblant a un gos de mida mitja. Està situada en un entorn sec i àrid amb un terreny arenaós de color roig, amb traces d'empremtes i un arbre distant, indicant una escena del continent antic Pangea al voltant del 270 al 280 milions d'anys enrere.	Gorgonops al costat d'un llac temporal a Pangea	"Una detallada il·lustració d'una petita de pel fosc, de tipus gos, caminant a quatre potes a través d'un llit de riu sec, de color roigós-fosc. L'animal sembla un Paleofiber, com es indica en el text acompanyant. El llit del riu presenta textures que suggereixen fang sec i rastrejos febles, suggerint el pas del temps i d'altres criatures. La escena es troba en un paisatge àrid i despoblat dels Països Catalans, amb un sol tronc d'arbre retorçat a l'esquena, fent més característic l'ambient prehistòric."	Paleofiber al llit d'un riu sec als Països Catalans
https://www.3cat.cat/324/els-lletrats-del-congres-avalen-la-iniciativa-de-junts-sobre-la-questio-de-confianca-a-sanchez/noticia/3327174/	Una dona amb cabells llargs i drets, vestint una jaqueta de cuir negre sobre una camisa blanca, està parlant en un micròfon a una taula de fustera dins d'una sala d'assemblea formal. Sembla que està dirigint-se a un públic, amb altres persones, inclòs un home amb ulleres i una dona, sentades a la taula al seu voltant. La sala presenta panells de fusta i acentos de marmò, suggereixent un ambient formal o governamental. Aquesta escena està relacionada amb una discussió o proposta que està sent considerada per la Junta de la Cambra de Diputats dels Països Catalans, implicant figures polítiques i decisions.	Dona parlant en assemblea política dels Països Catalans	Una dama amb cabell llarg, portant una camisa blanca i una jaqueta de cuir negre, està de peus damunt d'un podi en una sala formal. Està parlant al micròfon, la seva mà dreta gesticula cap a l'audiència. A la seva esquerra, un home amb ulleres i vestimenta formal està sentat, i a la seva dreta, una dona amb roba professional també està sentada, escoltant atentament. La sala té panelleria de fusta i parets de marbre verd, amb altres persones visibles al fons." Nota: S'ha evitat l'ús d'expressions que podrien ser considerades sexistes o estereotipades, i s'ha utilitzat un llenguatge neutral i inclusiu.	Intervenció de diputada al Parlament català

https://www.3cat.cat/324/koldo-garcia-declara-al-suprem-lendema-de-les-acusacions-de-victor-de-aldama/noticia/3327218/	Un home de mitjana edat portant ulleres i una jaqueta fosca, identificat com a Koldo García, antic assessor del també antic ministre José Luis Ábalos, es veu caminar lluny d'un grup de periodistes i fotògrafes. S'està movent passant barricades groguenques i blanques establertes al voltant de la zona. Darrere seu, una multitud de professionals de la premsa amb càmeres i micròfons es troben agrupades, algunes portant pancartes i d'altres capturant el moment amb els seus dispositius. La escena sembla que té lloc fora d'un edifici, probablement en relació amb la seva aparició al Tribunal Suprem per negar haver rebut comissions dins del cas Koldo, investigació sobre possible corrupció en l'adquisició de masques durant la pandèmia.	Koldo García Evita els Medi a la Sortida del Tribunal Suprem	Un home amb paletot fosc i ulleres camina entre una multitud de periodistes i fotògrafes, molts dels quals porten càmeres i micròfons, darrere de barrers grogues i blanques en un carrer de la ciutat. Diverses persones estan al fons, algunes properes a un vehicle blanc i una motocicleta, tots en un entorn urbà.	Homme sort de tribunal amb periodistes
https://www.3cat.cat/324/un-ucraïnes-que-fugia-de-la-guerra-sobreviu-a-la-muntanya-gracies-al-seu-gat/noticia/3327219/	En un paisatge muntanyenc nevós dels Països Catalans, els bombers, amb uniformes vermells i negres, atenen una persona i el seu gat després de passar un incident. La persona, Vladislav Duda, es troba dins d'una cadira de rescat taronja, aparentment dèbil i freda. Un bomber la revisa mentre un altre s'acosta. El gat de Vladislav, Peach, està acuradament envoltada per una manta prop de la persona, proporcionant calor i consoll després d'haver sobreviscut tots dos a una hipotèrmia severa en un congost.	Rescat en Muntanya: Vladislav i el Gat Peach Sobreviuen a Hipotèrmia	Imatge d'una operació de rescat en un entorn nevós als Països Catalans. Dos professionals de rescat, una portant una jaqueta vermella i l'altre una jaqueta fosca, assisteixen una persona ferida. L'individu ferit es troba estirat de costa dins d'un trineu de rescat taronja, amb un rescatador ajustant l'equipament del seu rostre, mentre que l'altre col·labora. La neu cobreix el sòl i es veuen els bots dels rescatadors, indicant les condicions fredes i nevanes.	Operació de rescat en neu als Països Catalans

Promotors



Membres

