

PRESENTACIÓ DEL PROJECTE

Assistència intel·ligent a la
descripció contextualitzada
de fotos per notícies

5 juny 2025



Presenta i modera



Marco Orellana

Gerent
CIDAI

Benvinguda institucional



Sr. Joan Mas i Albaigès

*Director, **CIDAI***

*Director Científic Digital, **Eurecat***



Sra. Rosa Romà Monfà

Presidenta

Corporació Catalana de Mitjans Audiovisuals- 3Cat



Sra. Maria Galindo

Secretària de Polítiques Digitals
Generalitat de Catalunya

Continguts

10:00h Benvinguda

10:15h Projecte “Assistència intel·ligent a la descripció contextualitzada de fotos per a notícies”

- Sr. Medir Plandolit, Cap Digital Informatius, **3Cat**
- Sr. Rafael Redondo, Cap Línia d’Imatge de l’Unitat de Tecnologies Multimèdia d’**Eurecat**

11:00h Avaluació del Projecte

- Sr. Rafael Bermúdez, recerca i serveis transversals **3Cat**

11:30h Preguntes i Cloenda

11:35h Cafè & Networking

12:00h Fi de l’acte

Presentació del Projecte “Assistència intel·ligent a la descripció contextualitzada de fotos per a notícies”



Sr. Medir Plandolit

Cap Digital Informatius

3Cat

Continguts

- Sobre el moment actual
- Sobre les corporacions públiques
- Sobre nosaltres

- **Desafecció** amb la democràcia.
- **Sobreabundància** de notícies i desinformació: dificultat per identificar continguts de qualitat.
- **Pèrdua d'interès** en les notícies.
- **Consum digital majoritari:** xarxes socials, cercadors i agregadors, principal via d'accés a les notícies.
- Gran domini de grans tecnològiques i gegants de continguts.
- Les corporacions públiques han de ser competitives per seguir **sent rellevants** i poder continuar sent actors clau en la construcció d'una democràcia sana.

Context europeu

- El 50% veu la IA com una prioritat en els propers 5 anys
- Només el 3% té una estratègia d'IA ben definida i implementada
- L'ús majoritari avui és per a millorar el contingut i la IA generativa és principalment per experimentar-hi

Usos principals dins de les redaccions: eficiència i productivitat

- Alliberar els equips de feines repetitives i tasques simples
- Arribar a més audiència amb una millor oferta.

Principals reptes i limitacions

- Limitacions financeres
- Dificultats i perfils tècnics
- Qualitat dels resultats
- Seguretat de les dades

Consideracions ètiques i valors de servei públic: TRANSPARÈNCIA I CONTROL HUMÀ

Creació de la Comissió d'IA Responsable:

“Vetllar per la difusió dels principis ètics, la formació en bones pràctiques i l'autoregulació de l'ús de tecnologies basades en la IA.”

Procés de transformació en l'àmbit informatiu

- Optimitzar i millorar processos de producció amb noves eines.
- Nous canals digitals per connectar amb més públics.
- Augmentar la qualitat, la veracitat i el rigor.



Rafael Redondo

Cap de Línia d'Imatge de l'Unitat de Tecnologies Multimèdia
Eurecat



ChatGPT

DESCRIPCIÓ FOTOS PER NOTÍCIES

- Descripció precisa del contingut, context i importància de la fotografia a l'hora d'acompanyar la notícia.
- Ràpida comprensió de la imatge.
- Millora de la narració, profunditat i context.
- Context precís: identificació de persones, llocs, objectes, esdeveniments,...
- Comprensió de la història, més enllà del moment capturat.
- Clars i concisos, ràpids i eficaços: poden animar als lectors a llegir l'article sencer.



DESCRIPCIÓ CONTEXTUALITZADA DE FOTOS

1. PEU DE FOTO

- Més enllà de la descripció visual.
- Criteris editorials.

2. TEXT ALTERNATIU

- Descripció textual de contingut no textual (textos, gràfics,...).
- Ha de complir la mateixa funció d'allò que serveixen els subtítols de pantalla dirigits a persones amb discapacitat visual.
- Tasca rutinària, sovint no considerada en fluxos de treball.



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

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



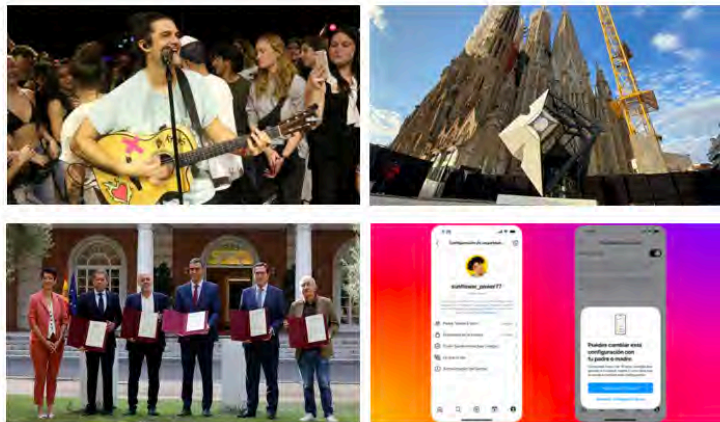
Assistir al professional per accelerar processos.



Trencar barreres d'accessibilitat, sense discriminació.

Contingut imatge, text, disposició i interacció d'elements,...

Large Vision Model



Large Language Model

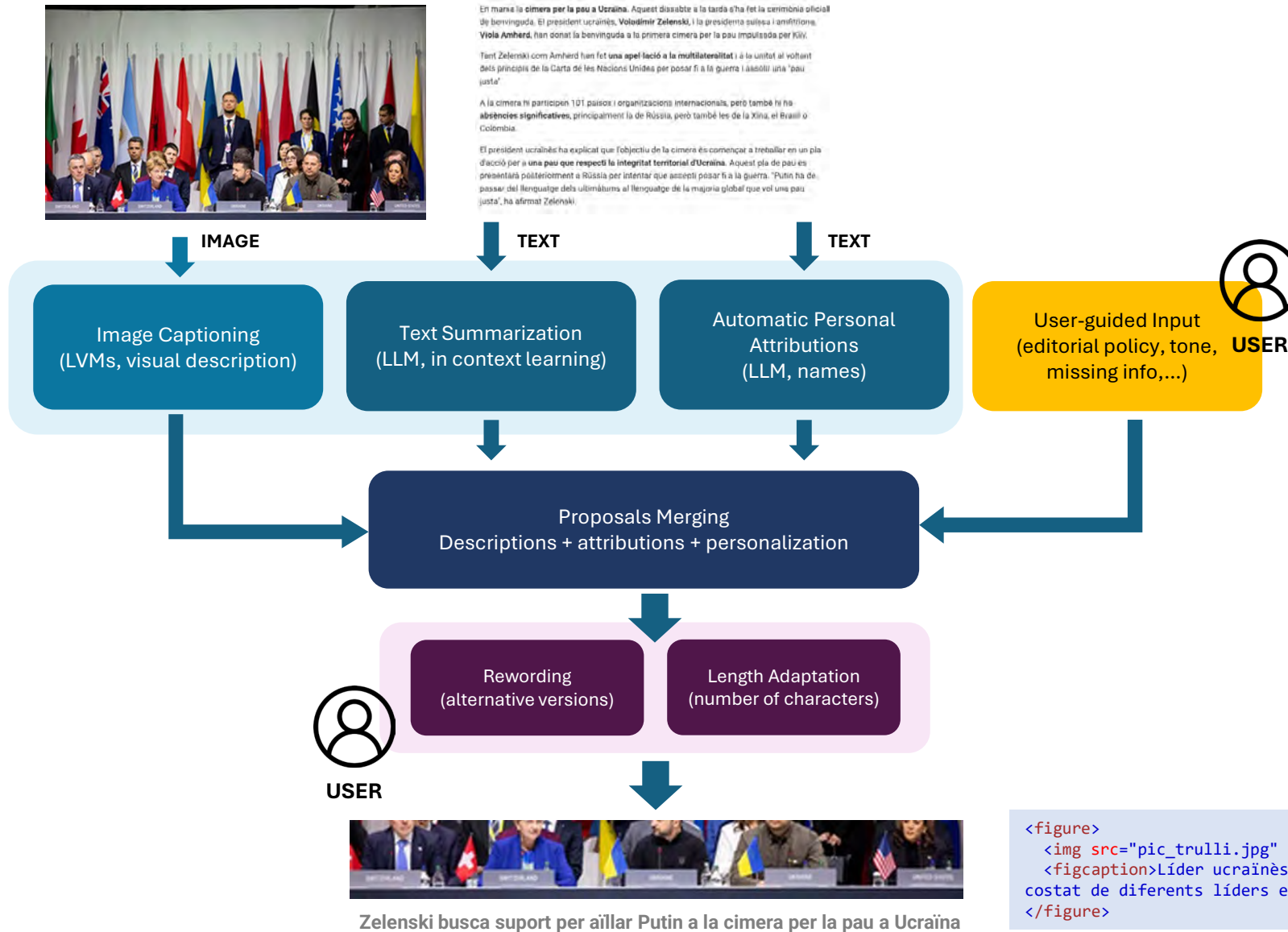
Natural Language Processing

- Extracció d'informació
- Anàlisi semàntic
- Anàlisi emocional
- Anàlisi del context
- Resposta preguntes
- Desambiguació
- ...

Informació textual, instruccions, context,... (ajuda enteniment del LVM)

+

PROMPT ENGINEERING



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	Imms-lab/llava-onevision-qwen2-0.5b-si	-	Apache 2.0	https://huggingface.co/Imms-lab/llava-onevision-qwen2-72b-ov-sft
CLIP		Fa servir GPT2, que té un rendiment molt baix	MIT license	rmokady/CLIP_prefix_caption: Simple image captioning model (github.com)
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	kyegomez/Reka-Torch: Implementation of the model: "Reka Core, Flash, and Edge: A Series of Powerful Multimodal Language Models" in PyTorch (github.com)
Reka	Flash-20240904	-	MIT license	kyegomez/Reka-Torch: Implementation of the model: "Reka Core, Flash, and Edge: A Series of Powerful Multimodal Language Models" in PyTorch (github.com)
InterntVL	InternVL - 26b	-	MIT license	OpenGVLab/InternVL: [CVPR 2024 Oral] InternVL Family: A Pioneering Open-Source Alternative to GPT-4o. 接近GPT-4o表现的开源多模态对话模型 (github.com)
InternVL	InternVL - 4b	Massa sortida (incontrolable)	MIT license	OpenGVLab/InternVL: [CVPR 2024 Oral] InternVL Family: A Pioneering Open-Source Alternative to GPT-4o. 接近GPT-4o表现的开源多模态对话模型 (github.com)

Llibre d'estil

- **Guia essencial** per a professionals de la comunicació.
- Defineix la manera com es **presenta la informació**.
- Influeix en la **percepció** i la **comprensió** del lector.
- L'ús adequat del llenguatge, el to, la claredat i la coherència **milloren la qualitat de la comunicació**.
- Un estil consistent pot construir una millor credibilitat i **confiança amb l'audiència**.



1. Denominacions
2. Paritat i equilibri de gènere
3. Imparcialitat, neutralitat i pluralisme
4. Fets violents
5. Violència de gènere i violència domèstica
6. Conflictos bèl·lics i catàstrofes
7. Suïcidi
8. Immigració
9. Les persones grans
10. Crisi climàtica

Llibre d'estil

1. Reestructuració en seccions temàtiques
→ Facilitar interpretació pel model
2. Reducció extensió
→ Adequació finestra de context del model
→ Coherència + precisió

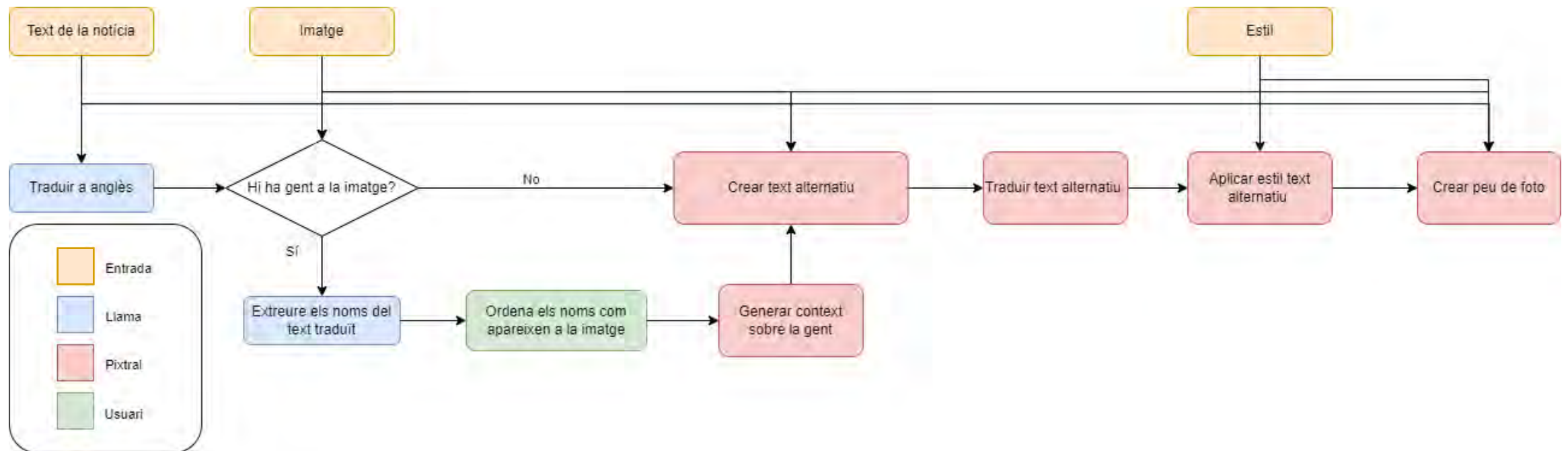


1. Denominacions
2. Paritat i equilibri de gènere
3. Imparcialitat, neutralitat i pluralisme
4. Fets violents
5. Violència de gènere i violència domèstica
6. Conflictes bèl·lics i catàstrofes
7. Suïcidi
8. Immigració
9. Les persones grans
10. Crisi climàtica

Instruccions

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.
4. Els models actuals han mostrat major efectivitat al ser **instruïts en angles**.
5. El **llibre d'estil** serà més efectiu en l'idioma original (català) pel que fa a la introducció de **biaixos i alteració semàntica** d'alguns termes.
 - *p.ex. corregint “homes” per “participants”, però també inclinació al terme “Països Catalans” més freqüentment del habitual o fora de context.*

Flux de Treball



Chain of Thoughts (I)

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}

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}

Chain of Thoughts (II)

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.





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 ue 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 correctament de text alternatiu i peu de foto en una notícia⁶ amb bon tractament del llibre d'estil. També s'observen un parell d'errors gramaticals i de traducció.



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.



Exemple generat incorrectament de text alternatiu i peu de foto en una notícia⁷ esbiaixat principalment pel context de la notícia. També s'observa una imprecisió en la disposició de les persones a la imatge.

Experimentació

- 8 redactors i editors de notícies, 1 persona responsable d'accessibilitat i 2 usuaris final.
- 13 les notícies diferents.
- Funcionalitat, Fiabilitat, Personalització, Facilitat, Rendiment.

Conclusions (I)

- **77%** dels text alternatius i un **70%** dels peus de fotos qualificats amb **bona o mitjana qualitat**.
- Només un **23%** de texts alternatius i **54%** de peus de fotos són publicables amb **mínima intervenció humana**, sempre present per garantir estàndards de qualitat elevats.
- UI valorada positivament per 3/4 participants en **facilitat d'ús** amb **rendiment** del sistema adequat.
- Els models LLMs oberts en **català no estan madurs** i els traductors no són precisos (encara).
- L'ús del **llibre d'estil** aporta claredat, coherència i ètica i per altra banda introdueix biaixos innecessaris, p.ex. geogràfics.
- Contexts de notícia massa extensos poden derivar en **suposicions errònies** sobre la imatge, alguna vegada amb **detalls superflus**.
- La metodologia semi-supervisada **d'identificació de persones és positiva** pels professionals per generar textos de major qualitat, especialment pels text alternatius. Tot i que altres models d'identificació basats en similitud facial podrien ajudar als LLMs de manera molt efectiva.

Conclusions (II)

- La **supervisió humana** continua sent necessària per assegurar **precisió i rellevància** en descripcions, amés de fomentar **credibilitat** i la **confiança** dels lectors.
- En la generació de descripcions de fotografies, no és tant l'estalvi de temps, sinó la millora de qualitat dels textos allà on actualment no hi és, cap a un **assistent periodístic més integral**.

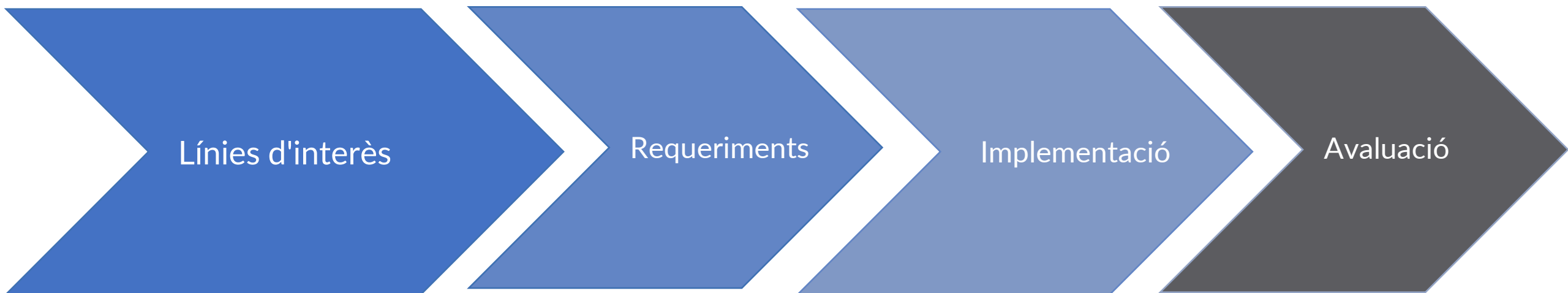
Avaluació del Projecte Assistència intel·ligent a la descripció contextualitzada de fotos per a notícies



Sr. Rafa Bermúdez

Recerca i Serveis Transversals
3Cat

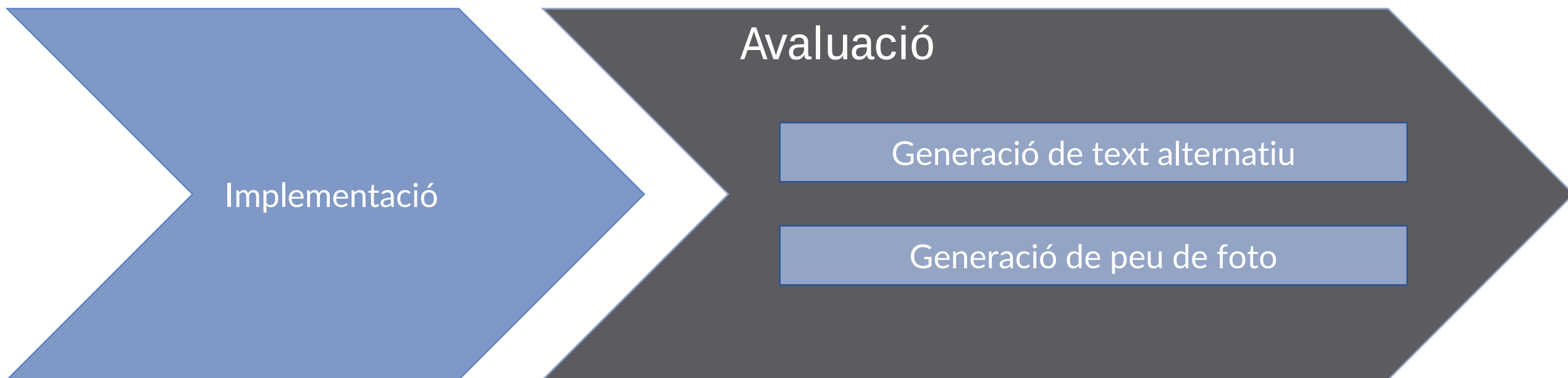
Timeline d'avaluació



Línies d'interès de 3Cat

Redacció: Millora de l'accessibilitat del portal amb la generació automàtica de text alternatiu d'alta qualitat

Requeriments



Qui?

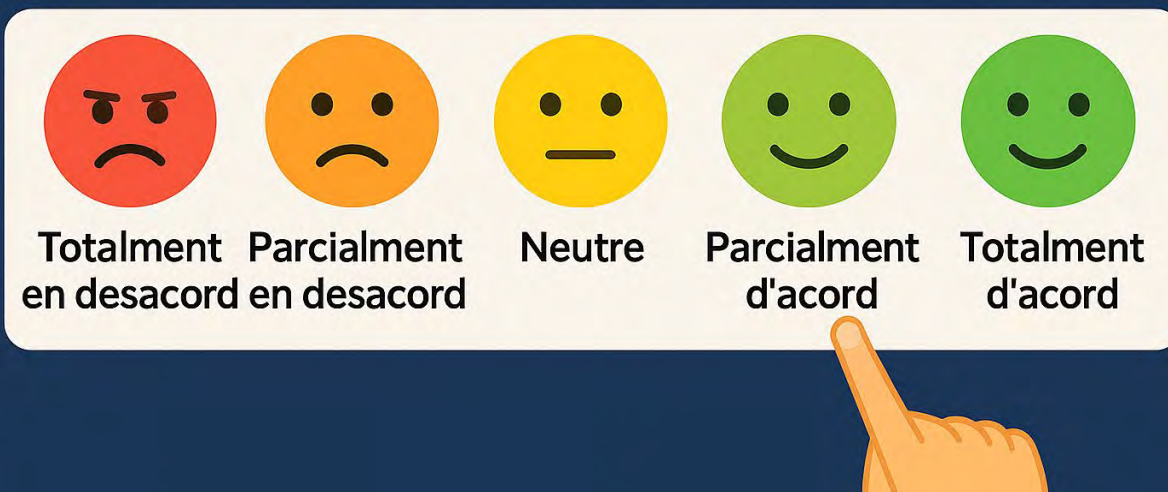
- Quatre àrees de l'organització
 - Redacció informatius
 - Accessibilitat
 - RSC
 - Documentació
- Representants amb llarga experiència i coneixement profund del domini

Com?

- Presentació d'exemples de resultats
- Proves dirigides sobre el software real
- Avaluació en escala Likert amb cinc afirmacions per mesurar:
 - Funcionalitat
 - Facilitat d'ús
 - Fiabilitat
 - Personalització
 - Rendiment

FUNCIONALITAT

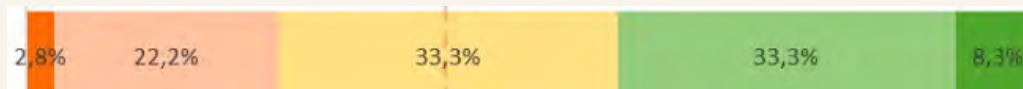
El text alternatiu o peu de foto generat descriu la part més important de la imatge tenint en compte el context



FUNCIONALITAT

El text alternatiu o peu de foto generat descriu la part més important de la imatge tenint en compte el context

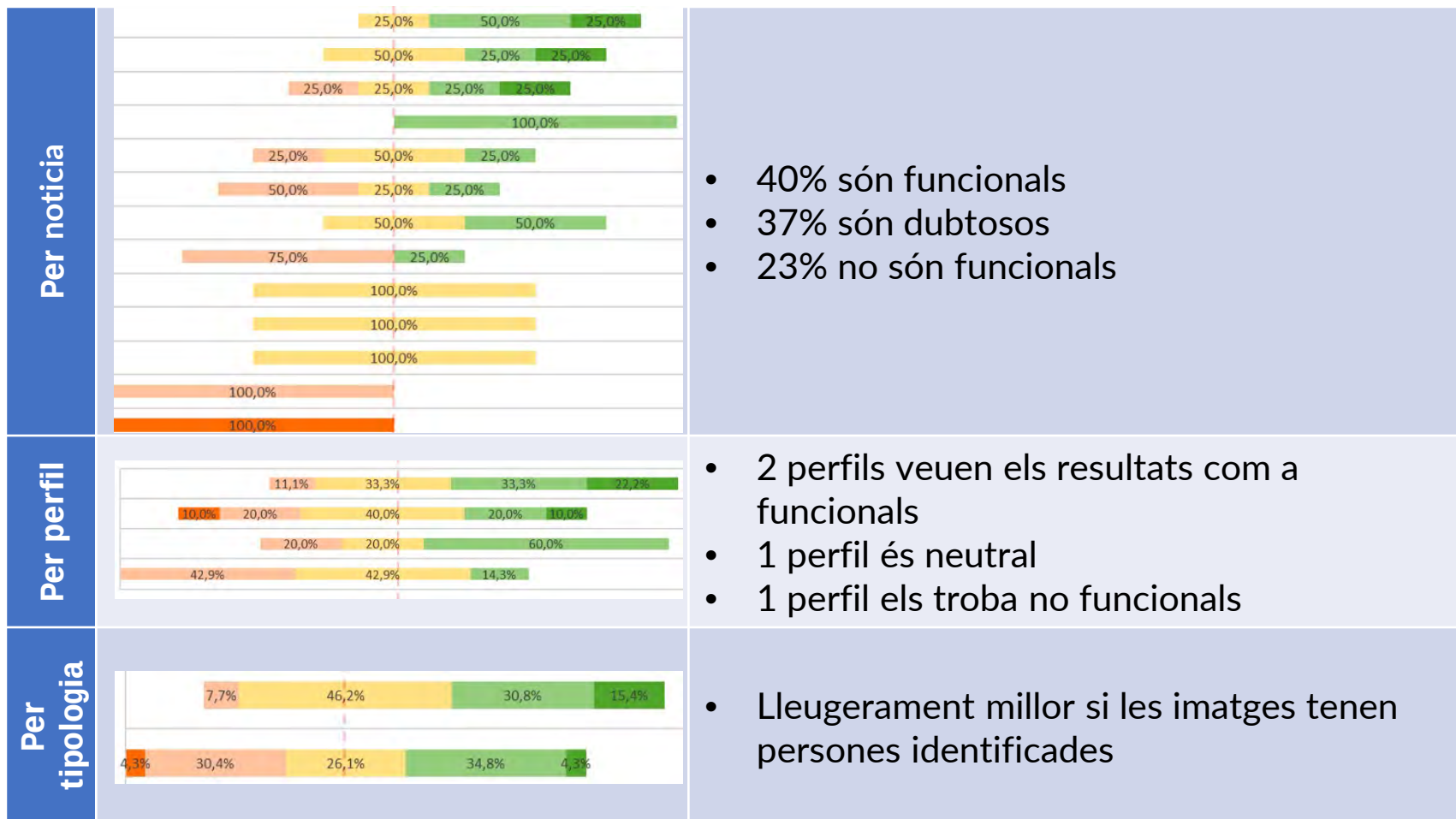
Text alternatiu



Peu de foto

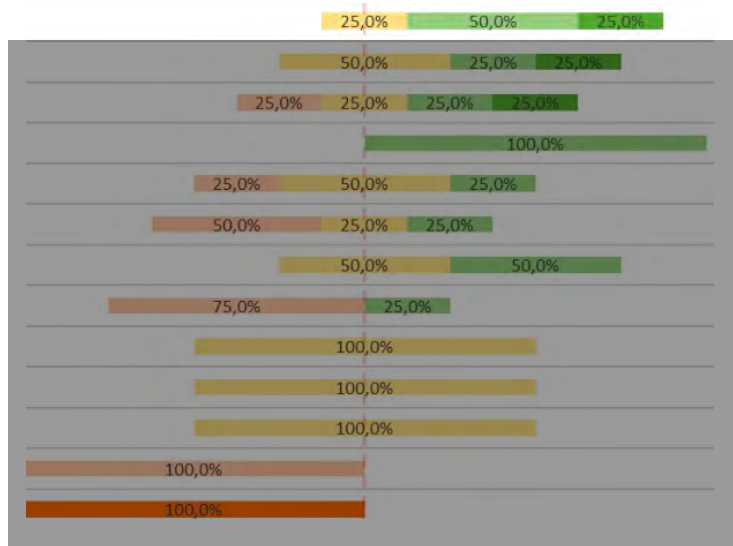


FUNCIONALITAT
Desglossat



FUNCIONALITAT

Més ben valorada



Josep Capella rep el Premi Josefina Carabias de periodisme parlamentari

El jurat n'ha valorat la trajectòria de 36 anys com a cronista



Peu de foto

Cerimònia d'entrega del Premi Josefina Carabias al periodista Josep Capella

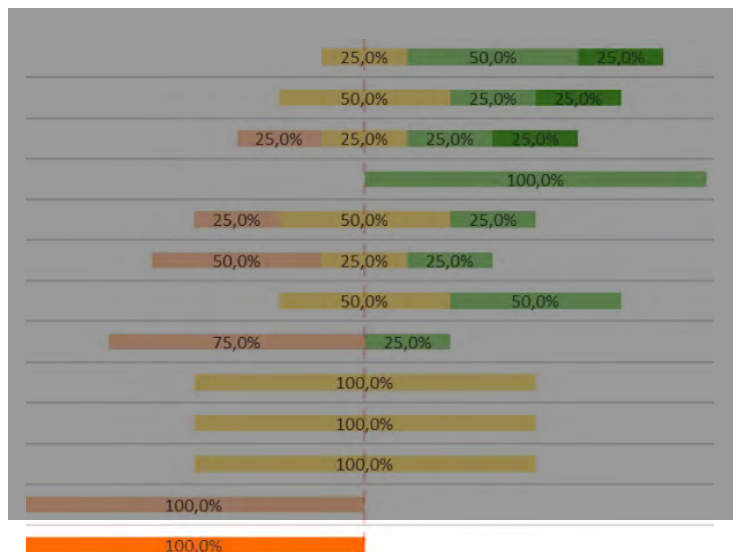
Avaluació del projecte

Text Alternatiu

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

FUNCIONALITAT

Pitjor valorada



Pedres gegants que cauen del cel: els meteoròlegs estudien un fenomen creixent a Catalunya

El Servei Meteorològic de Catalunya fa un estudi únic al món per preveure millor aquestes pedregades tan violentes



Peu de foto

Pedres de granit gegants comparades amb ous per a escala

Avaluació del projecte

Text Alternatiu

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.**

FIABILITAT

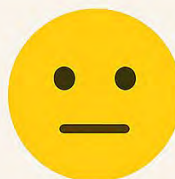
El text alternatiu o peu de foto es podria publicar
sense intervenció humana prèvia



Totalment
en desacord



Parcialment
en desacord



Neutre



Parcialment
d'acord



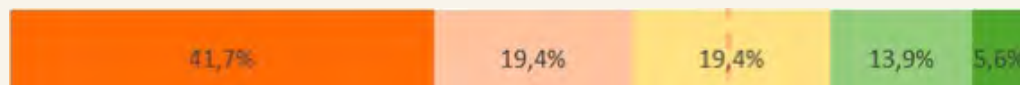
Totalment
d'acord



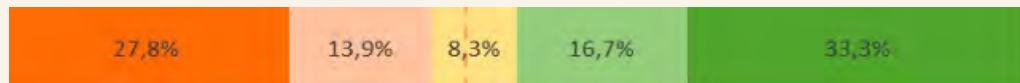
FIABILITAT

El text alternatiu o peu de foto es podria publicar sense intervenció humana prèvia

Text alternatiu



Peu de foto

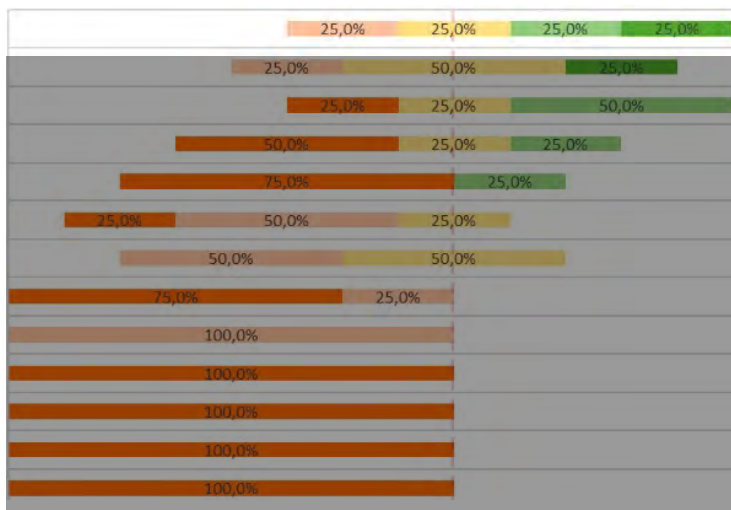


FIABILITAT
Desglossat

Per notícia	<table><tr><td>25,0%</td><td>25,0%</td><td>25,0%</td><td>25,0%</td></tr><tr><td>25,0%</td><td>50,0%</td><td>25,0%</td><td></td></tr><tr><td>25,0%</td><td>25,0%</td><td>50,0%</td><td></td></tr><tr><td>50,0%</td><td>25,0%</td><td>25,0%</td><td></td></tr><tr><td>75,0%</td><td>25,0%</td><td></td><td></td></tr><tr><td>25,0%</td><td>50,0%</td><td>25,0%</td><td></td></tr><tr><td>50,0%</td><td>50,0%</td><td></td><td></td></tr><tr><td>75,0%</td><td>25,0%</td><td></td><td></td></tr><tr><td>100,0%</td><td></td><td></td><td></td></tr><tr><td>100,0%</td><td></td><td></td><td></td></tr><tr><td>100,0%</td><td></td><td></td><td></td></tr><tr><td>100,0%</td><td></td><td></td><td></td></tr><tr><td>100,0%</td><td></td><td></td><td></td></tr></table>	25,0%	25,0%	25,0%	25,0%	25,0%	50,0%	25,0%		25,0%	25,0%	50,0%		50,0%	25,0%	25,0%		75,0%	25,0%			25,0%	50,0%	25,0%		50,0%	50,0%			75,0%	25,0%			100,0%				100,0%				100,0%				100,0%				100,0%				15% serien publicables sense intervenció 77% són clarament no publicables
25,0%	25,0%	25,0%	25,0%																																																			
25,0%	50,0%	25,0%																																																				
25,0%	25,0%	50,0%																																																				
50,0%	25,0%	25,0%																																																				
75,0%	25,0%																																																					
25,0%	50,0%	25,0%																																																				
50,0%	50,0%																																																					
75,0%	25,0%																																																					
100,0%																																																						
100,0%																																																						
100,0%																																																						
100,0%																																																						
100,0%																																																						
Per perfil	<table><tr><td>11,1%</td><td>33,3%</td><td>22,2%</td><td>11,1%</td><td>22,2%</td></tr><tr><td>50,0%</td><td>20,0%</td><td>30,0%</td><td></td><td></td></tr><tr><td>42,9%</td><td>42,9%</td><td>14,3%</td><td></td><td></td></tr><tr><td>60,0%</td><td>40,0%</td><td></td><td></td><td></td></tr></table>	11,1%	33,3%	22,2%	11,1%	22,2%	50,0%	20,0%	30,0%			42,9%	42,9%	14,3%			60,0%	40,0%				1 perfil amb una posició clarament contraria a la publicació																																
11,1%	33,3%	22,2%	11,1%	22,2%																																																		
50,0%	20,0%	30,0%																																																				
42,9%	42,9%	14,3%																																																				
60,0%	40,0%																																																					
Per tipologia	<table><tr><td>30,8%</td><td>15,4%</td><td>30,8%</td><td>15,4%</td><td>7,7%</td></tr><tr><td>47,8%</td><td>21,7%</td><td>13,0%</td><td>13,0%</td><td>4,3%</td></tr></table>	30,8%	15,4%	30,8%	15,4%	7,7%	47,8%	21,7%	13,0%	13,0%	4,3%	Lleugerament més fiable quan hi ha persones identificades al text alternatiu																																										
30,8%	15,4%	30,8%	15,4%	7,7%																																																		
47,8%	21,7%	13,0%	13,0%	4,3%																																																		

FIABILITAT

Més ben valorada



Els filtres de bellesa fan semblar menys llesta: com TikTok reforça els estereotips de gènere

Un estudi independent mostra l'impacte cognitiu dels filtres de bellesa que fan que les dones siguin percebudes com a menys intel·ligents que els homes



Peu de foto

Persona interaccionant amb aplicacions de xarxes socials al telèfon mòbil

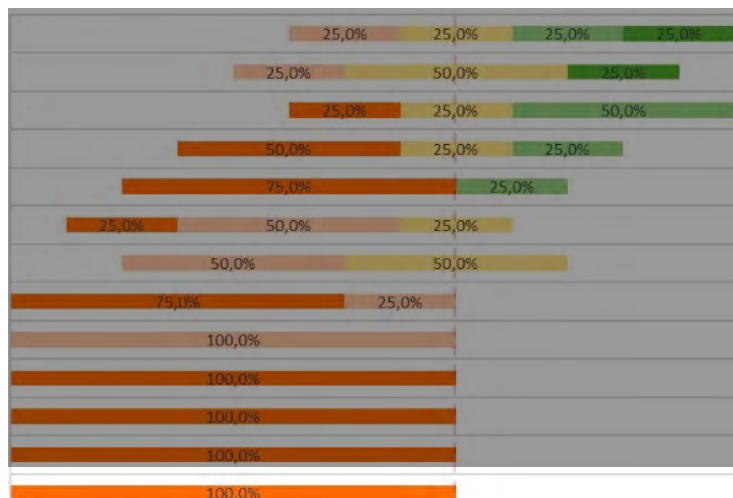
Avaluació del projecte

Text Alternatiu

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.

FIABILITAT

Pitjor valorada



La mesa ajorna la decisió sobre la proposta de Junts per una qüestió de confiança a Sánchez

La Mesa del Congrés vol estudiar a fons la proposta, que podria generar doctrina, i no deliberarà fins al febrer



Peu de foto

Dona parlant en assemblea política dels Països Catalans

Avaluació del projecte

Text Alternatiu

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.

PERSONALITZACIÓ

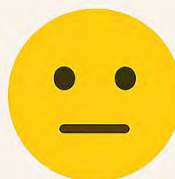
L'eina proveeix de totes les funcionalitats per poder aconseguir un resultat totalment adaptat a les meves necessitats



Totalment
en desacord



Parcialment
en desacord



Neutre



Parcialment
d'acord

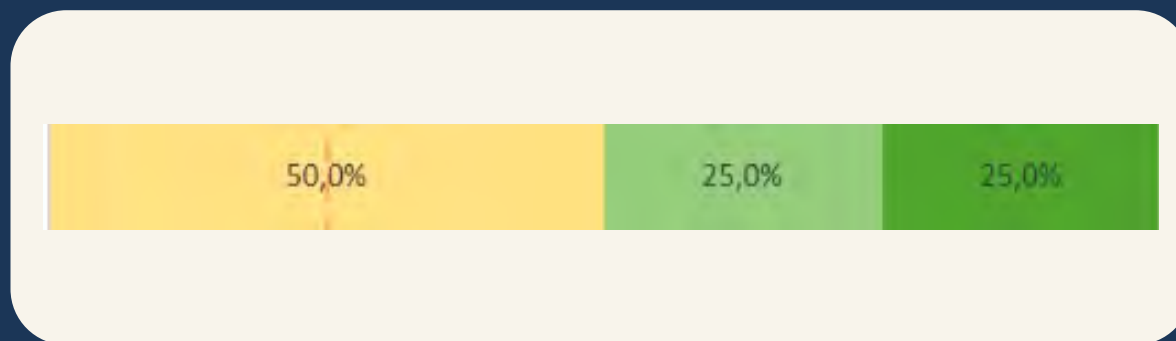


Totalment
d'acord



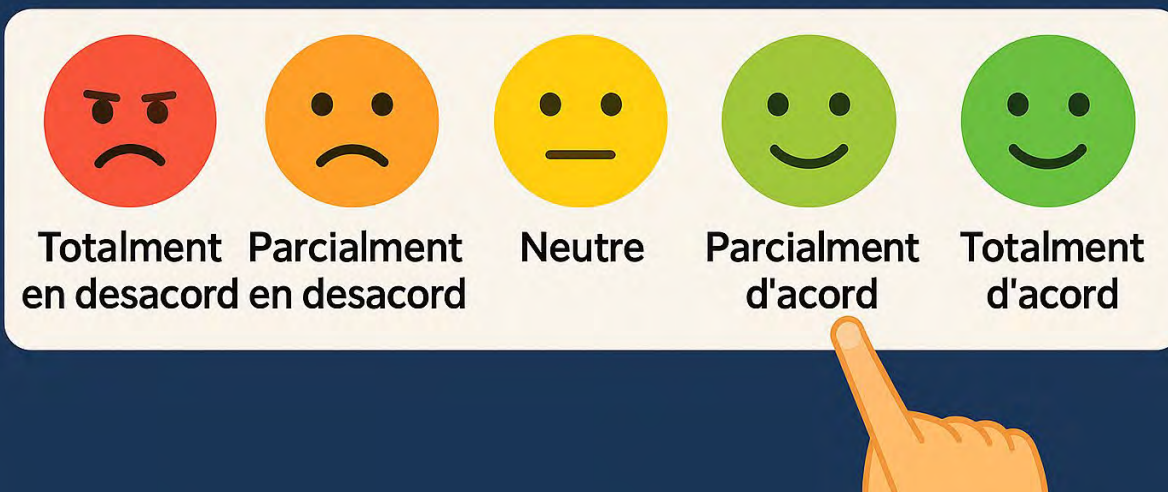
PERSONALITZACIÓ

L'eina proveeix de totes les funcionalitats per poder aconseguir un resultat totalment adaptat a les meves necessitats



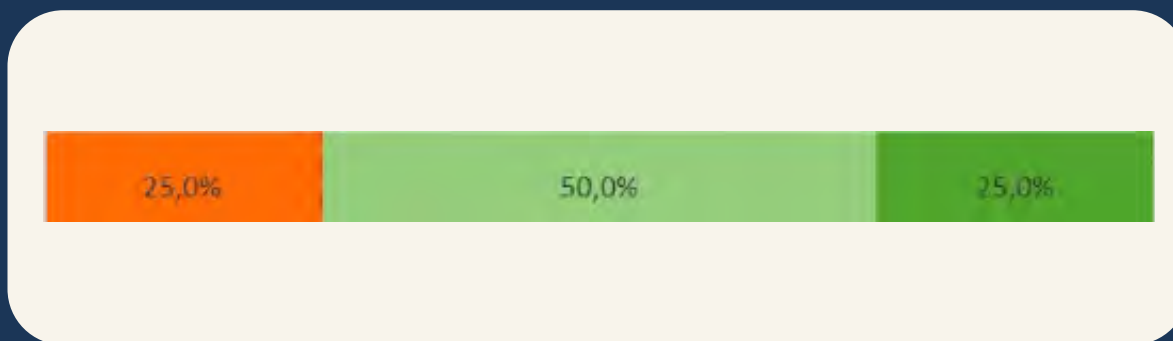
FACILITAT

L'eina te una interfície intuïtiva i fàcil
d'utilitzar



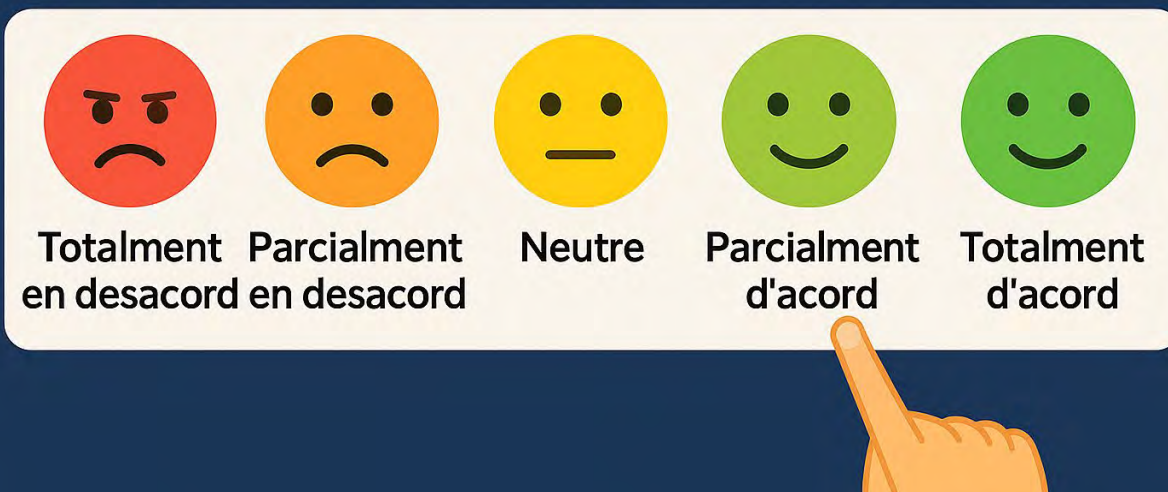
FACILITAT

L'eina te una interfície intuïtiva i fàcil
d'utilitzar



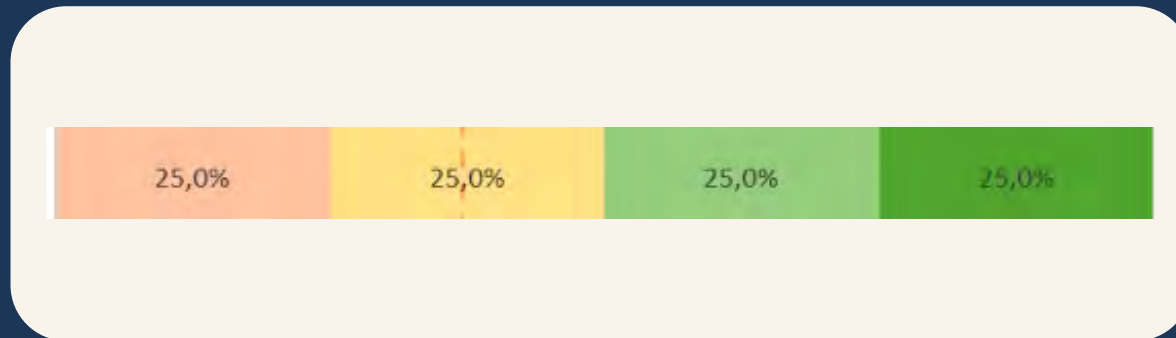
RENDIMENT

La generació del text alternatiu i peu de foto en
1 minut és un temps totalment acceptable



RENDIMENT

La generació del text alternatiu i peu de foto en 1 minut és un temps totalment acceptable

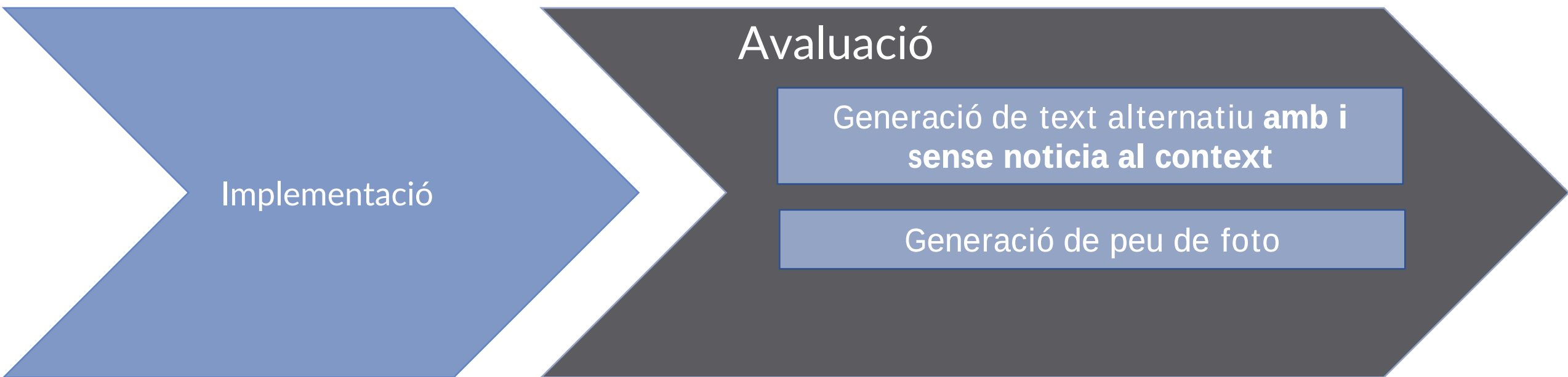


Conclusions

- L'estat actual no ens permet publicar automàticament el text alternatiu... ni tampoc els nostres principis
- Els resultats que inclouen persones identificades són lleugerament millor valorats, tant pel que fa a la funcionalitat com a la fiabilitat
- Els errors lingüístics impacten molt negativament a l'avaluació
- Tenim més clar quina eina volem: Un assistent que generi propostes de text alternatiu de forma automàtica en 1 minut
- Afegir la notícia al context sembla que millora el resultat



**Afegir la notícia
al context
sembla que
millora el
resultat**



Qui?

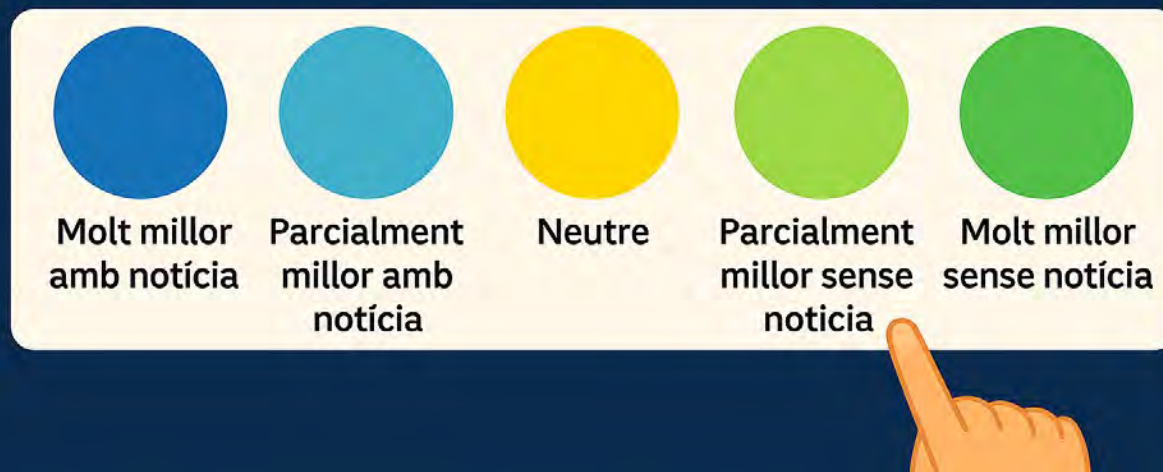
- Dues àrees de l'organització
 - Redacció informatius
 - Accessibilitat
- Usuaris reals de lectors de pantalla

Com?

- Formulari amb exemples de resultats generats amb i sense la notícia al context
- Es corregeixen els errors lingüístics per focalitzar l'avaluació

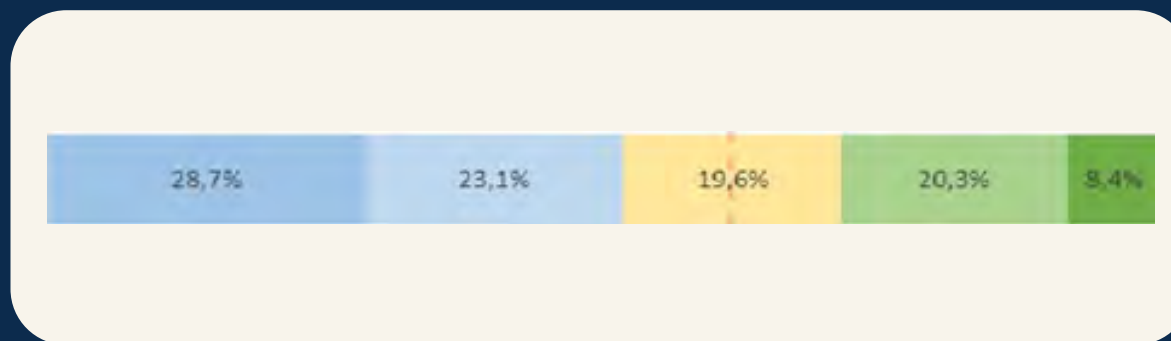
AMB O SENSE
NOTÍCIA?

Quina de les dues alternatives de generació de
text alternatiu et sembla millor



AMB O SENSE
NOTÍCIA?

Quina de les dues alternatives de generació de
text alternatiu et sembla millor



AMB O SENSE NOTÍCIA?



Els filtres de bellesa fan semblar menys llesta: com TikTok reforça els estereotips de gènere

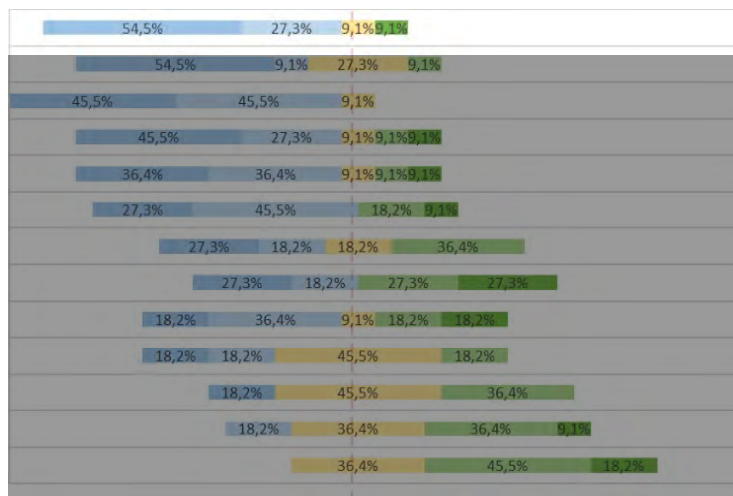
Un estudi independent mostra l'impacte cognitiu dels filtres de bellesa que fan que les dones siguin percebudes com a menys intel·ligents que els homes

Amb notícia al context

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 sueter blanc i probablement **està obrint l'aplicació de TikTok**, ja que està prominentament exhibida a la pantalla

Sense notícia al context

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**



AMB O SENSE NOTÍCIA?

Les millors valorades



Troben a Mallorca l'avantpassat més antic dels mamífers: un gorgonop, un carnívor que ponia ous

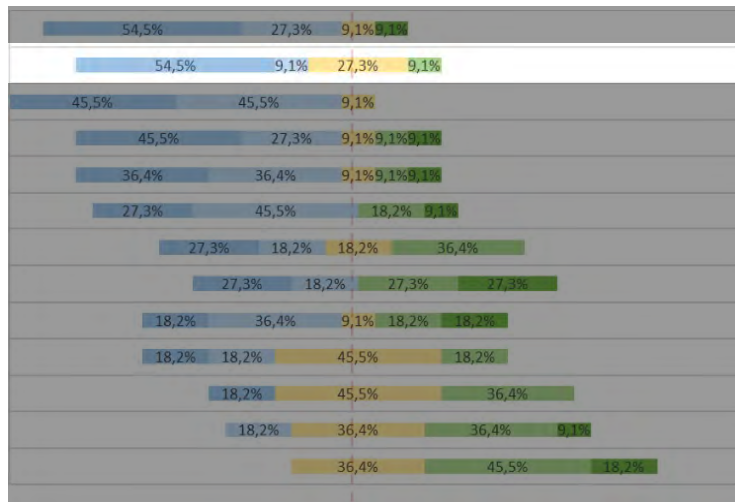
Un grup de recerca internacional liderat per l'Institut Català de Paleontologia Miquel Crusafont i el Museu Balear de Ciències Naturals troba l'avantpassat dels mamífers més antic del món

Amb notícia al context

Una il·lustració detallada d'un **gorgonops**, un antic rèptil similar a un mamífer, representat prop d'un llac temporal. Té un cos robust, potes curtes i un cap ampli amb dents de sabre prominents, semblant a un gos de mida mitja. Està situada en un entorn sec i àrid amb un terreny arenós 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**.

Sense notícia al context

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 rogenc-fosc. L'animal sembla un Paleofiber, com s'indica en el text acompanyant. El llit del riu presenta textures que suggereixen fang sec i patjades febles, suggerint el pas del temps i d'altres criatures. **L'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.



AMB O SENSE NOTÍCIA?

Les millors valorades



Josep Capella rep el Premi Josepina Carabias de periodisme parlamentari

El jurat n'ha valorat la trajectòria de 36 anys com a cronista

Amb notícia al context

Una cerimònia formal on **el periodista Josep Capella Matas** està darrere d'un podi, rebent 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 el bust d'una escultura. 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.

Sense notícia al context

Un esdeveniment formal que té lloc en un saló ornamentat amb motius vermells i daurats. 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 vestit 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 mocador blanc, i l'altre té cabell fosc. **Una gran bandera espanyola es mostra de manera prominent a la dreta de la imatge, afegint un context del país dels Països Catalans a la escena.**

AMB O SENSE NOTÍCIA?

La pitjor valorada



Locals comercials convertits en habitatge: què diu la normativa

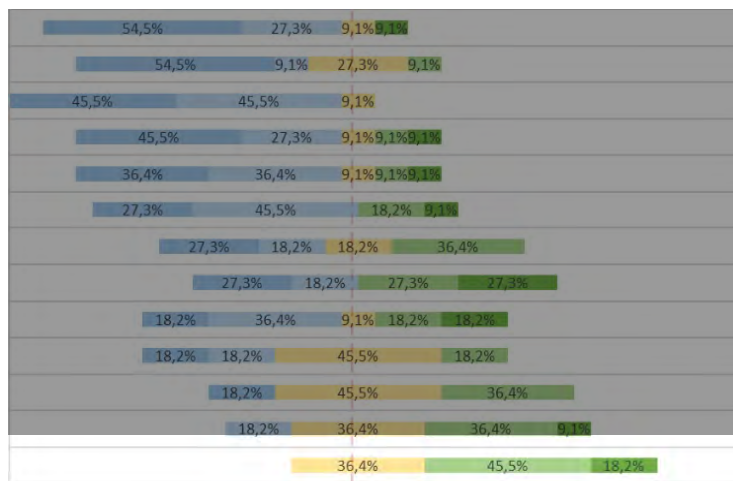
Davant la crisi de l'habitatge, propietaris de locals tancats aposten per la seva conversió en habitatges, però topen amb una normativa exigent i una tramitació dels permisos lenta

Amb noticia al context

La imatge mostra l'exterior d'un edifici comercial amb dos aparadors de comerç. Les persianes dels dos aparadors estan tancades, i el grafiti està prominentament exhibit en elles. L'edifici té una façana grisa amb una secció superior de maons vermells. Hi ha un arbre a la dreta de la imatge, al costat d'un contenidor. Una caixa elèctrica està fixada a l'edifici prop de la porta d'entrada. El carrer està pavimentat amb maons, i hi ha un pas de vianants davant de l'edifici. Aquesta escena sembla estar en una zona urbana dels Països Catalans, **reflectint la tendència creixent de l'ocupació de propietats comercials com a espais residencials.**

Sense noticia al context

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 una bústia de correus, amb un fanal davant de l'establiment. Els grafitis semblen estar en un estil ombrós i il·legibles, afegint un context urbà a l'escena.



Conclusions

- L'estat actual no ens permet publicar automàticament el text alternatiu... ni tampoc els nostres principis
- Els resultats que inclouen persones identificades són lleugerament millor valorats, tant pel que fa a la funcionalitat com a la fiabilitat
- Tenim més clar quina eina volem: Una eina senzilla que generi propostes de text alternatiu de forma automàtica en 1 minut
- Els errors lingüístics impacten molt negativament a l'avaluació
- Afegir la notícia al context ~~sembla que~~ millora el resultat... però s'hauria de limitar a la descripció de la imatge no afegir més informació ni suposicions al text generat
- L'aplicació del llibre d'estil també està portant a interpretacions o suposicions en l'àmbit geogràfic... Tot succeeix als països catalans.
- Opinió del usuaris: Els texts alternatius han de ser curts i precisos
- Línia d'evolució: Pot ser la base per tenir una eina que proposi una imatge del nostre arxiu fotogràfic per il·lustrar una notícia

Preguntas

Cloenda



Sr. Joan Mas i Albaigès

*Director, **CIDAI***

*Director Científic Digital, **Eurecat***

**Gràcies per la vostra
assistència**

CIDAI

Centre of Innovation
for Data tech
and Artificial Intelligence

Socis promotors



Membres



Presentació del Projecte
**Assistència intel·ligent a la
descripció contextualitzada de
fotos per notícies**

5 juny 2025

www.cidai.eu