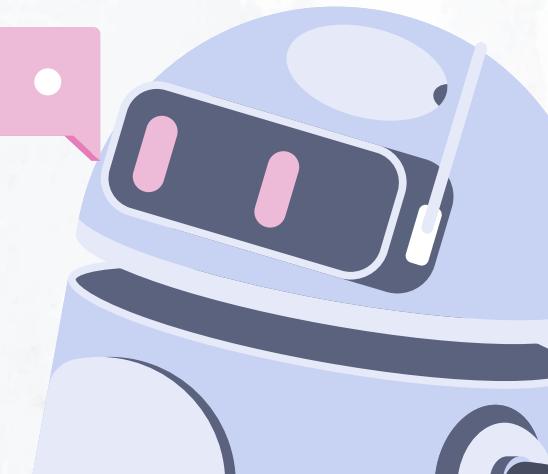
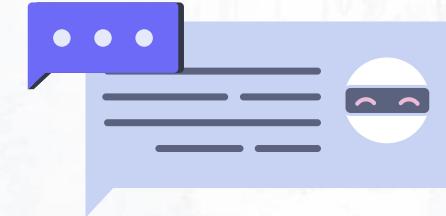


# Resultados y aprendizajes del uso de IA en la elaboración de materiales de lectura en el marco del proyecto LT-LiDER

Mireia Vargas Urpí



# Language and Translation: Literacy in Digital Environments and Resources

UAB, Dublin City University, TH Köln, Universität Wien, Université Grenoble Alpes,  
Euskal Herriko Unibertsitatea, Rijksuniversiteit Gröningen

LT-LiDER (ref. 2023-1-ES01-KA220-HED-000161531) has been funded with support from the European Commission. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



Co-funded by  
the European Union

# Índice

- 01 → ¿Qué dicen los estudios previos sobre el tema?
- 02 → Objetivos de nuestro estudio
- 03 → Metodología: redactar el *prompt*, evaluar los resultados
- 04 → Resultados de nuestro estudio
- 05 → A modo de conclusión

01 →

¿Qué dicen los estudios  
previos sobre el tema?

(IA)

# ¿La revolución de la IA?



## «Gran potencial»

ChatGPT tiene el **potencial** de mejorar la productividad del profesorado y ayudar en el desarrollo de materiales educativos (Li Belle et al., 2024: 12)

## ... pero faltan estudios prácticos

Hay un desequilibrio entre la investigación basada en percepciones sobre el potencial y aquella que se basa en resultados medibles en la práctica (Stockwell, 2024: 5)

## y, sobre todo, basados en más lenguas y más niveles

Hay un sesgo hacia el inglés, y es necesario estudiar «a wider range of language proficiency levels, learning objectives and educational levels» (Li Belle et al., 2024: 12-13)

# Breve panorámica

## (\*) ChatGPT como tutor vs. pensamiento crítico

Las IA pueden producir errores, por eso es interesante organizar actividades que desarrollen el pensamiento crítico de los alumnos o les hagan reflexionar sobre esto (Rovira-Esteva, 2025; Zhao et al., 2024).

## (\*) Estudios sobre competencias específicas

Importancia de la revisión humana (por parte del docente) en la creación de actividades, importancia de un *prompt* efectivo para mejores resultados.

## (\*) Sesgos e influencia del inglés

La mayoría de los modelos están entrenados con textos en inglés y esto puede provocar sesgos y/o resultados menos naturales en otras lenguas.

02 →

# Objetivos de nuestro estudio\*

\*Casas-Tost, Helena; González-Torrents, Isabel; Rovira-Esteva, Sara; Vargas-Urpí, Mireia (en prensa).

(IA)

# Objetivos

1. Evaluar la utilidad de la IAG como asistente del profesor de CLE.
2. Evaluar el potencial de los textos generados con IAG para el desarrollo de actividades de comprensión lectora para alumnos de nivel A2 (más o menos HSK3).
3. Comparar el desempeño de dos modelos de IAG (ChatGPT-4o y DeepSeek) para actividades de comprensión lectora.
4. Desarrollar una rúbrica para la evaluación de actividades generadas con IAG para determinar su usabilidad en la enseñanza de CLE.

# Premisas importantes

- Un texto para comprensión lectora no siempre es un texto «natural», porque tiene que ceñirse al nivel del alumnado.
- En el caso del chino, la opacidad del vocabulario desconocido limita mucho la posibilidad de introducir palabras nuevas en este tipo de ejercicios.
- Las preguntas deben poder evaluar la comprensión del texto. Si repiten el vocabulario o estructura del texto, son demasiado fáciles.

# 03 Metodología →

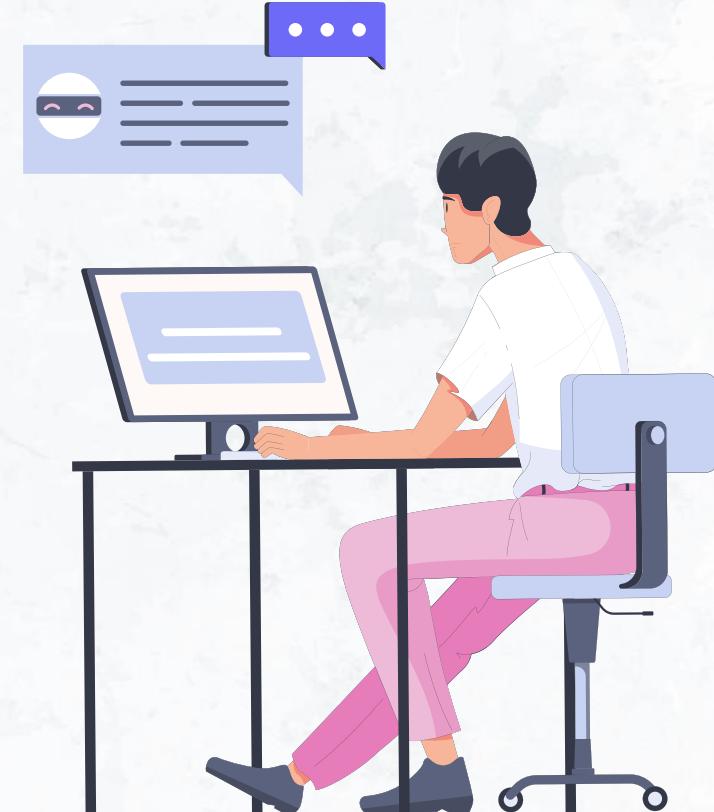
Primer paso:  
redactar el *prompt*

# Y, tú, ¿qué prompt usarías?

Pide a ChatGPT una actividad de comprensión lectora para tu clase, introduce tu *prompt* aquí:



<https://www.menti.com/alpsyj6ftnn>



# Consejos para el *prompting*

## (1) Cuanto más detallado mejor

Estrategias:

- Contextualizar / situar el *prompt*. Por ejemplo: «Eres un profesor de chino».
- Indicar la finalidad del texto (actividad de comprensión lectora), el nivel del alumnado, el tema y una extensión aproximada (aunque no cuenta los caracteres).
- Evitar humanizar al modelo (p. ej., «por favor») para evitar información innecesaria en el *prompt*.

## (2) Chain-of-thought

Consiste en «dialogar» con la IA a partir de la primera respuesta que genera. Por ejemplo: «Revisa el texto y utiliza solamente vocabulario del nivel HSK3».

# Ejemplo de uno de nuestros *prompts* \*

Eres un profesor de chino como lengua extranjera. Redacta un texto en chino para crear una actividad de comprensión lectora de nivel básico (correspondiente a A2 del CEFR o nivel 3 del HSK) sobre la rutina de un estudiante de intercambio en una universidad de Pekín, donde estudia chino. El texto debe incluir un mínimo de tres párrafos y tener suficiente contenido para poder hacer posteriormente una actividad de comprensión escrita. Asegúrate de ceñirte al nivel básico.

03 →

## Segundo paso: evaluar el resultado

(IA)

# Evalúa los textos generados

Lee los dos textos que encontrarás en el documento adjunto.

- ¿Cuál crees que se generó con ChatGPT y cuál crees que se generó con DeepSeek?
- ¿Qué te parecen estos textos?



# Evaluación según la rúbrica (I)\*

## Nivel

- Vocabulario y/o gramática adecuados
- Vocabulario y/o gramática demasiado alto y no aceptable (ejemplos)
- Vocabulario y/ gramática demasiado alto pero aceptable (ejemplos)
- Vocabulario y/o gramática demasiado bajo (ejemplos)
- «Aceptable» significa que los cambios necesarios son mínimos o que el texto incluye palabras que, aunque no se entiendan, se pueden deducir por el contexto.

# Evaluación según la rúbrica (II)

## **Corrección lingüística**

- Correcto
- Incorrecto, pero fácilmente corregible (ejemplos)
- Incorrecto y no aceptable por el volumen de incorrecciones (ejemplos)

## **Pragmática**

- Aunque entiendas lo que dice el texto, ¿crees que hay elementos que un nativo no diría así? ¿Cuáles y cómo los dirías?
- Aunque entiendas lo que dice el texto, ¿crees que hay ideas que un nativo no incluiría en un texto así? ¿Cuáles?

# Evaluación según la rúbrica (III)

## **Contenido:**

- Diversidad y riqueza léxica: sí/no
- Contenido suficiente para evaluar la competencia (CE): sí/no

## **Evaluación global:**

- Aceptable sin cambios
- Aceptable con cambios asumibles
- No aceptable

Tiempo total invertido en la corrección del texto:

# Herramienta adicional: Cathoven

-∞ HSK1 HSK2 HSK3 HSK4 HSK5 HSK6 HSK7-9 +∞

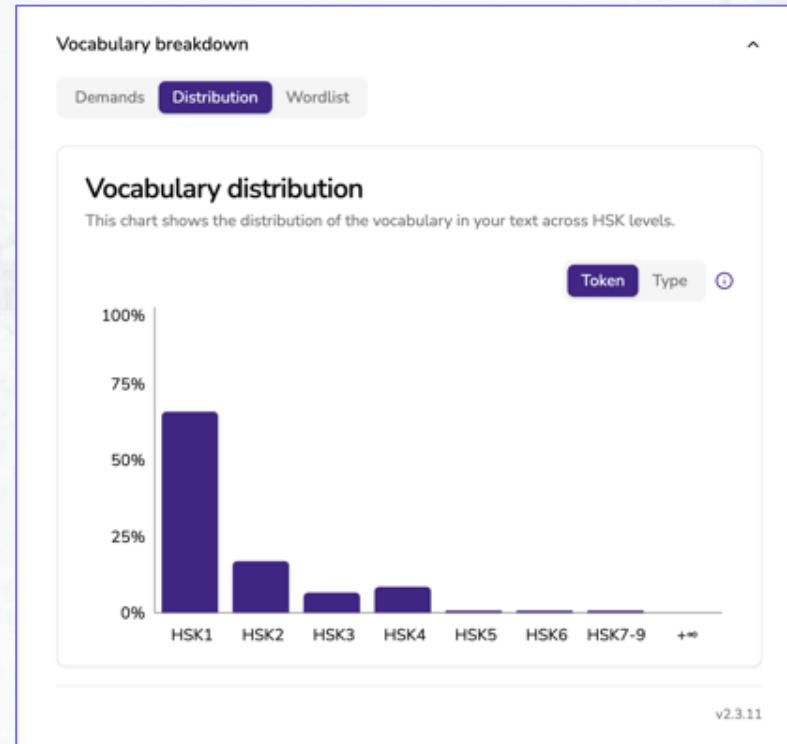
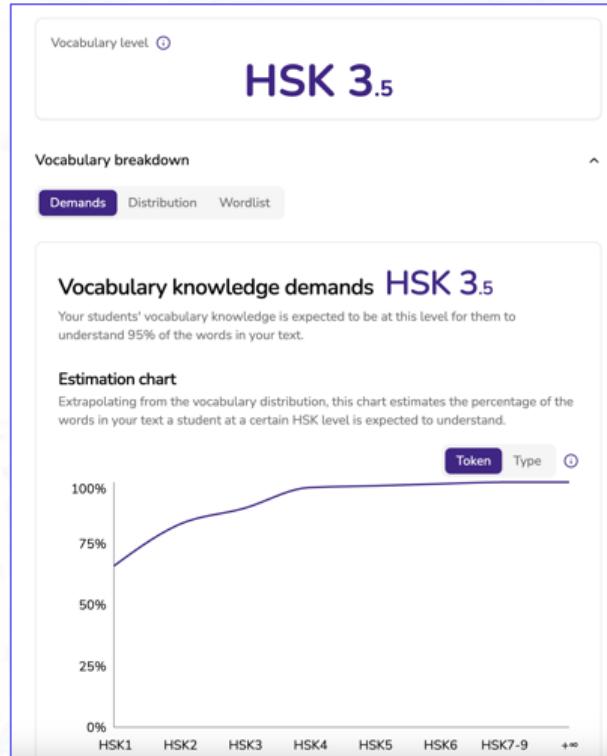
## 留学生的日常生活

王明是一名来自西班牙的留学生，他在北京的一所大学学习中文。他每天早上七点起床，洗脸、刷牙，然后去食堂吃早饭。他喜欢吃包子和喝豆浆。早上八点，他去教室上课。他的老师很友好，课堂内容很有趣，大家一起学习生词、语法和汉字。

中午十二点，王明和他的同学们一起去学校食堂吃午饭。他们常常吃米饭、炒菜和汤。吃完饭后，他有时候去图书馆看书，有时候在宿舍休息。下午两点到四点，他继续上课。老师会教他们如何用中文交流，大家一起练习口语。

晚上，王明喜欢去学校的体育馆打篮球，也会和朋友一起去附近的公园散步。他晚上十点睡觉，每天都很开心。他觉得在北京学习中文很有意思，因为他不仅学到了很多新知识，还交了很多中国朋友。

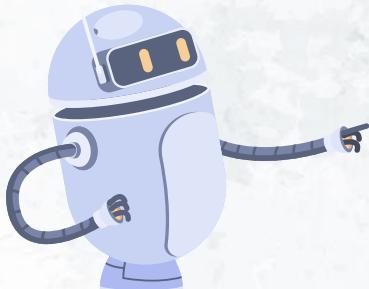
# Herramienta adicional: Cathoven



<https://www.cathoven.com/>

# Veamos ahora las preguntas...

Echa un vistazo a las preguntas generadas para el texto de ChatGPT.

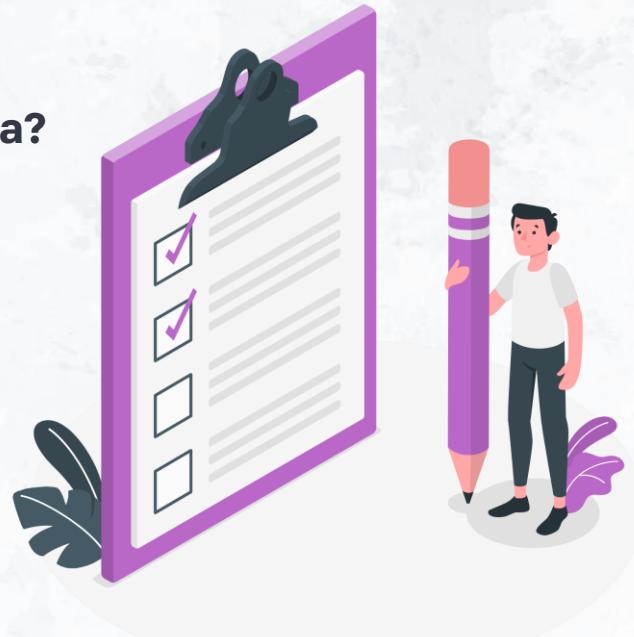


- ¿Qué te parecen?
- ¿Hay algún tipo de pregunta que te guste más que las otras?

# Rúbrica para las preguntas (I)

## 1. ¿Las preguntas se ajustan al nivel de lengua?

- Sí
- No, nivel demasiado alto
- No, nivel demasiado bajo



# Rúbrica para las preguntas (II)

## **2. ¿Todas las preguntas son adecuadas?**

- Sí
- No

En caso de que no, marca con colores las preguntas que no te parecen adecuadas.  
¿Por qué no lo son? (marca todas las opciones que consideres)

- La respuesta es demasiado explícita en el texto (respuesta muy obvia y/o formulación idéntica en el texto). [amarillo]
- Los distractores en las opciones de respuesta son inverosímiles (no salen en el texto). [naranja]
- No son pertinentes para el texto porque no se relacionan con su contenido. [rojo]

# Rúbrica para las preguntas (III)

## 3. Evaluación global

- Aceptable sin cambios
- Aceptable con cambios mínimos/asumibles
- No aceptable



04 →

# Resultados de nuestro estudio

(IA)

# Análisis de 8 actividades según las rúbricas anteriores

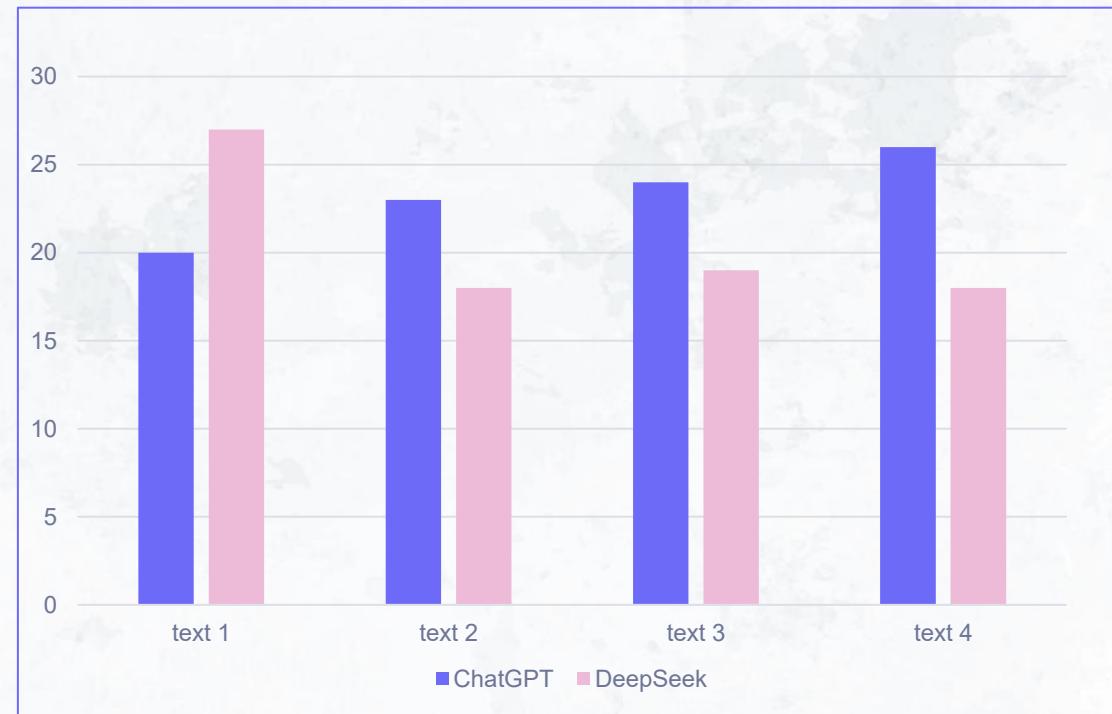
	Texto ChatGPT-4o Text	Preguntas ChatGPT-4o	Texto DeepSeek-R1	Preguntas DeepSeek-R1
<b>Actividad 1 - Rutinas</b>	Aceptable, cambios menores	Opción múltiple: no aceptable V/F: Aceptable, cambios menores	Aceptable sin cambios	Opción múltiple: no aceptable V/F: Aceptable, cambios menores
<b>Actividad 2 - Vacaciones</b>	Aceptable sin cambios	V/F: Aceptable, cambios menores	Versión 1: No aceptable Versión 2: Aceptable, cambios menores	(Según la versión 2) V/F: no aceptable
<b>Actividad 3 – Mi familia china</b>	Aceptable, cambios menores	V/F: Aceptable, cambios menores	Aceptable, cambios menores	V/F: no aceptable
<b>Actividad 4 - Piso</b>	Aceptable, cambios menores	Sin consenso	Aceptable, cambios menores	V/F: no aceptable

# Análisis de 8 actividades según Cathoven

	Texto	Texto
	ChatGPT-4o	DeepSeek-R1
Actividad 1	3,5	3,1
Actividad 2	3,4	(6,0) 4,00
Actividad 3	2,9	2,7
Actividad 4	4,2	4,6
Media	3,5	3,6

# Encuesta sobre la naturalidad

- Preguntamos a 30 personas nativas de chino.
- Escala de Likert (1-5).
- El gráfico muestra la agrupación de las respuestas de 4 y 5 (textos «muy naturales».)



05 →

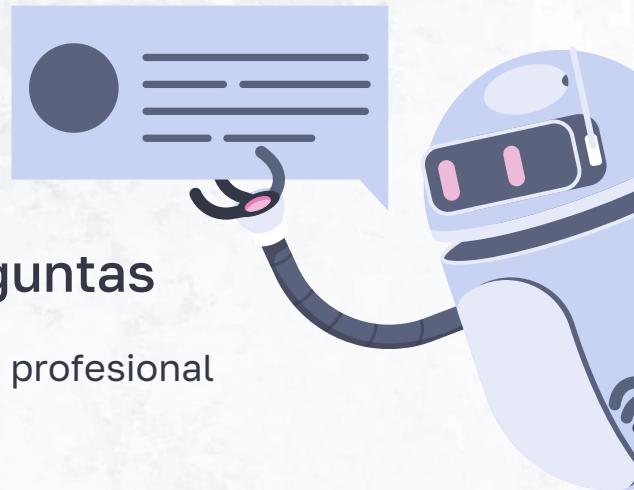
A modo de conclusión...

(IA)

# ¿En qué medida fue útil la IA?

## (+) Buenos resultados en la generación de textos

- Importancia del *prompt*
- Mejores resultados en nivel A2/HSK3 que con nivel A1
- Posiblemente, cuanto más nivel, mejores resultados



## (-) Menos eficiente en la generación de preguntas

- Las preguntas requieren enfoque pedagógico / experiencia profesional
- Tienen que servir para evaluar la comprensión lectora

# Otros aspectos

## (\*) ChatGPT vs. DeepSeek

Mejores resultados en el caso de ChatGPT (según las rúbricas y la encuesta)

## (\*) Evaluación de textos y preguntas

Rúbrica y Cathoven, herramientas complementarias y útiles para evaluación objetiva

## (\*) Futuras líneas de investigación

- Identificar los rasgos que definen la «naturalidad» en estos textos
- Probar otras opciones más avanzadas de las IA (GPT personalizado, temperatura, etc.)

# Referencias

- Casas-Tost, Helena, Isabel González-Torrents, Sara Rovira-Esteva & Mireia Vargas-Urpí.  
2025a. Materiales de análisis de actividades de comprensión escrita generadas con  
inteligencia artificial. *GELEA2LT Working Papers*. <https://ddd.uab.cat/record/313580>.
- Casas-Tost, Helena, Isabel González-Torrents, Sara Rovira-Esteva & Mireia Vargas-Urpí.  
2025b. Herramienta de evaluación de actividades de comprensión escrita generadas  
con inteligencia artificial. <https://ddd.uab.cat/record/313581>.
- Casas-Tost, Helena, Isabel González-Torrents, Sara Rovira-Esteva & Mireia Vargas-Urpí.  
En prensa. Using GAI for reading comprehension activities in the teaching and learning  
of Chinese as an additional language.

# Referencias

- Li, Belle, Victoria L. Lowell, Chaoran Wang & Xiangning Li. 2024. A Systematic Review of the First Year of Publications on ChatGPT and Language Education: Examining Research on ChatGPT's Use in Language Learning and Teaching. *Computers and Education: Artificial Intelligence* 7.  
<https://doi.org/10.1016/j.caeai.2024.100266>
- Rovira-Esteva, Sara. 2025. Teaching or Cheating: The Dark Side of ChatGPT as a Learning Companion for Beginner Chinese Students. *Journal of Technology and Chinese Language Teaching* 16(1). 52-84.  
<http://www.tclt.us/journal/abstractshow.php?id=162>.
- Stockwell, Glenn. 2024. ChatGPT in Language Teaching and Learning: Exploring the Road We're Travelling. *Technology in Language Teaching and Learning* 6 (1): 2273.  
<https://doi.org/10.29140/tltl.v6n1.2273>.
- Zhao, Qun (肇群), Hsu Yu-Yin (許又尹) & Huang, Chu-ren (黃居仁). 2024. Make LLMs Your Portable Teachers of Chinese as a Foreign Language -- Prompt Patterns of an AI-Teacher. In *Proceedings of the 12th International Conference and Workshops on Technology and Chinese Language Teaching*, 131–39. Los Angeles: TCT.

# Gracias :)

mireia.vargas@uab.cat

CRÉDITOS: Esta plantilla para presentaciones es una creación de **Slidesgo**, e incluye íconos de **Flaticon**, infografías e imágenes de **Freepik** y contenido de **Eliana Delacour**

Las ilustraciones de las diapositivas 22 y 24 son de **Storyset**.

