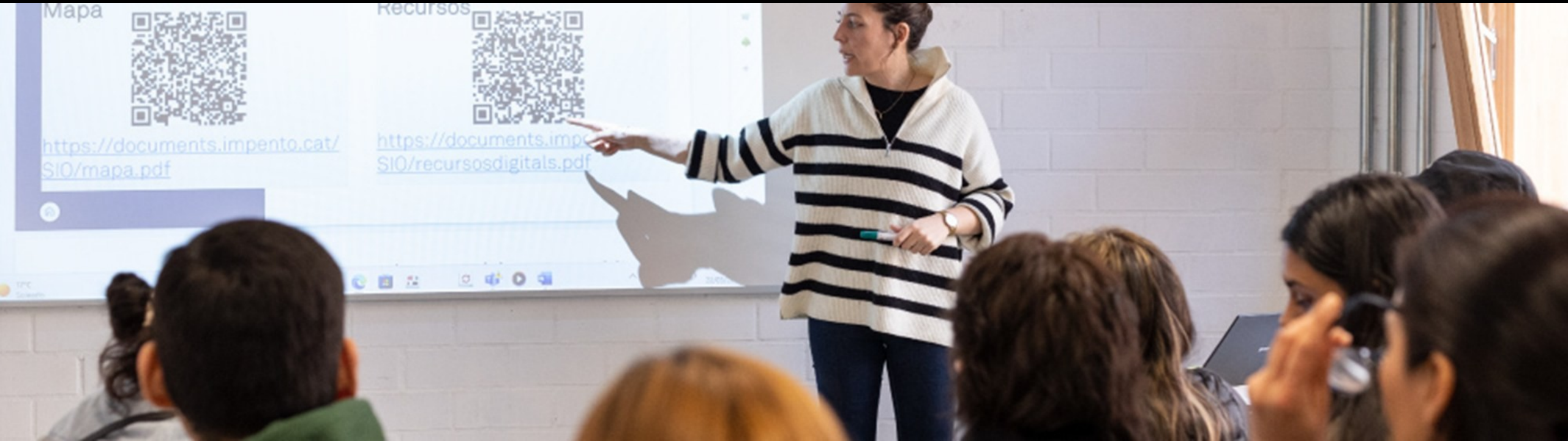


BARCELONA ACTIVA · CIBERNÀRIUM

Laboratori pràctic d'IA amb Hugging Face Spaces

Veure una capacitat, provar una demo, observar-ne els límits
i endur-se una conclusió útil.

3 hores · 25 mini-pràctiques · Daniel M. Casas



3 hores, 2 pauses i molt de laboratori

Alternem blocs curts de tendències amb pràctica real. Cap bloc teòric supera els 30 minuts.

00:00	● Benvinguda i mapa de la sessió	10'
00:10	● Tendències · Models frontera i agents	30'
00:40	● Laboratori 1 · Imatge i creació visual	25'
01:05	● PAUSA	7'
01:12	● Tendències · Visió, robòtica i mobilitat	25'
01:37	● Laboratori 2 · Visió per computador	30'
02:07	● PAUSA	7'
02:14	● Tendències · Regulació, feina i Catalunya	25'
02:39	● Laboratori 3 · Àudio, vídeo i flux multimodal	21'

**100%
pràctica
guiada**

Què és una Space i per què ens serveix

Un directori d'aplicacions d'IA amb demos d'equips oficials, empreses, recerca i comunitat.

SENYALS D'ESTAT QUE HAS DE SABER LLEGIR

Running

Activa i llesta per fer servir.

Sleeping

Dormida per inactivitat: es desperta al primer clic.

Paused

Aturada pel propietari. No arrencarà.

Runtime error

Fallada del codi o de dependències.

ZeroGPU / cua

GPU compartida: pot caldre esperar torn.

Regla d'or del laboratori

No busquem que tot surti perfecte.

Busquem aprendre a provar, comparar i decidir si una eina serveix per a un cas real.

Una Space caiguda també és contingut: ensenya que la fiabilitat forma part del criteri de selecció.

› [Documentació oficial de Spaces](#)

Abans de començar: què necessites i què no has de pujar

Requisits previs

- Navegador actualitzat.
- Compte gratuït de Hugging Face (recomanat, no obligatori).
- Auriculars per a les pràctiques d'àudio.
- Imatges de prova NO sensibles.
- Un document de prova inventat, sense dades reals.
- Una carpeta local per desar captures, resultats i notes.

Privacitat: què NO puges mai

- Contractes, factures o documents confidencials reals.
- DNI, passaport o dades bancàries.
- Històries clíniques o dades de salut.
- Imatges de menors i àudios privats.
- Contrasenyes i claus API.
- Qualsevol dada sensible d'empresa o de clients.

Moltes Spaces són demos públiques o comunitàries. Serveixen per aprendre, no per treballar amb informació real.

01

Imatge: crear, editar i netejar

- › Explorar el directori de Spaces i llegir-ne els estats
- › Text → imatge amb Stable Diffusion 3.5
- › Imatge → imatge, eliminació de fons i upscaling

25 MINUTS

Primer contacte i generació d'imatge

Obre cada Space en una pestanya nova. Si veus cua de ZeroGPU, passa a l'alternativa i torna-hi després.

1

Exploració inicial de Spaces

10'

Revisa categories, entra en una Space «Running» i localitza App / Files / Community. Entrega: 3 Spaces candidates amb una frase d'utilitat.

› huggingface.co/spaces

2

Registre i configuració bàsica

7'

Crea o revisa el compte per poder desar i participar. Comprova que no fas servir credencials compartides.

› huggingface.co/join

3

Generació d'imatge des de text

10'

Converteix una descripció en imatge. Entrega: imatge final + el prompt exacte que has fet servir.

› stabilityai/stable-diffusion-3.5-large

4

Edició image-to-image

12'

Transforma una imatge amb una instrucció. Entrega: abans/després i una nota del que ha canviat.

› diffusers/unofficial-SDXL-Turbo-i2i-t2i

Netejar i escalar: fons i resolució

5

Eliminació de fons

7'

Treu el fons d'una imatge. Si falla, fes servir Xenova, que s'executa al navegador i no depèn de GPU.

› [not-lain/background-removal](https://not-lain.com/background-removal)

6

Upscaling (millorar resolució)

10'

Puja la resolució d'una imatge petita i compara el resultat amb l'original al 100%.

› [bookbot/Image-Upscaling-Playground](https://bookbot.net/image-upscaling-playground)

PROMPT DE REFERÈNCIA · PRÀCTICA 3

«Il·lustració editorial d'una classe en línia sobre intel·ligència artificial a Barcelona, estil net i corporatiu,

llum suau, paleta magenta i grisos, sense text dins la imatge, format apaïsat.»

Canvia una sola variable cada vegada (estil, llum, paleta, enquadrament) per entendre què fa cada paraula.

02

Visió per computador

- › Llegir text d'una imatge (OCR) i descriure-la automàticament
- › Preguntar a una imatge i detectar-hi objectes
- › Segmentació, canvi d'angle i reconstrucció 3D aproximada

30 MINUTS

Del píxel al significat

És la mateixa família de tecnologia que hi ha darrere del control de qualitat industrial, la lectura de matrícules o la conducció autònoma.

7

OCR: extreure text d'una imatge

10'

Fes servir una imatge amb una frase inventada. Entrega: text extret i llista d'errors detectats.

› [nielsr/TrOCR-Scene-Text-Recognition](#)

8

Descripció automàtica d'imatge

7'

Genera una descripció (caption) i escriu-ne tu la versió humana. Compara-les.

› [hysts/image-captioning-with-blip](#)

9

Preguntes sobre una imatge (VQA)

10'

Fes dues preguntes sobre la mateixa imatge i valora si la resposta és correcta o inventada.

› [hysts/BLIP2](#)

10

Detecció d'objectes

10'

Localitza objectes amb caixes i etiquetes. Entrega: captura i observacions sobre falsos positius.

› [NeerajCodz/ObjectDetection](#)

Segmentar, girar la càmera i sortir del pla

24 Segmentació d'imatge

10'

Separa regions o objectes amb màscares. Corre al navegador amb WebGPU: no depèn de cua de GPU.

› [Xenova/segment-anything-webgpu](#)

25 Canviar l'angle d'una foto

12'

Novel view synthesis: el model imagina l'escena des d'un altre angle. Anota què conserva i què s'inventa.

› [multimodalart/qwen-image-multiple-angles-3d-camera](#)

23 Imatge → model 3D

15'

Converteix una foto d'un objecte senzill en un 3D aproximat. Entrega: captura i nota de fidelitat.

› [tencent/Hunyuan3D-2](#)

Compte amb la paraula «reconstrucció»

El resultat és plausible, no fidel. És una imatge generada, no una mesura de la realitat. Útil per a prototipatge i comunicació; arriscat per a peritatges, documentació tècnica o proves.

03

Àudio, text i documents

- › Transcripció automàtica i veu sintètica
- › Traducció, resum i anàlisi de sentiment
- › Preguntar a un document sense comprometre dades reals

25 MINUTS

La veu i la llengua com a interfície

11 Àudio a text (Whisper)

12'

Transcriu un àudio breu i corregeix-lo. Compta quants errors fa amb accents i noms propis.

› [openai/whisper](#)

12 Text a veu

10'

Crea una locució breu. Entrega: àudio generat i el text exacte que has fet servir.

› [k2-fsa/text-to-speech](#)

13 Traducció automàtica

7'

Tradueix un text i revisa'n els matisos. Busca una frase on la màquina perdi la ironia o el registre.

› [UNESCO/nllb](#)

14 Resum de text

10'

Sintetitza un text llarg i corregeix el resum. Què ha desaparegut que era important?

› [razakhan/text-summarizer](#)

Classificar opinions i preguntar a un document

15 Anàlisi de sentiment

7'

Classifica opinions en positiu/negatiu. Entrega: taula text / etiqueta / comentari teu.

› [nazianafis/Sentiment-Analysis](#)

16 Preguntes a PDFs

12'

Puja el document INVENTAT de prova i fes-li 3 preguntes. Entrega: preguntes i respostes.

› [cvachet/pdf-chatbot](#)

ATURA'T ABANS DE PUJAR RES

La pràctica 16 és la més perillosa de la sessió: convida a pujar un PDF. Fes servir només el document de prova inventat.

Abans de pujar qualsevol fitxer a una demo pública, revisa'n tres coses: la política de privacitat, la llicència i on s'executa el model. Si no ho pots comprovar en un minut, no ho pugis.

04

Vídeo, música i flux multimodal

- › Generació de vídeo i de música: on és el límit real avui
- › Comparar dues Spaces amb criteri professional
- › Encadenar diverses IA en un mateix flux de treball

21 MINUTS

Generar moviment i so — i saber comparar

21

Generació de vídeo

15'

Prova text-to-video o image-to-video. Entrega: clip o captura i el temps d'espera real.

› [Lightricks/ltx-video-distilled](https://lightricks.com/ltx-video-distilled)

22

Generació musical

12'

Crea un fragment musical breu. Entrega: àudio i prompt. Alternativa al navegador: MusicGen Web.

› [facebook/MusicGen](https://facebook.com/MusicGen)

19

Comparar dues Spaces que fan el mateix

15'

La pràctica que més s'assembla a una decisió real de feina. Omple una taula amb quatre columnes: qualitat del resultat, velocitat, facilitat d'ús i fiabilitat (s'ha caigut? hi havia cua?). Després decideix quina recomanaries a un client i per què.

› [not-lain/background-removal](https://not-lain.com/background-removal) vs › [Xenova/remove-background-web](https://Xenova.com/remove-background-web)



Mini projecte final: una fitxa feta amb quatre IA

Encadena el que has après. En 15 minuts has de tenir una peça de comunicació completa, generada i verificada per tu.

17

Imatge per a xarxes

Genera la base visual del post amb Stable Diffusion 3.5 i escriu-ne el copy.

› stabilityai/stable-diffusion-3.5-large

18

Locució d'introducció

Crea l'àudio de benvinguda amb Edge-TTS. Entrega: àudio final i guió.

› [inoai/Edge-TTS-Text-to-Speech](https://ino.ai/Edge-TTS-Text-to-Speech)

20

Flux multimodal

Uneix imatge, àudio i text en una sola fitxa coherent i revisa-la.

› huggingface.co/spaces

Entregable de la sessió: una fitxa amb imatge, àudio i resum + tres línies sobre què has hagut de corregir a mà.

Tres tallers més que tens al Trello del curs

Si vas ràpid o vols continuar a casa, aquestes pràctiques amplien la sessió cap a assistents i recerca.

1

Custom GPT «Tendències IA BCN»

Munta un assistent per a principiants amb instruccions i knowledge ja escrits, en català i amb exemples de Barcelona. Inclou la versió equivalent per a Gemini.

[› Obrir al Trello](#)

2

Investigació profunda (Deep Research)

Dissenya el teu prompt de recerca i executa'l a Gemini, ChatGPT o DeepSeek. Inclou 12 temes preparats: IA i ocupació, geopolítica, ciutats intel·ligents, democràcia i desinformació...

[› Obrir al Trello](#)

3

NotebookLM sobre la teva recerca

Converteix l'informe en targetes d'estudi, mapa conceptual i resum executiu d'una pàgina. L'eina organitza; el criteri final és teu.

[› NotebookLM](#)

Ruta recomanada a classe · i la resta, com a pràctica autònoma

Les 25 pràctiques sumen molt més de 3 hores. Aquestes 9 són el nucli que farem junts; les altres 16 queden documentades al Trello per continuar pel teu compte.

1 · Explorar Spaces

3 · Text → imatge

5 · Eliminar fons

7 · OCR

9 · Preguntes a una imatge

11 · Àudio a text

12 · Text a veu

25 · Canviar l'angle

20 · Flux multimodal

Autònoma (documentada al Trello): 2, 4, 6, 8, 10, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 24.

› [Llista completa del laboratori al Trello del curs](#)

Criteri: com decidir si una eina et serveix

Provar és fàcil. El valor professional és saber respondre aquestes quatre preguntes davant d'un client o d'un cap.

1

Funciona prou bé?

Compara sempre amb el que fas ara a mà. Si has de corregir més del que estalvies, no serveix.

2

És fiable?

Una demo que es cau, dorm o té cua no és una eina de producció. La disponibilitat forma part de la qualitat.

3

És segura?

On s'executa? Qui veu les dades? Quina llicència té la sortida? Si no ho saps, no hi posis res real.

4

Qui respon si falla?

Amb una demo comunitària, ningú. Per a un ús professional cal un proveïdor amb condicions per escrit.

Cap eina és bona o dolenta en abstracte: ho és per a una tasca, un volum i un nivell de risc concrets.

Per on seguir després d'avui

CIBERNÀRIUM

Enginyeria de prompts

3 sessions de 3 hores. El pas natural després d'aquesta sessió: escriure instruccions que funcionin de manera consistent.

RUTA SEGÜENT

Vibe coding: programar amb IA

Crear aplicacions i webs reals amb assistència d'IA, partint del que has vist aquí.

AUTÒNOM

IA local amb LM Studio i Ollama

Executar models al teu propi ordinador: més privacitat i control, sense enviar dades al núvol.

Tot el material de la sessió, targeta a targeta:

› Laboratori pràctic d'IA amb Hugging Face Spaces (Trello)

Gràcies!

Daniel M. Casas

Full Stack Developer · Màster en Intel·ligència Artificial i Big Data

Formador tecnòleg a Cibernàrium — Barcelona Activa

Continuïtat de la ruta formativa: cibernarium.barcelonactiva.cat

