

Repensem l'avaluació amb la IA: reptes i oportunitats

39es Jornades per a professors de català



16 de juliol de 2025 – Universitat Jaume I - Castelló

Sònia Sáez

Departament d'Educació - Generalitat de Catalunya

ssaez@xtec.cat

Objectius del taller

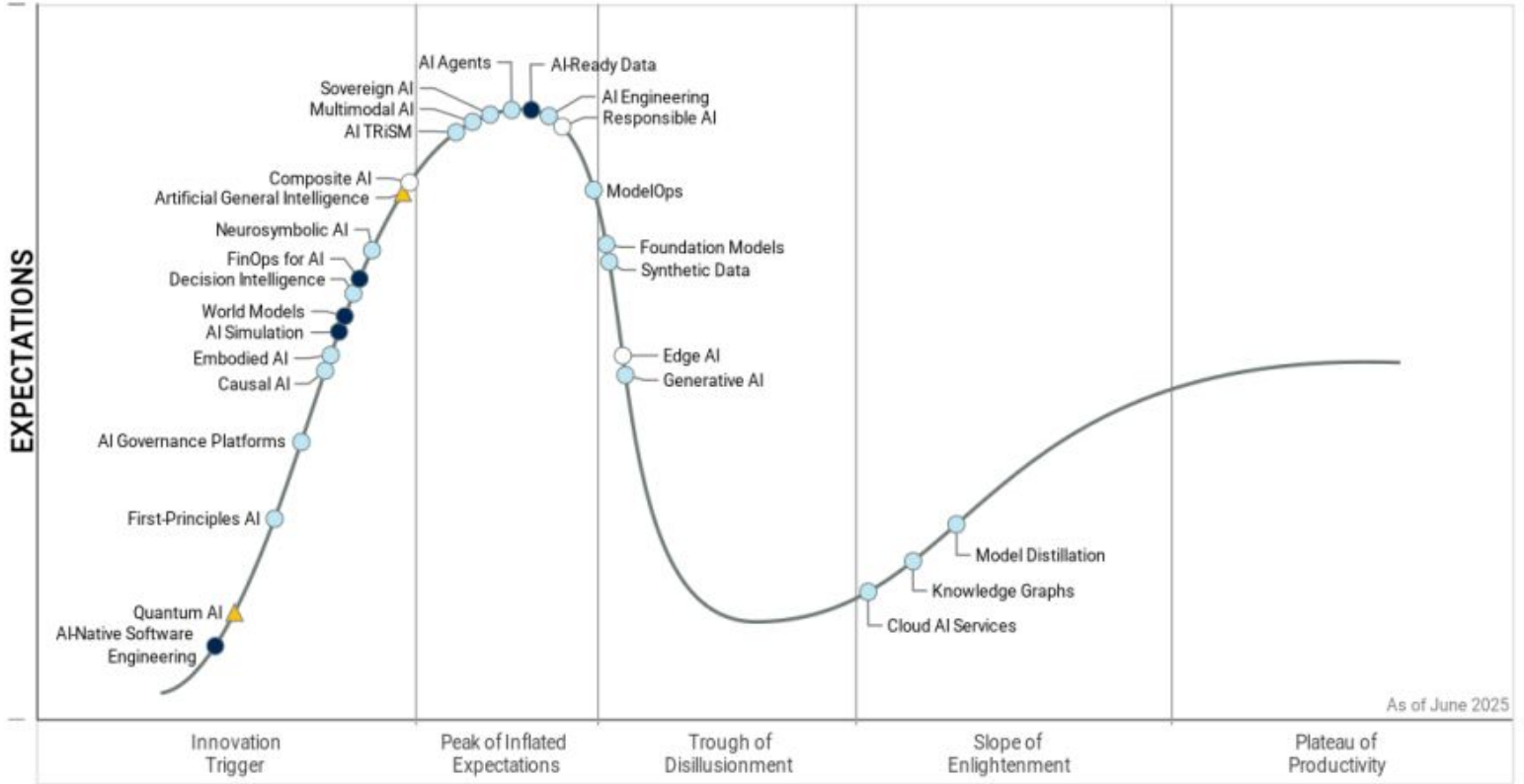
- **Analitzar on som i on anem amb la IA respecte l'avaluació.**
- **Explorar les funcions de la IA en la correcció de textos escrits i orals.**
- **Reflexionar sobre els reptes i oportunitats de la IA en l'avaluació.**
- **Replantejar el disseny de tasques d'avaluació 'resistents' a la IA.**

Estructura del taller

- Introducció: consideracions sobre la IAG per l'avaluació
- Activitats:
 - Bloc I – Expressió escrita
 - Bloc II – Expressió oral
 - Bloc III – Tasques d'avaluació “resistents” a la IA
- Conclusions: la IAG com a assistent en l'avaluació.

On som
i on
anem
amb la
IA

Hype Cycle for Artificial Intelligence, 2025



As of June 2025

Plateau will be reached: ○ <2 yrs. ● 2-5 yrs. ● 5-10 yrs. ▲ >10 yrs. ✗ Obsolete before plateau

Consideracions prèvies a l'ús de la IA generativa

- La seva evolució vertiginosa.
- La seva base probabilística.
- Els seus biaixos de gènere, raça i cultura.
- La seva tendència a inventar contingut (al·lucinacions).
- El perill de l'erosió del pensament crític i de la pèrdua d'habilitats.

Calculant la base predictiva de la IA generativa

“Et donaré una frase inacabada, a la qual li faltará l'última paraula, i vull que em donis un parell de suggeriments de quina seria aquesta última paraula.

Genera les paraules en ordre, començant per la més probable, juntament amb la probabilitat estimada per a cadascuna.”

1. Els nens van sortir de l'escola per ...
2. Els meus alumnes van acabar la classe molt ...
3. No en tenia prou amb el que em va dir així ...
4. Tot es va solucionar en un obrir i ...
5. Tot i haver-li explicat les raons, ell no va ...

Fraser	Copilot	Grok	ChatGPT
Els nens van sortir de l'escola per ...	anar (45%) jugar (35%)	tornar (70%) jugar (20%)	tornar (70%) jugar (25%)
Els meus alumnes van acabar la classe molt ...	contents (40%) cansats (35%)	cansats (65%) contents (25%)	cansats (65%) contents (30%)
No en tenia prou amb el que em va dir així ...	que (60%) doncs (30%)	que (80%) com (15%)	també (70%) doncs (25%)
Tot es va solucionar en un obrir i ...	tancar (95%) plegar (5%)	tancar (90%) fer (5%)	tancar (90%) parpallejar (8%)
Tot i haver-li explicat les raons, ell no va ...	entendre (45%) cedir (35%)	entendre (75%) acceptar (20%)	entendre (65%) acceptar (30%)

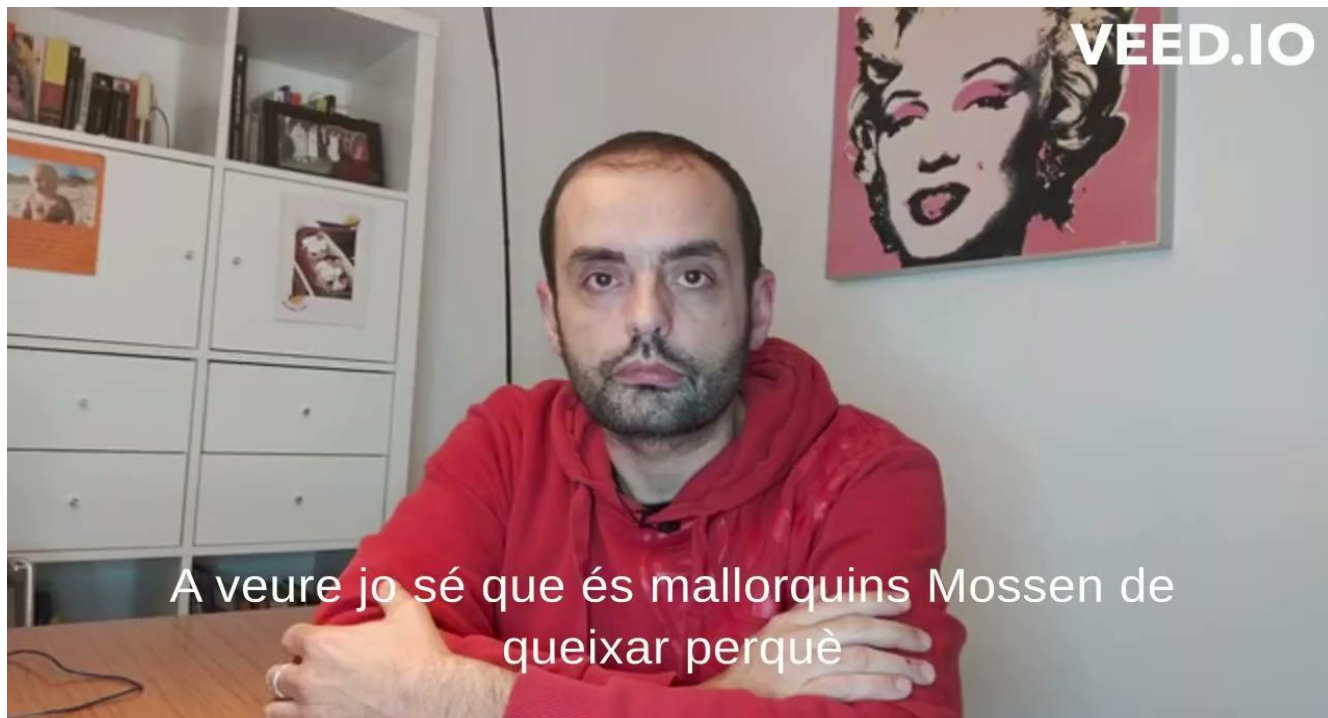
Formulació d'indicacions (*prompts*) amb la IAG

- **Les indicacions han de ser molt específiques i requereixen iteracions.**
- **Els resultats generats varien amb cada iteració.**
- **Els resultats han de ser verificats per detectar explicacions errònies i omissions.**
- **La IA fa propostes automàtiques de noves indicacions amb cada resultat.**

Tasca de comprensió oral – què podem fer amb els subtítols?

Identificar i corregir els errors més evidents dels subtítols automàtics (errors de transcripció, gramàtica, puntuació, o incoherències).

Debatre: Per què la IA comet aquests errors? Com pot aquesta activitat, on l'alumne corregeix la IA, ser una tasca d'avaluació formativa per a la comprensió oral i per a l'atenció als detalls lingüístics? Com fomenta l'autonomia de l'alumne?



Des de l'arribada de la IA generativa...

- Què està canviant en l'avaluació de llengües?
- Quins reptes planteja als docents?
- Quines oportunitats tenim a l'abast?
- Quins aspectes de l'avaluació creus que la IA pot fer millor que tu?

Reptes d'avaluar amb la IA

Biaixos en l'entrenament i en la resposta



Superficialitat en el judici lingüístic



Desconnexió amb el context educatiu



Manca d'infraestructura i equitat tecnològica



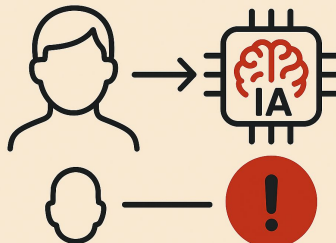
Tractament inadequat dels errors i la retroacció



Deslegitimació del rol docent



Dependència o delegació excessiva

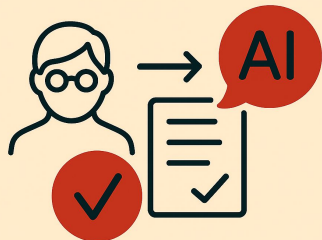


Privacitat i tractament de dades de l'alumnat



Oportunitats d'avaluació amb la IA

Suport a la funció docent



Estalvi de temps en la correcció



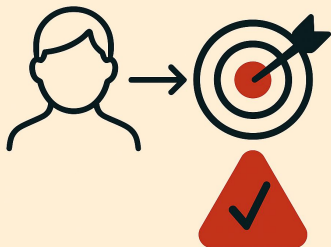
Retroacció més extensa i immediata



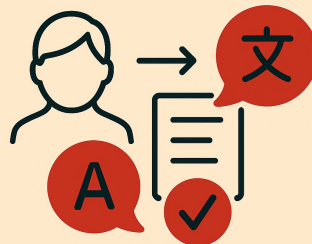
Personalització i iteració de la retroacció



Millora de l'objectivitat



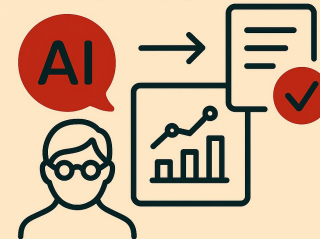
Adaptabilitat multilingüe



Gestió eficient de grans volums de dades



Visualització del progrés i generació d'informes



Bloc I: Avaluació de l'expressió escrita

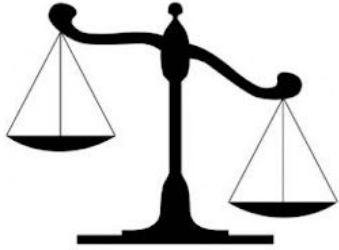
Activitat 1 – Mediació Escrita – Nivell C2

Qüestions pel debat:

- Quina retroacció ha donat la IA?
- Quin tipus d'errors ha detectat?
- Quines omissions hi trobem?
- Què faria el docent que la IA no fa?



Correcció automàtica vs. retroacció del professorat



Els sistemes automàtics guanyen en la detecció d'errors: **ortogràfics, de puntuació, gramaticals, i les propostes de sinònims.**

L'humà guanya en la detecció de l'**assoliment de la tasca, inconsistències del contingut, subtileses, trets socioculturals, coherència i estructura del discurs.**

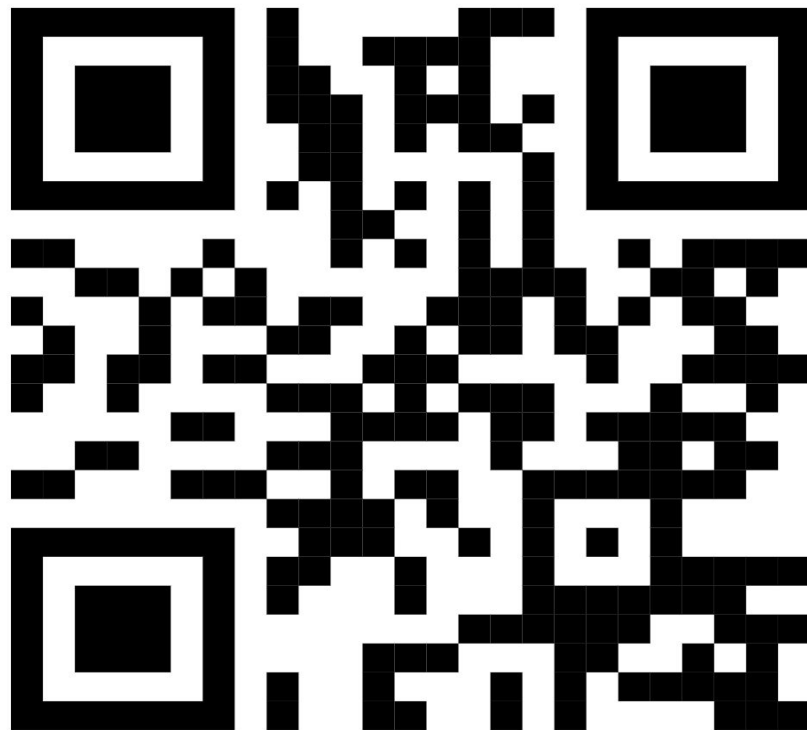
Pu, Jianmei: *Learner Uptake with automated Feedback and Teacher Feedback in an online OFL Writing Context*, 2023

Bloc I: Avaluació de l'expressió escrita

Activitats 2 i 3 – Dos estudis de cas

Qüestions pel debat:

- En l'avaluació d'aquestes tasques, la IA ho ha fet bé?
- Aquestes tasques tradicionals promouen el pensament crític?
- En aquestes tasques: avaluem l'estudiant o l'eina?



Bloc I: Avaluació de l'expressió escrita



Activitat 4 – Rúbriques d'avaluació

- Quina qualitat tenen les rúbriques generades?
- Són útils?
- Quins aspectes (complexitat sintàctica, ventall lèxic) són difícils d'avaluar per la IA?



Bloc II: Avaluació de l'expressió oral

Diàlegs generats per la IA - *ElevenLabs*

- versió 1 
- versió 2 

Podem utilitzar la IA per crear models d'interacció oral, però com podem utilitzar-los per subratllar la importància d'elements com la prosòdia i la producció d'un llenguatge més 'natural'?

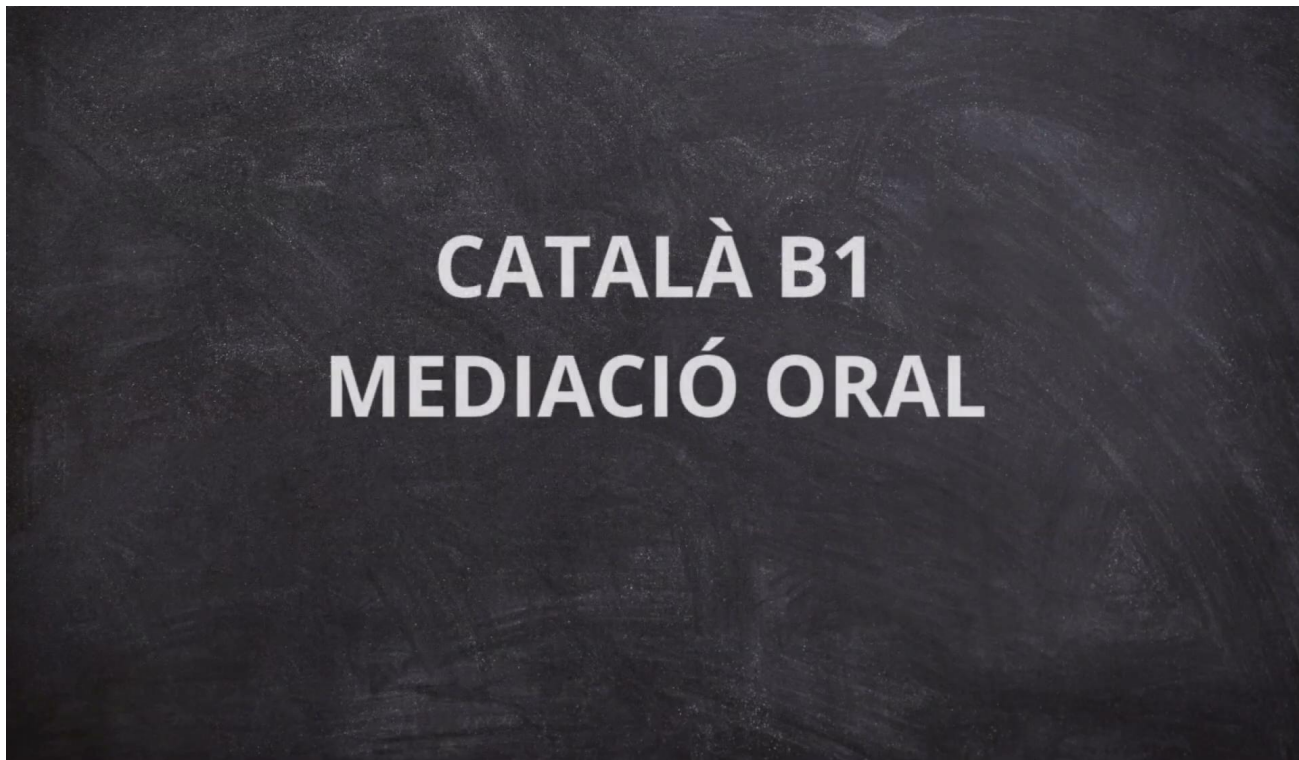
TikTok clip fet amb [Veed.io](https://veed.io)

- Aquest vídeo ha estat creat a partir del diàleg de la conversa anterior, amb la eina [AI Gen](https://veed.io), que transforma el diàleg en un monòleg.
- Poden triar-se quines imatges són les més adients per a cada imatge.
- [Veed.io](https://veed.io) genera el vídeo a partir d'aquestes indicacions.



Bloc II: Avaluació de l'expressió oral

Com donar retroacció als alumnes amb els subtítols o transcripció de la IA?



Subtítols fets amb [OpusClip](#)



Anàlisi de les emocions del diàleg amb *ChatGPT*

Anna:

- *Emocions predominants: frustració, tristesa, solitud, necessitat d'afecte.*
- Se sent ignorada i desatesa, però vol recuperar la connexió emocional amb l'Èric. Expressa els seus sentiments de manera clara, però busca comprensió i una resposta empàtica.

Èric:

- *Emocions predominants: cansament, defensa, culpa, voluntat de millora.*
- Se sent pressionat i incomprès, però quan pren consciència del dolor de l'Anna, mostra empatia i predisposició a canviar. Comença defensiu, però acaba assumint responsabilitat i oferint una solució.

Fins a quin punt podem confiar en la IA per "avaluar emocions" o aspectes subtils de la comunicació oral?

Bloc III: Tasques d'avaluació “resistents” a la IA



El nou paradigma: Multimodalitat i Integració de Destreses

- Com la IA ens permet dissenyar tasques que **integren diverses destreses** (llegir per parlar, escoltar per escriure, etc.) i **formats** (text, àudio, vídeo, imatge).
- Com l'avaluació ja no és lineal sinó **més complexa i holística**.

Replantejament de les tasques d'avaluació amb la IA

Cal dissenyar tasques que:

- Els estudiants no puguin resoldre sense el seu propi criteri o pensament.
- Permetin observar el procés de l'elaboració i no només el producte.
- Integrin la IA com eina formativa per als estudiants.

Com integrar la IA d'una manera efectiva en la pràctica avaluadora?

- 1) Contextualitza les tasques en experiències personals**
(opinions, vivències, projectes locals)
- 2) Fes que la producció final sigui només una part del procés**
(prèvia lectura, discussió, etc.)
- 3) Demana que mostrin el procés, no només el resultat** (ex: esborranys, notes, revisions)
- 4) Inclou interacció real o retroacció en viu** (oral o per escrit)

REFLEXIONS FINALS

- 1) Després d'aquestes activitats, com **repenseu el vostre rol** en l'avaluació del català a l'era de la IA?
- 2) Quins són els aspectes que **la IA no podrà mai avaluar de manera efectiva** i que continuaran sent el valor diferencial del professor humà?
- 3) Com podem educar els nostres alumnes per fer **un ús ètic i responsable de la IA com a eina d'aprenentatge i autoavaluació**, no com a substitut de l'esforç?

Gràcies per la vostra participació!!!



Bibliografia I

Jessie S. Barrot (06/2021). [Using automated written corrective feedback in the writing classrooms: effects on L2 writing accuracy.](#)

Qing-Ke Fu. (02/2022) [A review of AWE feedback: types, learning outcomes, and implications.](#)

Atsushi Mizumoto a, Masaki Eguchi. (08/2023) [Exploring the potential of using an AI language model for automated essay scoring.](#)

Uttkarsha Bhosale,(June 27, 2023). ["Limitations of DeepL Write and ChatGPT in Scholarly Editing: How human editors overpower AI tools."](#) Enago Academy.

Rüdian, Dittmeyer, Pinkwart (03/2022) [Challenges of using auto-correction tools for language learning.](#)

Rüdian, Dittmeyer, Pinkwart: [Using Data Quality to compare the Prediction Accuracy based on diverse annotated Tutor Scorings.](#)

Blanka Klímová, Prodhah Mahbub Ibna Seraj. (03/2023) [The use of chatbots in university EFL settings: Research trends and pedagogical implications.](#)

Jianmei Pu. (04/2023) [Learner Uptake with Automated Feedback and Teacher Feedback in an Online EFL Writing Context.](#)

Bibliografia II

Hernández-Leo, Davinia, [Com ho avaluem? Repensem la universitat: impacte de la intel·ligència artificial \(IA\) en l'aprenentatge. Taula interuniversitària per a repensar l'experiència d'aprenentatge a la universitat a Catalunya](#), Barcelona, 13 juliol 2023.

Martínez Belchí, Evaristo; [La irrupción de la Inteligencia Artificial: una perspectiva práctica de sus posibilidades para la elaboración de pruebas de evaluación de lenguas extranjeras](#)

Mena Octavio, Manuela; González Argüello, Vicenta; Pujolà, Joan-Tomàs; [Guia ChatGPT ELE](#)

ROMÁN MENDOZA, Esperanza; [Formular preguntas para comprender las respuestas: ChatGPT como agente conversacional en el aprendizaje de español como segunda lengua](#)

SHARPLES, Mike. [Towards social generative AI for education: theory, practices and ethics](#), 2023.

[Exemples de Prompts per a professors d'ELE](#)