AI Congress Barcelona

20 - 21 - 22 OCT 2026

Auditori L'illa

Un esdeveniment de:

CIDAI Centre of Innovation for Data tech and Artificial Intelligence

Organitzat per:

eurecat

Alineat amb l'Estratègia Catalunya IA 2030:

Generalitat de Catalunya

Segueix-nos a:

#AICongressBCN

aicongress.barcelona

Patrocinadors OR:

CaixaBank TECH

conversia compliance is our business

forvis mazars & infini DIGITAL EFFICIENCY

orange Empresas

raona coding corporate futures

softeng Maximize the cloud. Securely.

Patrocinadors PLATA:

apdcat Autoritat Catalana de Protecció de Dades

LABORAL kutxa

Col·laboradors:

Barcelona

BSC Barcelona Supercomputing Center Centro Nacional de Supercomputación

CVC Centre de Visió per Computador

i2cat

UPC Universitat Politècnica de Catalunya Barcelona Tech

IDEAI Intelligent Data Science and Artificial Intelligence Research Center

Congress Barcelona

Un esdeveniment de: CIDAI Centre of Innovation for Data tech and Artificial Intelligence

Organitzat per: eurecat

Alineat amb l'Estratègia Catalunya IA 2030: Generalitat de Catalunya

20 - 21 - 22 OCT 2026
#AICongressBCN
aicongress.barcelona

MÉS ENLLÀ DELS MODELS DE LLENGUATGE: ELS MODELS MULTIMODALS

Patrocinadors OR: CaixaBank TECH · conversia compliance is our business · forvis mazars & infini DIGITAL EFFICIENCY · orange Empresas · raona coding corporate futures · softeng Maximize the cloud. Securely.

Patrocinadors PLATA: apdcat Autoritat Catalana de Protecció de Dades · LABORAL kutxa

MÉS ENLLÀ DELS MODELS DE LLENGUATGE: ELS MODELS MULTIMODALS

Presenta:



Rafael Redondo
Director de la Unitat de Tecnologies Multimèdia
Eurecat

Ponents:



Carles Rams
CEO
Ebantic



Nicolás Carrasco
CTO
Unnic AI

Congress Barcelona

Un esdeveniment de: CIDAI Centre of Innovation for Data tech and Artificial Intelligence

Organitzat per: eurecat

Alineat amb l'Estratègia Catalunya IA 2030: Generalitat de Catalunya

20 - 21 - 22 OCT 2026
#AICongressBCN
aicongress.barcelona

MÉS ENLLÀ DELS MODELS DE LLENGUATGE: ELS MODELS MULTIMODALS

Patrocinadors OR: CaixaBank TECH · conversia compliance is our business · forvis mazars & infini DIGITAL EFFICIENCY · orange Empresas · raona coding corporate futures · softeng Maximize the cloud. Securely.

Patrocinadors PLATA: apdcat Autoritat Catalana de Protecció de Dades · LABORAL kutxa



Carles Rams
CEO
Ebantic

PONÈNCIA:

ACN-IA: UNA WEB HÍBRIDA PER NAVEGAR I DIALOGAR AMB UNA HEMEROTECA MULTIMODAL

ebantic.
Management solutions for digital media

AI CONGRESS BARCELONA 2026 · CASO DE USO

Del archivo audiovisual a la conversación

UN PROYECTO DE

acn watchity AMPLIFY NUBILUM aureacolor ebantic. Management solutions for digital media

¿Cómo transformar un archivo de más de 250 mil vídeos, 900 mil imágenes, 750 mil audios y más de un millón de textos informativos en un **nuevo modelo de negocio** para la Agència Catalana de Notícies?

EL PUNTO DE PARTIDA

Un gran archivo que solo se encontraba por el texto

noticias

+1M

fotografías

Ą E Ć Ć Ħ

audios

+750K

vídeos

+250K

Un fondo que crece desde 1999

Indexado solo por texto

La búsqueda dependía del texto de la noticia y de los pies de foto.

Buscador estructurado

Había que saber exactamente qué buscar, y cómo estaba catalogado.

Lo audiovisual, sin indexar

Lo que se ve y se dice en fotos, audios y vídeos era invisible al buscador.

Nueva necesidad de negocio

La Agència quería vender su contenido con un modelo de cesta de la compra.

EL OBJETIVO

Preguntar al archivo en lenguaje natural

Navegación estructurada y diálogo conversacional con IA conversacional con IA sobre el fondo de la Agència Agència Catalana de Notícies

“Vídeos de [persona] hablando de [tema]”

¿Quién? ¿Qué? ¿Cuándo? ¿Dónde?

Web híbrida

Buscador estructurado y chatbot conviven en la misma interfaz. El usuario elige cómo buscar.

Multimodal

Texto, audio, vídeo e imagen en un solo resultado: la respuesta mezcla clips, fotos y audios.

Integrado en la compra

El chat no se queda en la consulta: lleva al contenido a la cesta y cierra la venta.

LA RESPUESTA: SEIS ORGANIZACIONES, UN MISMO ARCHIVO

¿Y esto cómo lo hacemos?

acn

Cliente y dueña del archivo

Plantea el reto, aporta el fondo y valida la solución en entorno real.

watchity

Infraestructura del archivo

Almacenamiento cloud (S3) del fondo audiovisual de la Agència.

AMPLIFY

IA aplicada al archivo

18 modelos de IA que generan los metadatos de fotos, audios y vídeos.

NUBILUM

Curación de metadatos

Matching manual de caras y voces: entrena y afina el reconocimiento.

aureacolor

Diseño UX/UI del chat

Diseño gráfico y funcional de la experiencia conversacional.

ebantic.
Management solutions for digital media

Integración y gestión

CMS, portal y APIs. Integra el chat en la plataforma y coordina el proyecto.

LA ARQUITECTURA

Cómo encajan las piezas

1

Archivo en la nube

Fotos, audios y vídeos en S3

watchity

2

Indexación con IA

18 modelos + texto de la noticia

AMPLIFY

3

CMS y APIs

Metadatos, portal y chat LLM vía API

ebantic.
Management solutions for digital media

4

Portal con chat IA

Experiencia diseñada para el usuario

aureacolor

NUBILUM

Curación humana
Matching de caras y voces

acn

Usuarios y clientes de la ACN
Preguntan y compran

Entender el archivo como lo haría una persona

¿Quién?

Personas que aparecen y hablan

- Reconocimiento facial
- Reconocimiento de locutor
- Diarización
- NER: personas

¿Qué?

Lo que se ve, se dice y se cuenta

- Descripción con VLM
- Detección de objetos
- Logos y marcas
- Transcripción
- Keywords del habla

¿Cuándo?

Fechas, eventos y el momento exacto

- NER: fechas y eventos
- OCR de rótulos
- Transcripción con tiempos

¿Dónde?

Lugares citados y escenarios

- NER: lugares
- OCR de carteles
- Descripción de escena

¿Cómo?

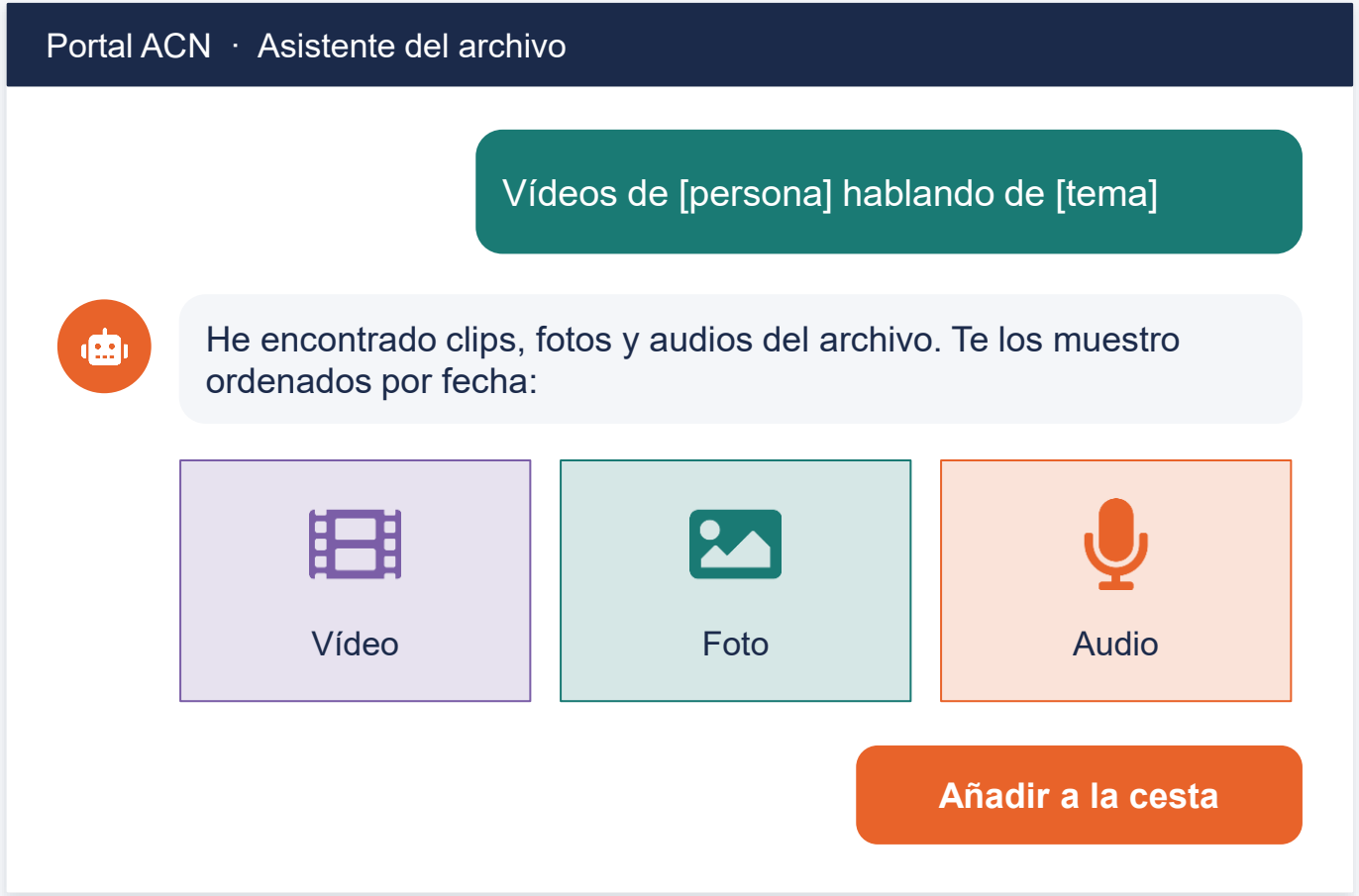
Formato, contexto y lo que ocurre

- Tipología de contenido
- Resumen automático
- Tipos de plano
- Segmentación semántica

18 modelos de IA + texto de la noticia → **Índice + embeddings** → **Chat LLM en el portal**

LA EXPERIENCIA

Buscar conversando… y llegar a la cesta



Lenguaje natural

El cliente pregunta como hablaría con un documentalista.

Una sola respuesta multimodal

Vídeo, foto y audio juntos, con el momento exacto.

Del hallazgo a la compra

El contenido encontrado va directo a la cesta.

LO QUE NOS LLEVAMOS

Cada uno, en lo que mejor sabe hacer

Colaborar

Ninguna de las seis organizaciones lo habría resuelto sola. El valor está en cómo encajan las piezas.

Humanos en el bucle

Los documentalistas validan caras y voces. La IA propone, las personas confirman.

El contexto como catalizador

El texto de las noticias, el que usan los periodistas, da sentido a lo que la IA ve y oye.

Validar en entorno real

Requisitos de usabilidad definidos y probados con la propia Agència, no en el laboratorio.

ebantic.
Management solutions for digital media

Gracias

Un modelo replicable en otras agencias, televisiones y archivos públicos

crams@ebantic.com

UN PROYECTO DE

acn watchity AMPLIFY NUBILUM aureacolor ebantic. Management solutions for digital media

Congress Barcelona

Un esdeveniment de: CIDAI Centre of Innovation for Data tech and Artificial Intelligence

Organitzat per: eurecat

Alineat amb l'Estratègia Catalunya IA 2030: Generalitat de Catalunya

20 - 21 - 22 OCT 2026
#AICongressBCN
aicongress.barcelona

MÉS ENLLÀ DELS MODELS DE LLENGUATGE: ELS MODELS MULTIMODALS

Patrocinadors OR: CaixaBank TECH · conversia compliance is our business · forvis mazars & infini DIGITAL EFFICIENCY · orange Empresas · raona coding corporate futures · softeng Maximize the cloud. Securely.

Patrocinadors PLATA: apdcat Autoritat Catalana de Protecció de Dades · LABORAL kutxa



Nicolás Carrasco
CTO
Unnic AI

PONÈNCIA:

MÉS ENLLÀ DEL CHATBOT: IA APLICADA A LA PRESSUPOSTACIÓ INDUSTRIAL

AI CONGRESS 2026 · MODELOS MULTIMODALES

Modelos que deciden como humanos

Spoiler: ninguno decide como un humano. Lo que decide es **cómo se reparte la decisión**.

Nicolás Carrasco · CTO, Unnic AI

Percibir no es **decidir**.

Y lo que nos pedían era decidir.

01 · EL ENCARGO

Cuánto cuesta algo que nunca has fabricado

A un fabricante de transformados metálicos le llega una petición de oferta con un modelo 3D, un PDF de plano por pieza y un correo. Un técnico decide **qué operaciones lleva cada pieza y cuánto cuesta**.

Tarda entre una y dos semanas. Y dos técnicos con la misma documentación no dan la misma cifra.

95 %

de las ofertas comerciales respetan ese escandallo técnico.

Esa decisión **es** la oferta.

02 · LAS MODALIDADES

Tres señales distintas, y una lo es ella sola

Geometría 3D

El modelo del conjunto

— Medidas y estructura exactas

— Jerarquía de piezas

— Ni material ni acabado

Plano técnico

Texto + layout + visión

— Vector, símbolo y texto a la vez

— Un documento multimodal solo

— Uno por pieza: hay que casarlo

Correo

Lenguaje natural

— Lotes, plazos y acabados

— Ambiguo por naturaleza

— Lo decisivo, a veces al final

Aquí no hay ni audio ni vídeo: en la industria documental las modalidades son otras.

03 · DÓNDE EL MODELO MULTIMODAL SÍ GANA

Un plano solo lo resuelve un modelo que lo mire entero

- **Solo OCR —** sacas el texto y pierdes lo que significa
- **Solo extracción vectorial —** sacas la geometría y pierdes las notas
- **Por separado y luego juntado —** pierdes la relación: qué símbolo va con qué arista

Integrar e interpretar simultáneamente es literalmente la definición de esta sesión. Y aquí se cumple.

Document AI · modelos que fusionan texto, layout e imagen en una sola representación

Y aun leyéndolo perfecto,
todavía no tienes un precio.

Tienes datos.

Percibir es el primer paso, no el único

01
PERCIBIR
Leer modelo, planos y correo

02
IDENTIFICAR
Separar piezas y conjuntos

03
INFERIR
Completar lo que no dice ningún fichero

04
DECIDIR
Qué operaciones lleva cada pieza

05
CALCULAR
Tiempos, material y escalado por lote

06
DUDAR
Qué asumo y qué tengo que consultar

Ninguna modalidad lo contiene

Saber cuándo no sabes

Dos pasos que ninguna modalidad cubre

● la contiene ◑ en parte — no

	Geometría 3D modelo del conjunto	**Plano técnico** texto + layout + visión	**Correo** lenguaje natural
01 · Percibir	●	●	●
02 · Identificar	●	◑	—
03 · Inferir lo ausente	—	—	—
04 · Decidir operaciones	◑	◑	◑
05 · Calcular	●	◑	●
06 · Dudar	—	—	—

EL HUECO

No es falta de potencia del modelo: la información no existe en ningún canal legible.

06 · LA TRAMPA

Por qué no basta con un solo modelo grande

01	**Siempre responde**	Siempre da una cifra y siempre suena razonable. No la auditas ni se la explicas al cliente.
02	**La similitud engaña**	El espacio de representación no está alineado con el coste: dos piezas casi idénticas difieren en un orden de magnitud.
03	**Predice la media**	Entrenado con tu histórico, se rompe en la pieza nueva — que es donde está el margen.
04	**No hay datos para afinarlo**	Harían falta miles de escandallos etiquetados. Hay cientos, heterogéneos, y el valor está justo en lo no visto.

07 · LA ARQUITECTURA

Cada señal, a la técnica que sabe leerla



Fusionar no es concatenar: es casar entidades



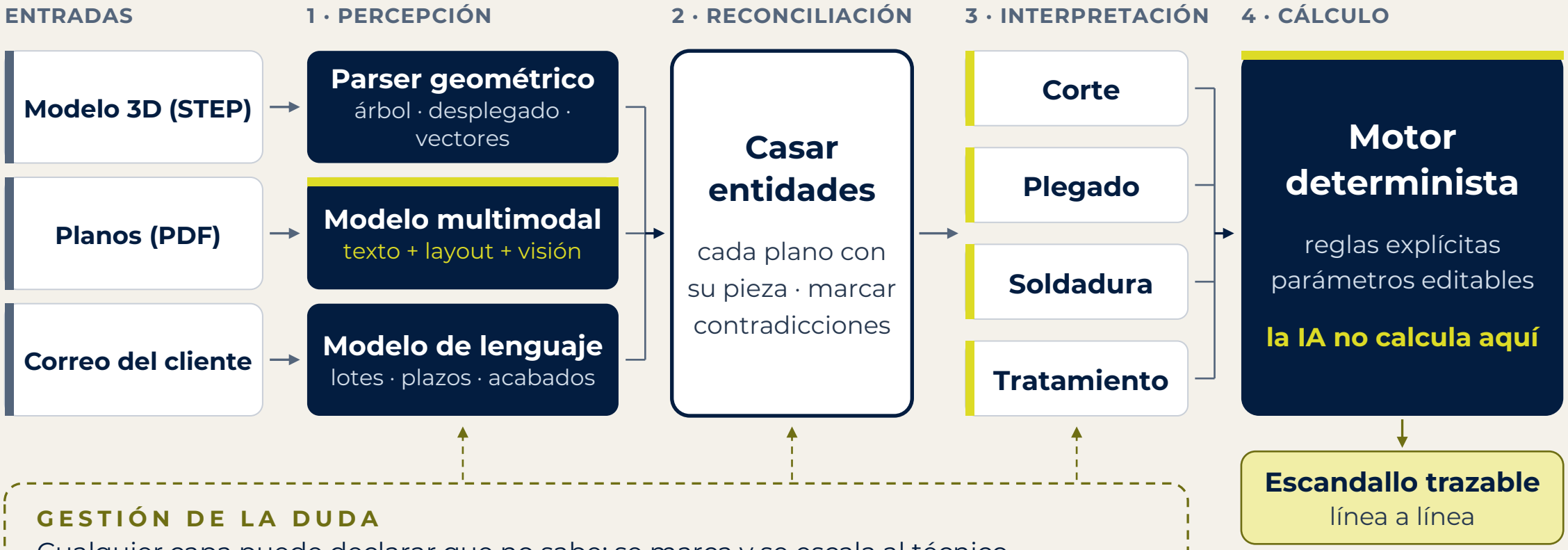
07 · LA ARQUITECTURA

La IA interpreta. El motor calcula



Arquitectura neurosimbólica: percepción aprendida, cálculo simbólico y trazable

Y cualquier capa puede declarar que no sabe



Predicción selectiva: el sistema decide cuándo abstenerse en lugar de responder

08 · LAS DECISIONES DE DISEÑO

Tres decisiones sostienen todo lo anterior

01

La IA no calcula

En industria nadie firma un presupuesto que no puede defender. La trazabilidad no es cumplimiento: es la condición para que alguien lo use.

02

Declara lo que no sabe

El valor no está en acertar el cien por cien. Está en saber en qué porcentaje no puedes confiar.

03

Se edita, no se reentrena

Sube el acero y alguien cambia un número en un panel. Si eso exige reentrenar, el sistema está muerto a los seis meses.

09 · DÓNDE SE ROMPE

Los tres límites reales, sin maquillar

La soldadura

Lo que más pesa en el coste no está en el modelo 3D y a menudo tampoco en el plano.

El 15 % de desplegados

Falla porque el cliente modeló la pieza como sólido. El input no trae lo que el proceso necesita.

La variabilidad de formato

Cada cliente dibuja a su manera. Lo que funciona con uno se degrada con el siguiente.

El límite no está en el modelo. Está en la información que existe.

unnic.
AI CONSULTING

Un modelo multimodal ve lo mismo que vuestro mejor técnico. Lo que no tiene son sus veinte años decidiendo qué mirar.

Eso no se entrena: **se arquitectura**.

Nicolás Carrasco · CTO, Unnic AI

nicolas@unnic.ai

+34 610 75 76 89

unnic.ai

Conectamos en LinkedIn

Escanea el código

AI Congress Barcelona

20 - 21 - 22 OCT 2026

Auditori L'illa

Un esdeveniment de:

CIDAI Centre of Innovation for Data tech and Artificial Intelligence

Organitzat per:

eurecat

Alineat amb l'Estratègia Catalunya IA 2030:

Generalitat de Catalunya

Segueix-nos a:

#AICongressBCN

aicongress.barcelona

Patrocinadors OR:

CaixaBank TECH

conversia compliance is our business

forvis mazars & infini DIGITAL EFFICIENCY

orange Empresas

raona coding corporate futures

softeng Maximize the cloud. Securely.

Patrocinadors PLATA:

apdcat Autoritat Catalana de Protecció de Dades

LABORAL kutxa

Col·laboradors:

Barcelona

BSC Barcelona Supercomputing Center Centro Nacional de Supercomputación

CVC Centre de Visió per Computador

i2cat

UPC Universitat Politècnica de Catalunya Barcelona Tech

IDEAI Intelligent Data Science and Artificial Intelligence Research Center