

Ingurumen Neurketa Tailerra: Ingurumena Babestu eta Ebaluatzeko AA

Taller de Medición Ambiental: IA para la Protección y Evaluación del Medio Ambiente

Prof. Dr. Ibai Laña

Prof. Dr. Javier Del Ser

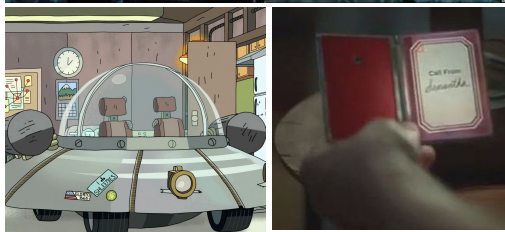
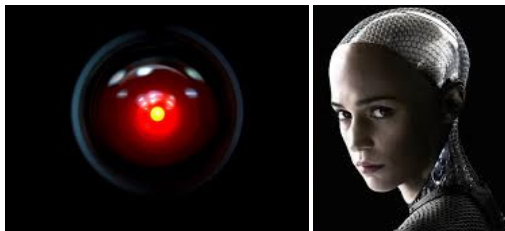
9/9/2026

Agenda

- Medioambiente, Datos, IA
- *Hands on* IAGEN
- Enfoque *clásico*
- Conclusiones

Medioambiente, Datos, IA

¿Qué es IA?



allá por 2022...

pero ahora...

mira, te doy una serie de ideas para un curso que estoy dando y necesito contenidos para las slides. Es un curso de verano de IA y medio ambiente y necesito contenidos para una slide en primer lugar que defina lo que es ia. puede estar bien usar algun recurso gráfico también

He preparado la presentación completa para tu curso de verano sobre **Inteligencia Artificial y Medio Ambiente**, estructurada en un formato visual y profesional de alta calidad.

Incluye una diapositiva principal dedicada a **definir qué es la IA** utilizando un recurso gráfico explicativo y conceptual, así como un recorrido completo por el programa del curso (aplicaciones climáticas, biodiversidad, redes inteligentes, huella de carbono de los centros de datos y balance sostenible).

 Curso de Verano: IA y Medio Ambiente

3 sept, 16:51

[Abrir](#)

¡Tu presentación de diapositivas sobre **IA y Medio Ambiente** está lista! Puedes revisarla en la ventana interactiva y comentarme si deseas ajustar algún texto, añadir algún tema específico del temario o modificar el estilo visual.

   ...

veamos...



 CURSO DE VERANO 2026

IA y Medio Ambiente

Oportunidades, Desafíos Climáticos y Sostenibilidad Tecnológica para el Futuro del
Planeta

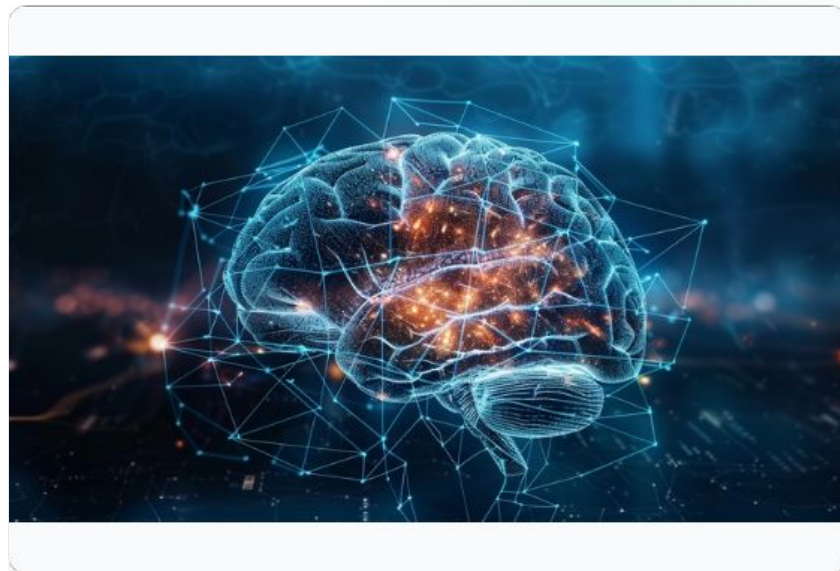


¿Qué es la IA?



La **Inteligencia Artificial (IA)** es la rama de las ciencias de la computación que desarrolla sistemas capaces de realizar tareas que simulan la cognición humana.

-  **Aprendizaje Autónomo:** Identifica patrones complejos en grandes volúmenes de datos sin programación explícita.
-  **Predicción y Modelado:** Anticipa comportamientos y variables futuras basadas en tendencias históricas.
-  **Toma de Decisiones:** Optimiza recursos, procesos y sistemas en tiempo real de forma automatizada.



Fundamentos Técnicos



Machine Learning

Algoritmos estadísticos que permiten a las computadoras aprender de los datos climáticos y mejorar su precisión con la experiencia.



Deep Learning

Redes neuronales profundas para el procesamiento de imágenes satelitales, visión por computador y análisis de audio ambiental.



Big Data Ambiental

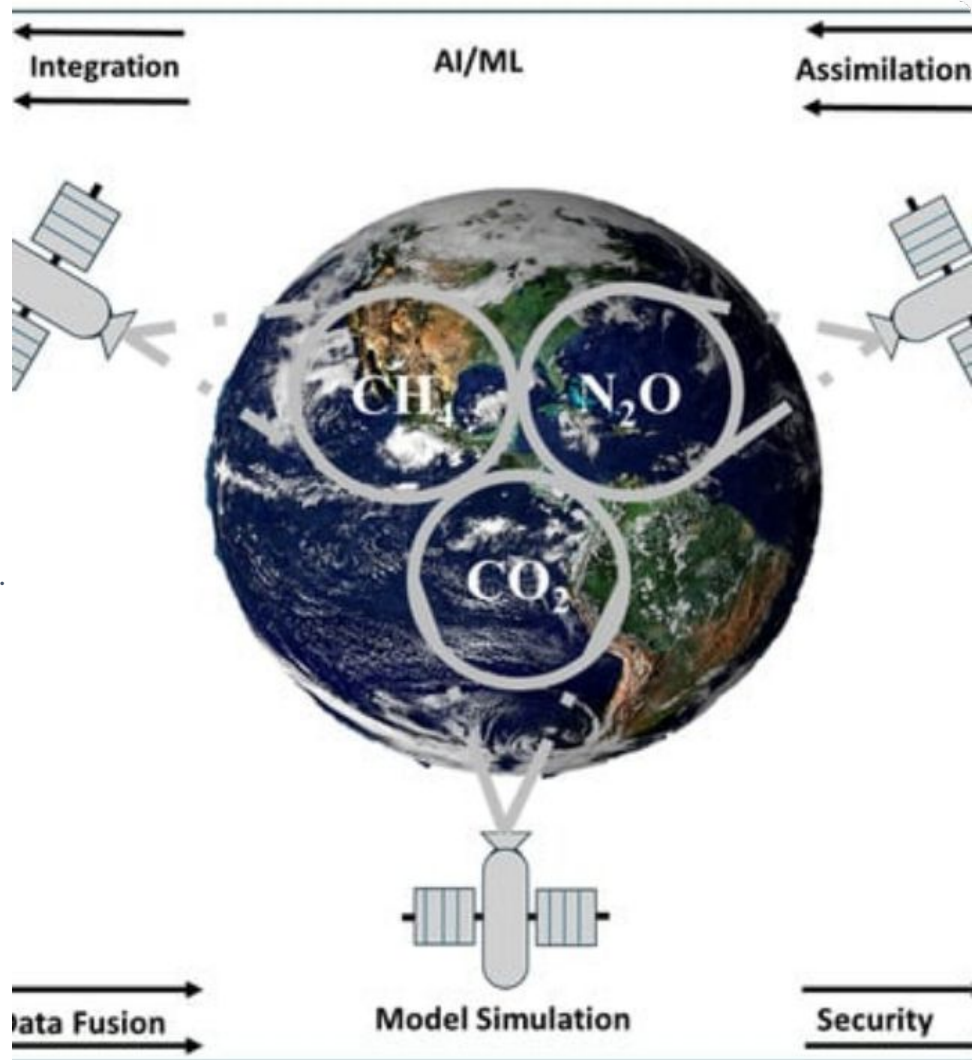
Procesamiento masivo de datos sensoriales en tiempo real procedentes de estaciones meteorológicas, satélites y sensores marinos.



Modelado Climático

Predicción y Simulación

La IA permite procesar petabytes de datos satelitales para simular modelos de calentamiento global con una precisión sin precedentes. Facilita la detección temprana de eventos meteorológicos extremos (huracanes, sequías e inundaciones), permitiendo diseñar mejores planes de adaptación local.



Áreas de Aplicación

Ciudades Inteligentes

Sincronización de semáforos en tiempo real mediante IA para reducir congestiones de tráfico, disminuyendo hasta un 10% las emisiones urbanas de CO2.

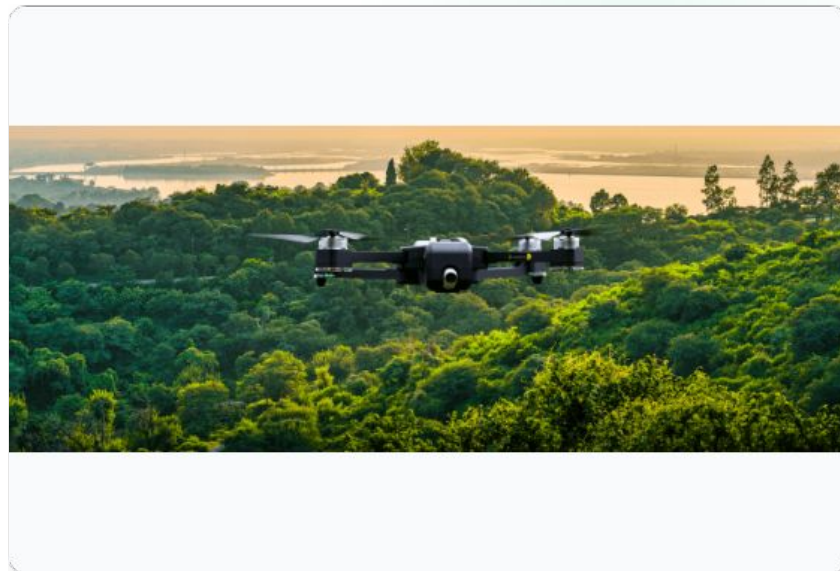
Agricultura de Precisión

Optimización del uso de agua y fertilizantes mediante sensores de suelo y visión artificial, reduciendo el desperdicio agrícola y protegiendo acuíferos.



Biodiversidad y Naturaleza

-  **Fototrampeo Automatizado:** Identificación instantánea de especies en miles de imágenes nocturnas usando visión por computador.
-  **Monitoreo Acústico:** Detección de motosierras o disparos en selvas protegidas mediante el análisis de patrones de audio en vivo.
-  **Control de Deforestación:** Análisis automático de imágenes de dron para alertar sobre talas ilegales en cuestión de minutos.



Redes Eléctricas Inteligentes



Integración Energética

Las fuentes de energía renovable (solar y eólica) son intermitentes por naturaleza. La IA predice la producción energética según el pronóstico del tiempo.

Ajusta dinámicamente la oferta y la demanda en las redes eléctricas (Smart Grids), reduciendo la necesidad de recurrir a plantas térmicas contaminantes.



Antes de 2022 había IA

y big data, y machine learning, y data scientists...

hay algo común a todo ello, y a las IAs modernas: los datos

pero

Hasta hace poco...

DATOS

INTEGRACIÓN

Desde diferentes
fuentes de datos

Recolectando datos
en formato común

LIMPIEZA

Datos perdidos

Eliminar ruido

Eliminar outliers

Datos duplicados

Eliminar inconsistencias

TRANSFORMACIÓN

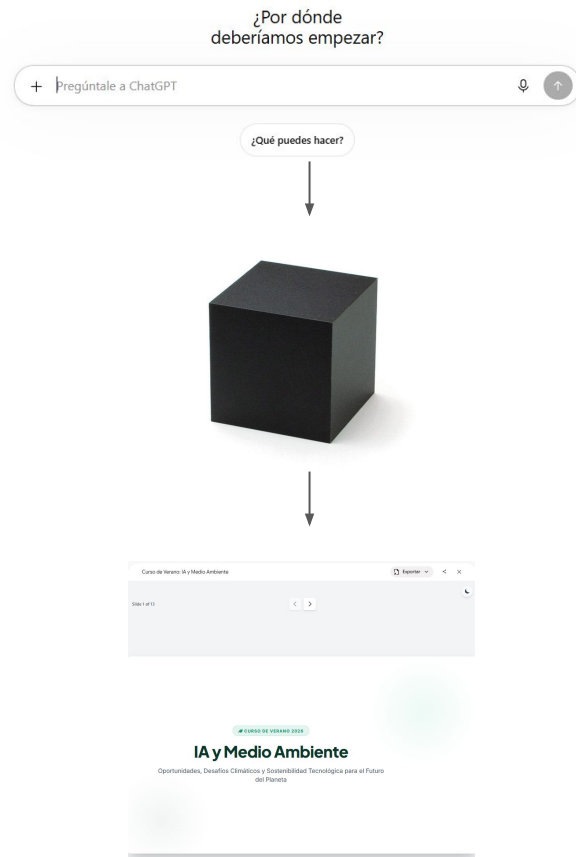
Escalado de
características

Datos categóricos

Datos desbalanceados

Reducción dimensiones

Pero ya por defecto...



Y qué hay del Medioambiente?

Está altamente monitorizado y un gran generador de datos sobre los que se toman decisiones.

La IA se puede utilizar para:

- Encontrar patrones
- Realizar predicciones
- Modelar sistemas físicos
- Optimizar sistemas y procesos
- Encontrar relaciones
- ...

The screenshot shows the top section of the European Environment Agency Datahub website. At the top, there is a header with the European Union flag, the text 'An official website of the European Union | How do you know?', 'Environmental information systems', and a language selector set to 'EN'. Below this is the EEA logo and the text 'European Environment Agency | Datahub'. To the right are links for 'About', 'Featured data', and 'EEA data policy'. A large blue banner with the word 'Datahub' in white is prominent. Below the banner, a paragraph states: 'The EEA's knowledge work relies on high-quality data. Through our network and other institutional partners across 38 European countries, we collect, quality-assure and quality-check data on a wide set of topics and legislation related to the environment, climate and sustainability. This datahub allows you to explore and download these data.' Below this text is a 'Featured' section with three items: 'Featured data Copernicus' (with a satellite image), 'Featured data Statistical data' (with a line graph), and 'What's next About datahub' (with a blue abstract image).

An official website of the European Union | How do you know? ▾ Environmental information systems ▾ EN

 **European Environment Agency** | **Datahub**

[About](#) [Featured data](#) [EEA data policy](#)

Datahub [Share](#)

The EEA's knowledge work relies on high-quality data. Through our network and other institutional partners across 38 European countries, we collect, quality-assure and quality-check data on a wide set of topics and legislation related to the environment, climate and sustainability. This datahub allows you to explore and download these data.

Featured



Featured data
Copernicus



Featured data
Statistical data



What's next
About datahub

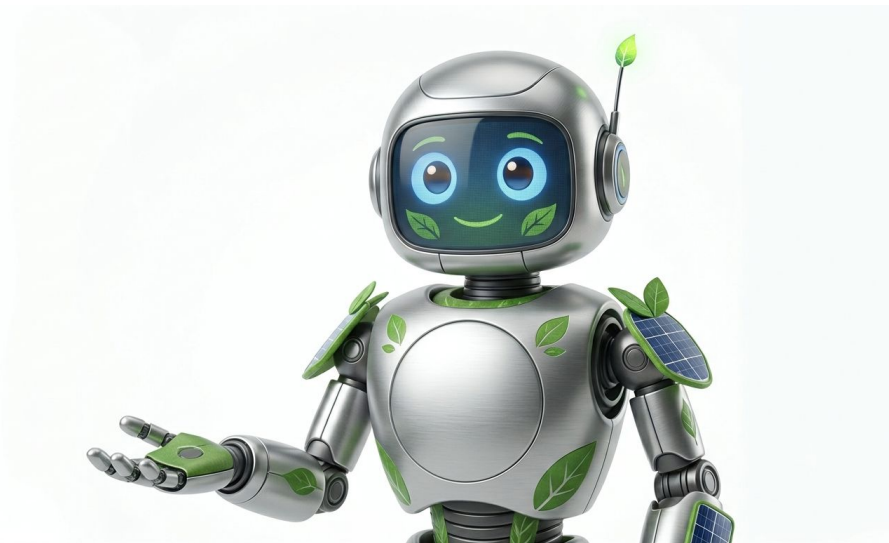
Estado del arte IA en medioambiente

Table 2: Summary of Emerging Applications of AI and ML in Environmental Health Science
(Adapted from Kittichai et al., 2021; Watts et al., 2018)

Application Area	AI/ML Techniques	Input Data	Primary Outputs/Goals
Air Quality Monitoring and Pollution Prediction	Random Forest, SVR, DNN, LSTM	Meteorology, satellite AOD, land use, traffic, sensor data	PM _{2.5} and NO ₂ forecasting, pollution mapping, short-term alerts
Water Quality Assessment and Anomaly Detection	ANN, Decision Trees, K-Means, Autoencoders	pH, BOD, turbidity, DO, sensor data	Real-time water quality prediction, anomaly detection, contamination alerts
Vector-Borne Disease Prediction	Random Forest, XGBoost, CNN, SVM	Temperature, rainfall, NDVI, GIS layers, disease records	Outbreak prediction, vector hotspot mapping, habitat classification
Climate Change Health Impact Modeling	LSTM, Regression, Decision Trees, NLP	Climate projections, temperature, flood data, disaster records, social media	Heatstroke prediction, flood risk, disease shifts, mental health surveillance
Toxic and Chemical Exposure Risk Analysis	QSAR (RF, SVM), DNN, GNN, Ensemble Models	Molecular descriptors, SMILES strings, toxicogenomic data	Chemical classification, toxicity prediction, exposure biomarkers

Hands on IAGEN

Intentaremos tomar decisiones utilizando un dataset y un agente preparado para darnos respuestas sobre los datos.



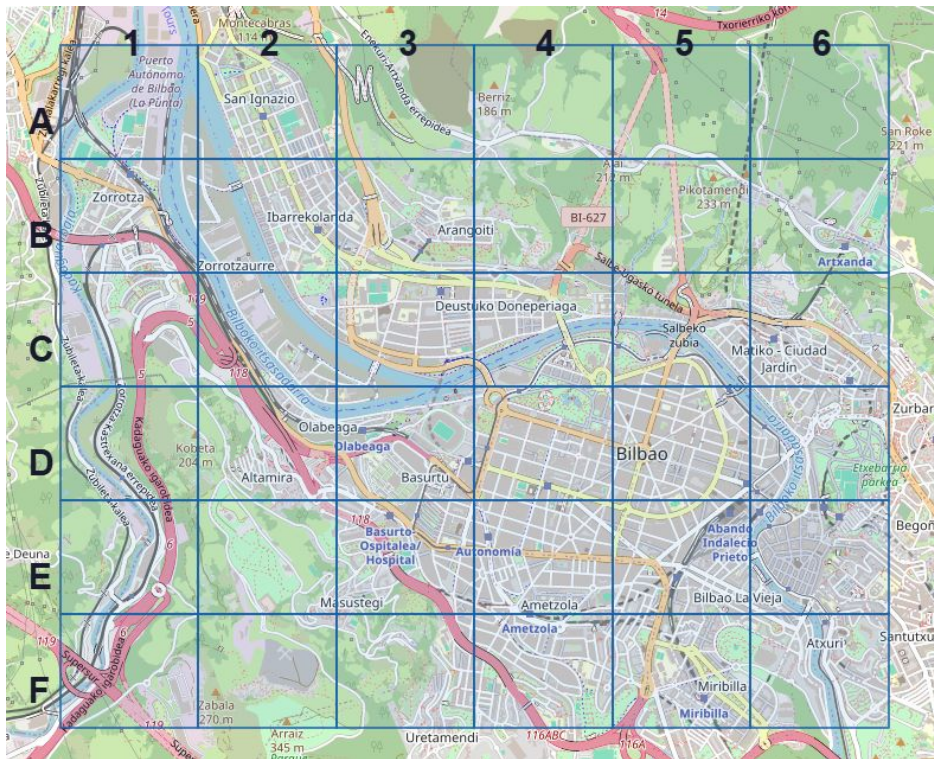
Ingurubot



Objetivo:

Un técnico de medioambiente sale a una hora del día con la misión de diseñar una ruta que maximice la probabilidad de encontrar los picos de contaminación reales de la ciudad ese día, sabiendo que esos picos no son aleatorios: dependen determinados patrones y eventos que ocurren en la ciudad.

El agente tiene acceso a los datos producidos por una red de sensores y puede dar información (y opinión) sobre los mismos.



Ingurubot

Valiéndose de la información que puede dar el agente a partir de los datos:

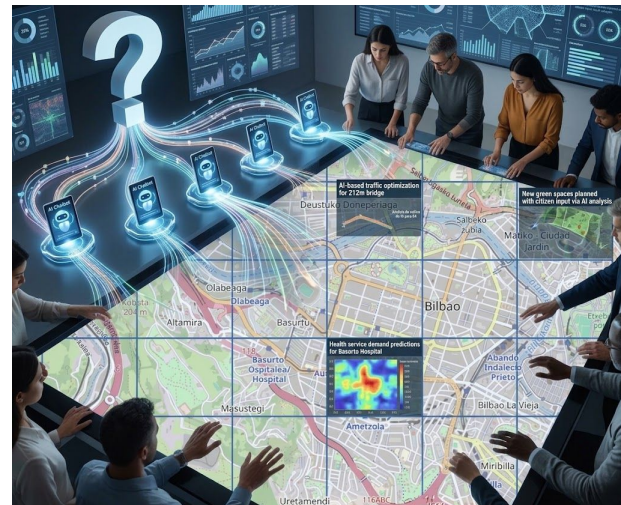
- ¿Cómo se podría diseñar una ruta para la inspección?
- ¿A qué se podrían achacar las principales fuentes de contaminación?
- ¿Qué eventos tienen impacto en las mediciones tomadas?



Enfoque clásico

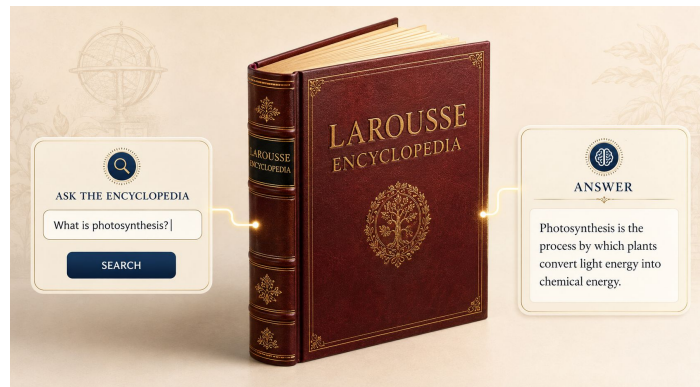


- Habéis preguntado a InguruBot por datos de un área geográfica (Bilbao)
- Habéis extraído información de él, porque queréis resolver un problema formulado: encontrar rutas.
- ¿Habéis pensado sobre el problema formulado? ¿Creéis que abordar un problema con IA es resolverlo con *prompts*? Vamos a verlo.



¿Debemos usar IA generativa?

- Cuando queremos consultar datos sin saber de la tecnología de la base de datos que lo almacena.
- Cuando queremos hacer operaciones sencillas de procesamiento (e.g. comparar espacial/temporalmente valores de la rejilla).
- Cuando queremos narrar sobre los datos, aprovechando el conocimiento de propósito general del modelo.



Si queremos explorar los datos, la IA generativa (InguruBot) sirve para mucho, porque traduce lo que queremos hacer (expresado en lenguaje) y devuelve la respuesta, incluso aumentada.

¿Podemos hacerlo mejor?

- El problema es encontrar la mejor ruta (mejor = $f(\text{polución})$). Pero:
 - Hay otras variables en la base de datos (temperatura, humedad).
 - La ciudad son calles, no celdas. Atravesar una celda puede llevar más tiempo que otra, aunque sean del mismo tamaño.
 - Cuando uno camina por la calle, el tiempo avanza, y con ello cambian las variables asociadas a su celda.
- Cualquier solución que déis está sesgada por el enunciado del ejercicio: una pregunta + un ChatBot.
- Si no *pensamos* y *definimos* bien el problema, obtendremos respuestas, pero nos llevarán tiempo o serán tan completas como nos fiemos del “bicho”...



Decisiones que una IA no debe decidir por sí misma

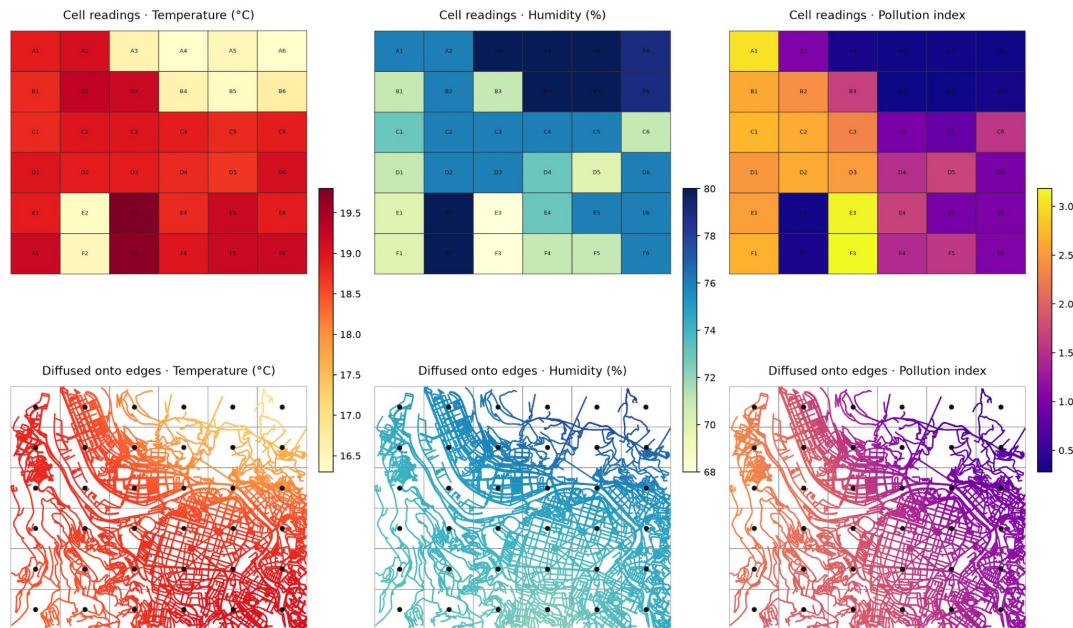
Decisión	¿Por qué no es un prompt?
Celda versus calle	No es lo mismo cruzar el Casco Viejo que la ribera de Deusto.
Peatón vs. coche vs. bici vs. ...	Diferentes calles, diferente ética en el uso de medios de transporte.
Un contaminante vs. tres	NO _x , PM _{2.5} y PM _{10} no reflejan lo mismo.
Exposición media vs. la peor calle	Un pico puntual de polución en un puente vs. un paseo largo con polución moderada.
Rangos de confort 18–24 °C, 40–60 %	El camino lo ha de hacer una persona.
Hora puntual vs. tiempo real a lo largo de la ruta	Caminata de ~70 min a 1.33 m/s... ¿a qué hora pasa por cada celda?

- La IA generativa puede darnos una lista de posibilidades para modelar nuestro caso de uso.
- Nos “da ideas” sobre qué tener en cuenta, teniendo en cuenta los datos a los que tiene acceso, e información sobre lo que nos interesa.
- Somos nosotros quienes debemos definir qué aplica a nuestro caso, valiéndonos – o no – de las ideas que nos proporciona la IA (centauro IA + humano). Veamos cómo hacerlo.

Hagamos el caso de uso más realista: celda versus calle (1/2)

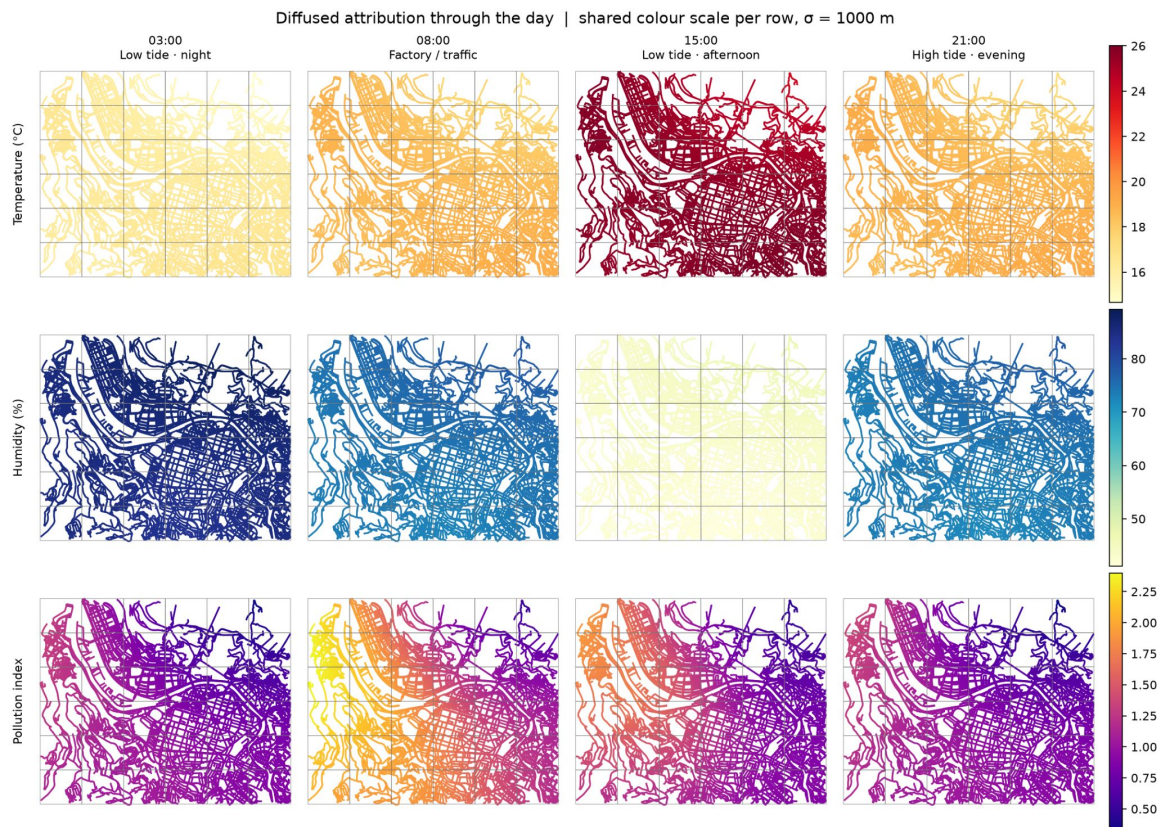
- Tenemos lecturas a nivel de celda, pero un peatón experimenta a nivel de calle.
- Debemos repartir las medidas de celdas a lo largo de las calles de Bilbao.
- Extraemos un grafo (aristas: calles, nodos: intersecciones) de la ciudad.
- En la figura (fila inferior), vemos 3 ejemplos del resultado de una asignación por mezclas de gaussianas con $\sigma = 1000$ m de las medidas de los sensores.
- Es una aproximación que puede tener alternativas (e.g., simuladores alimentados por medidas, integración de sensores virtuales).

Spatial attribution at 08:00 | normalized Gaussian kernel, $\sigma = 1000$ m



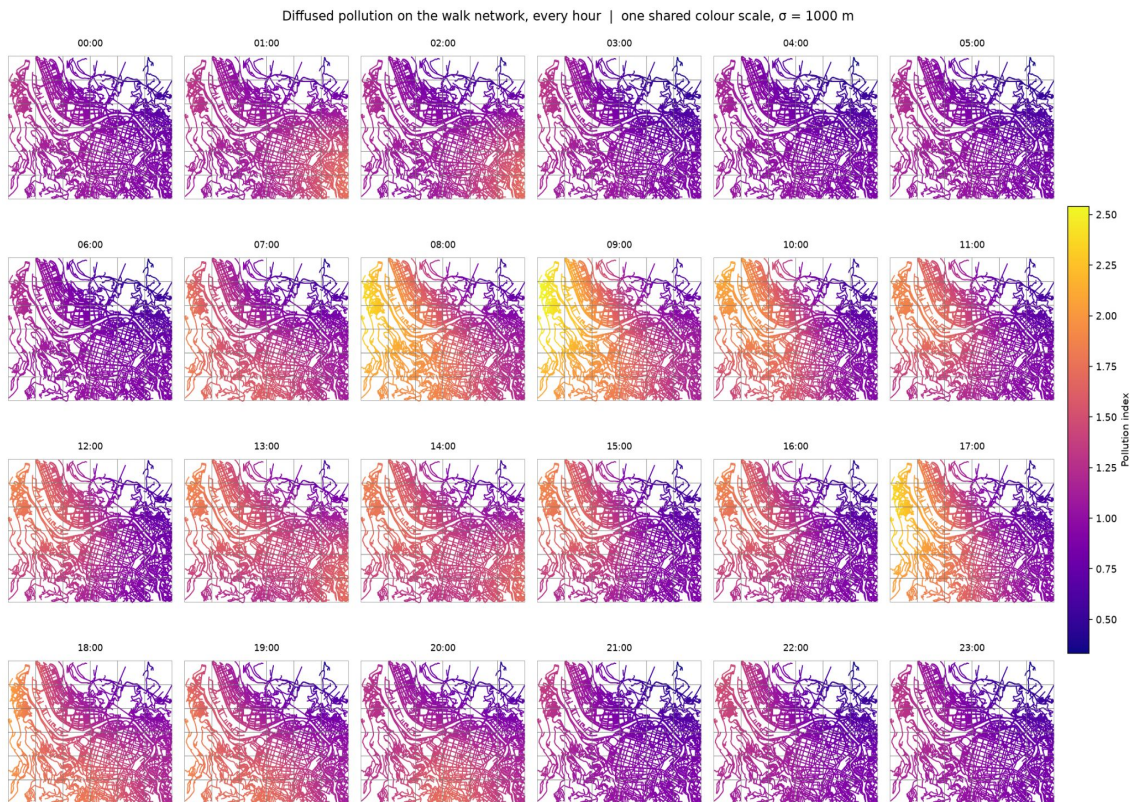
Hagamos el caso de uso más realista: celda versus calle (2/2)

- Definimos un índice de polución como la media de los tres contaminantes.
- Analizamos la asignación espacial de las variables temperatura, humedad e índice de polución:
 - Temperatura y humedad prácticamente estables para las diferentes horas.
 - El índice de polución sí que cambia *espacialmente y temporalmente*, afectando a la “calidad” de las rutas.
- Ese suavizado espacial es por la asignación gaussiana empleada, es una simplificación mejorable.



Hagamos el caso de uso más realista: variabilidad diaria de la polución

- La polución cambia a lo largo del día (lógicamente), tanto en intensidad media como en los patrones espaciales resultantes de la asignación gaussiana.
- Contaminación nocturna.
- Zonas industriales y carreteras en horas punta (jornada laboral).
- Si una persona inicia un paseo de 2 horas de duración a las 7:00 de la mañana, la polución a esa hora es muy diferente a la de las 8:00 y la de las 9:00.
- Necesitamos tener en cuenta la polución a la hora que una persona atraviesa una calle determinada, no a la hora de salida.

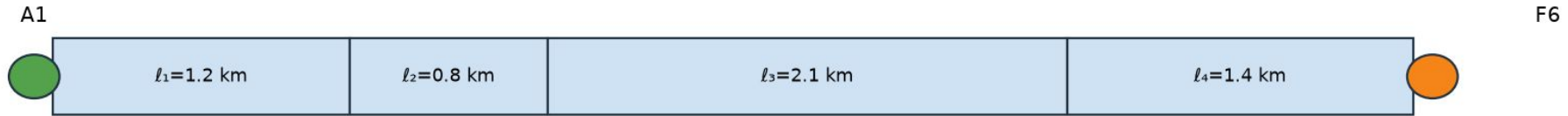


¿Por qué un objetivo y no varios? (1/5)

- Supongamos que fijamos el índice de polución como objetivo a **minimizar**:
 - Buscamos que el peatón no pase por zonas de mucha polución.
 - Buscamos la ruta que **minimice la suma de ese índice** a lo largo de todas las calles que componen una ruta que conecta dos puntos de dos celdas determinadas (supongamos A1 y F6).
- Pero... :
 - ¿Y si por ello la persona tiene que pasar por una calle de **extremo calor**? (por e.g. exposición solar, pocos árboles que den sombra...).
 - ¿Y si por ello tiene que andar **muchos kilómetros**?
 - ¿Y si por ello pasa por zonas de **mucha humedad**? (a sudar...).
- Tengamos en cuenta estos cuatro objetivos, y busquemos no una ruta, sino varias rutas que balanceen de manera óptima entre ellos:
 - Son objetivos “contrapuestos” (Pareto).
 - Definamos dichos objetivos matemáticamente (tranquilidad...).

¿Por qué un objetivo y no varios? Distancia total (2/5)

$$f_1 = \sum_i \ell_i$$

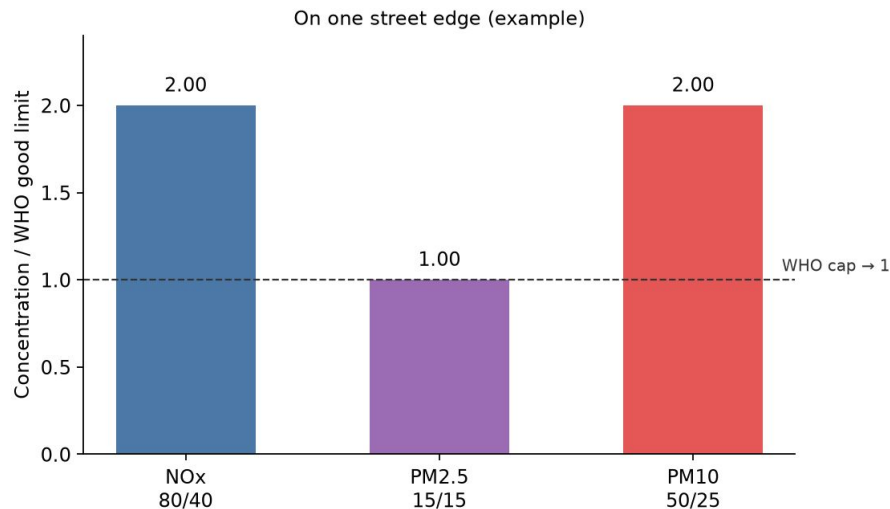


$$f_1(P) = L(P) = \sum_i \ell_i = 5.5 \text{ km}$$

- El primer objetivo no es más que la longitud total de la ruta.
- Queremos rutas cortas entre los dos puntos seleccionados.
- En este objetivo no estamos teniendo en cuenta nada sobre polución, temperatura o humedad → ¡solo distancia!

¿Por qué un objetivo y no varios? Polución (3/5)

$$f_2 = \frac{1}{L} \sum_i \ell_i \pi_i$$



$$\pi = \frac{1}{3}(2.00 + 1.00 + 2.00)$$

$$\pi = 1.67$$

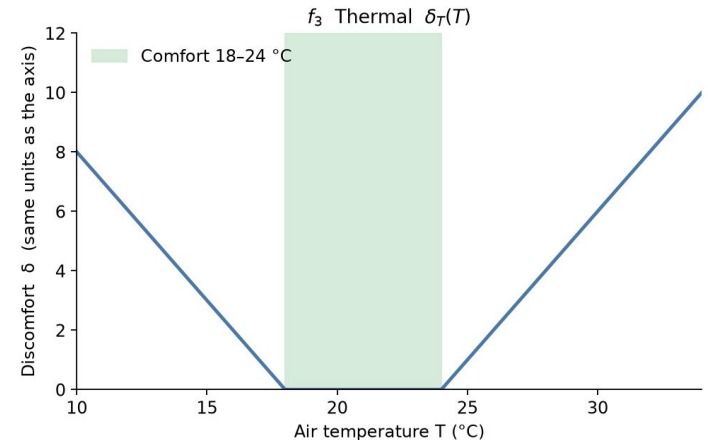
$\pi = 1 \rightarrow$ all three at the WHO cap
 $\pi > 1 \rightarrow$ dirtier than "good"

- Usamos los límites de tolerancia de WHO para cada contaminante para normalizarlos (WHO: World Health Organization).
- El valor del objetivo para esa calle es la media de los valores normalizados de los tres contaminantes. Después agregamos para todas las calles de la ruta (pesado por la longitud de la calle), y normalizamos entre la longitud total de la ruta.

¿Por qué un objetivo y no varios? Temperatura (4/5)

$$f_3 = \frac{1}{L} \sum_i \ell_i \delta_T(T_i)$$

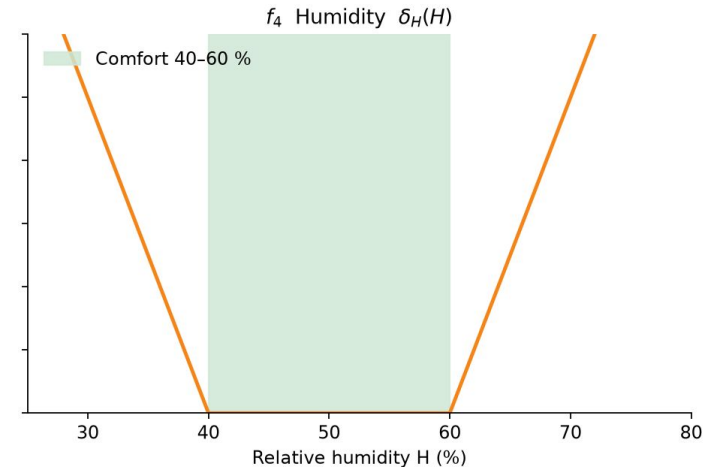
- Definimos un umbral de tolerancia térmica que el peatón es capaz de soportar (18°C - 24°C).
- Penalizamos la ruta si la temperatura de cualquiera de sus calles se sale de este umbral.
- La penalización es mayor:
 - cuanto más se desvíe la temperatura del límite del umbral numéricamente más cercano.
 - cuanto más larga sea la calle en la que la temperatura está fuera de confort térmico.
 - cuantas más calles sufran esa desviación fuera del rango de confort.
- La definición de δ_T aporta lo primero, el pesado por la longitud de la calle aporta lo segundo, y la suma a lo largo de la ruta aporta lo tercero.
- De nuevo normalizamos entre la longitud total de la ruta para que los objetivos sean comparables.



¿Por qué un objetivo y no varios? Humedad (5/5)

$$f_4 = \frac{1}{L} \sum_i \ell_i \delta_H(H_i)$$

- Adoptamos una definición muy similar para la humedad: asumimos que el peatón está cómodo cuando la humedad relativa está en el rango 40%-60%:
- Para ello, usamos una función similar a la usada para el confort térmico, donde penalizamos cuando la humedad en una calle de la ruta sale fuera de esos márgenes.
- De nuevo, la penalización es proporcional a cuánto “se sale” de ese margen, sea por encima o por debajo.
 - Obviamente, podemos explorar otros modelos que e.g. dependan del peatón.



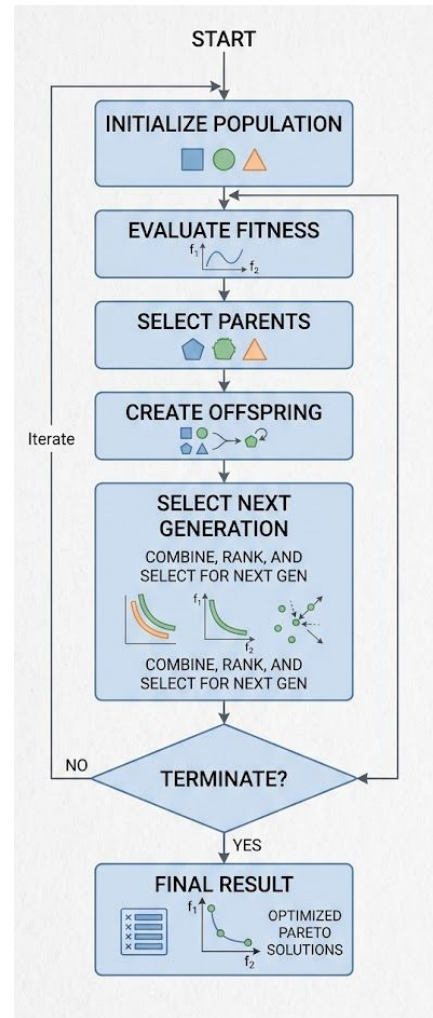
Vale, buscamos minimizar esos cuatro objetivos. ¿Cómo?

- Tenemos un problema de optimización guiado por 4 objetivos **contrapuestos**.
- Parámetros del problema: σ , velocidad peatón, nodos origen/destino.
- Las variables a optimizar son:
 - **Hora de salida** del peatón.
 - **Nodos que componen la ruta** entre los nodos origen (A1) y destino (F6).
- **No hay una solución única**, ya que una ruta puede minimizar la distancia pero ser fatal para la polución, y/o para la temperatura.
- Sí que existe un **conjunto óptimo de rutas**, para las que sabemos que mejorar un objetivo implica necesariamente el empeoramiento en al menos uno de los demás objetivos (**conjunto de Pareto**).



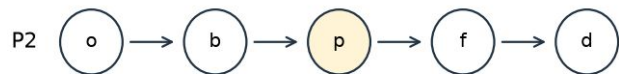
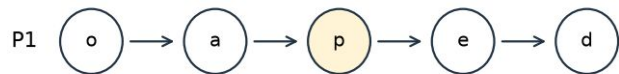
Resolvemos el problema con IA, pero no la generativa

- Para resolverlo, usaremos **algoritmos evolutivos para optimización multi-objetivo**:
 - Algoritmos que **aprenden** por sí mismos a encontrar el conjunto de rutas óptimas (en el sentido de Pareto).
 - Manejan un conjunto (población) de soluciones, y generan poco a poco nuevas soluciones a partir de las anteriores.
- Los cambios son mezclas y perturbaciones entre las **representaciones numéricas** de las soluciones (números que codifican hora y nodos).
- La evaluación de cada nueva solución supone calcular f_1 , f_2 , f_3 , f_4 teniendo en cuenta los valores de T^a , humedad e índice de polución a la hora de salida (también variable).
- La selección de las soluciones que se mantienen en la población depende del metaheurístico seleccionado (hay un montón: nuestros resultados con NSGA2).

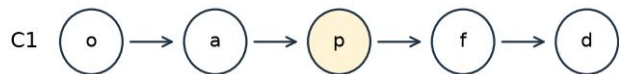


Operadores de mezcla (cruce) y mutación (perturbación)

Crossover — splice at a shared node



pivot p → child keeps one parent until p, then the other



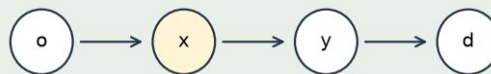
If no shared node: shortest-path bridge; loops are stripped

Mutation — hour or a stretch of street

With probability ~ 0.4 : change departure hour

$h \rightarrow$ a different hour in 0 ... 23

Otherwise: replace a random subpath



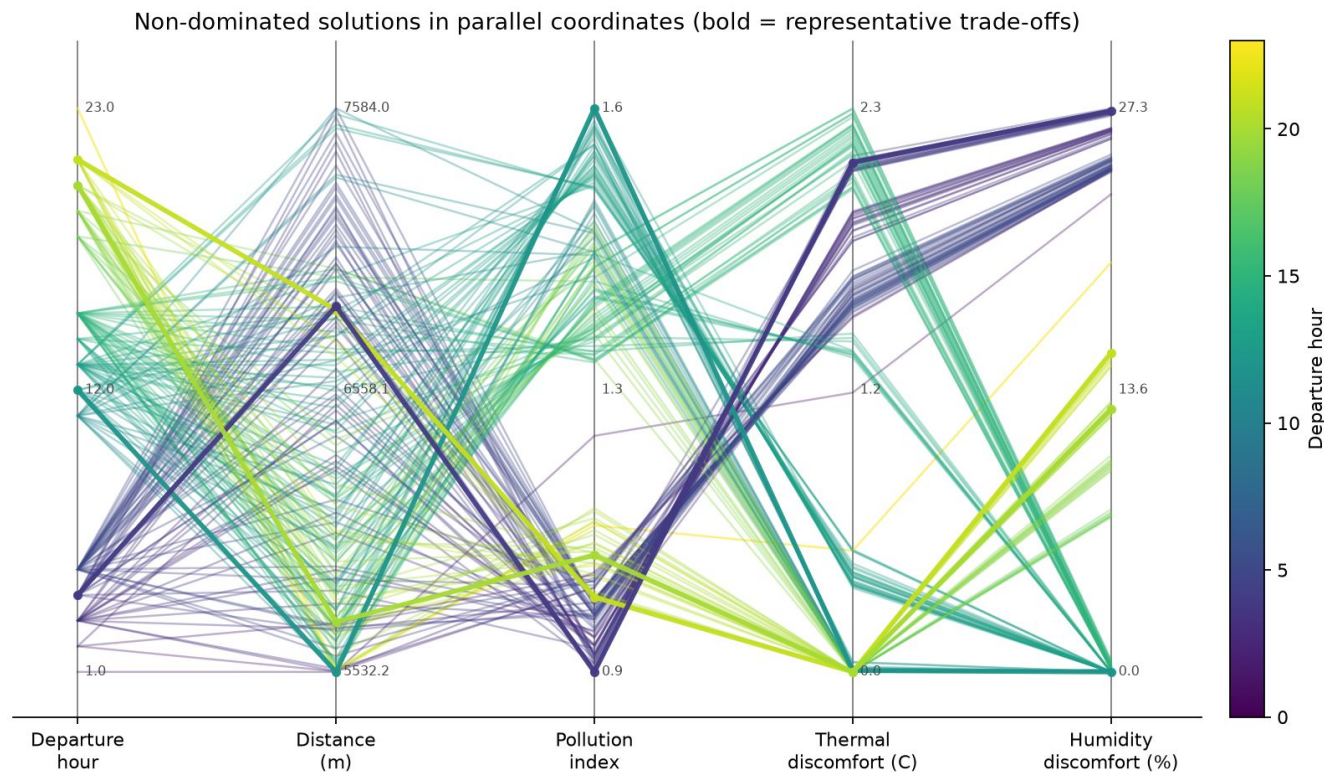
$x \dots y$ replaced by another randomised walk

- Están controlados por una configuración del algoritmo que impone la frecuencia a la que se aplica (probabilidades de cruce y mutación).
- Mejor búsqueda si están definidos de acuerdo al problema.

Configuración del escenario

- Velocidad del peatón: 1.33 m/s (= 4.7 km/h).
- Celdas origen/destino: A1/F6.
- Nodos origen/destino: aleatoriamente de cada celda.
- Población del algoritmo evolutivo: 200.
- Asignamos las variables a las calles de la ciudad en función de la hora a la que se pase por ellas.
- 30,000 evaluaciones máximas = 150 generaciones de búsqueda.
- Pintaremos los resultados como:
 - Diagramas de ejes paralelos: un eje por objetivo, una línea por cada ruta Pareto-óptima descubierta atravesando los cuatro ejes por el valor que tiene la ruta para cada uno de los objetivos.
 - Rutas pintadas en el mapa de Bilbao.

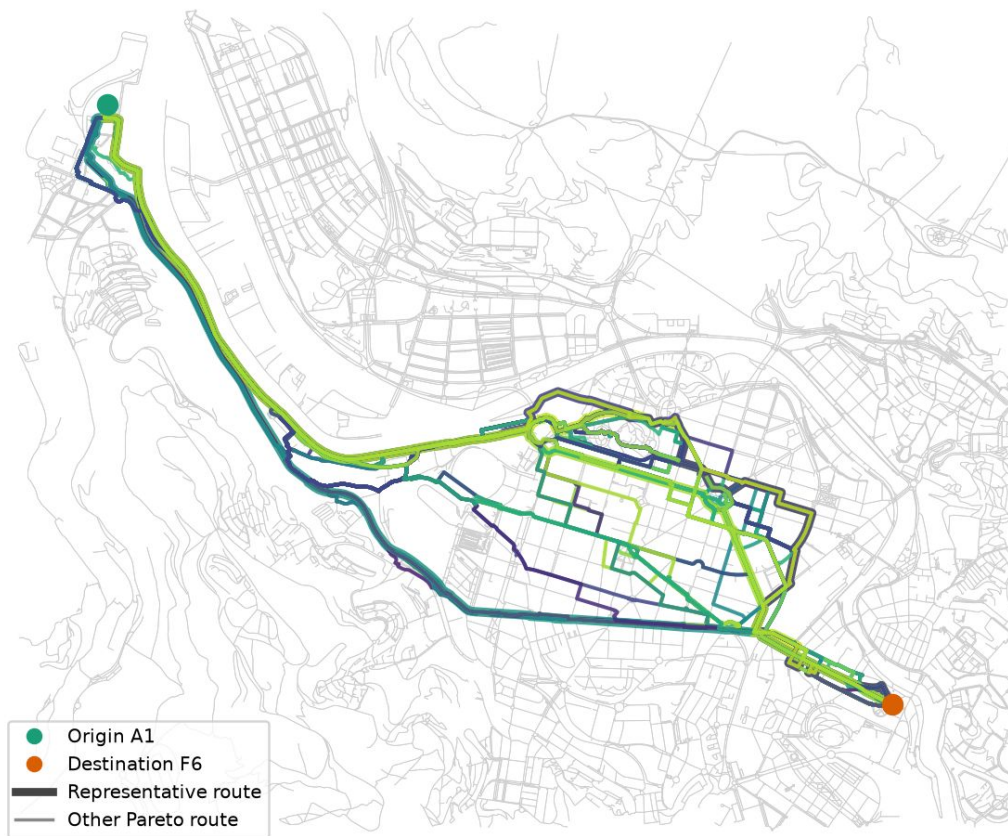
Resultados (1/2)



- Añadimos la hora de salida como un eje adicional (izquierda).
- Vemos efectivamente que hay Paretos entre los objetivos.
- Hay rutas que uno no habría descubierto si no hubiésemos formulado un problema multi-objetivo (soluciones que cruzan = rutas que balancean entre los diferentes objetivos).
- Polución mínima a las noches, máxima en horas centrales, intermedias a última hora.

Resultados (2/2)

- Representamos varias soluciones en el mapa.
- Verde, origen en A1.
- Naranja, destino en F6.
- Pensemos que como las horas de salida cambian, la “ciudad” que ve el peatón en cuanto a los objetivos es diferente.
- ¿Cuál recomendarías para un adolescente que camina a la escuela? ¿Cuál recomendarías para una persona con asma? ¿Cuál recomendarías para alguien que odia la humedad porque suda?



Conclusiones

El verdadero potencial de la IA bien aplicada:

- 1) Usar IA generativa para entender bien los datos e inspirarnos con hipótesis, posibilidades, supuestos... (¡ojito con los tokens!).
 - 2) Usar IA clásica + nuestra capacidad para entender problemas ambientales y traducirlos en problemas que proporcionen valor práctico
- La IA moderna tiene un poder transformativo ENORME, pero hay que saber orientarla, gobernar lo que nos dice, y aterrizar problemas que den valor.
 - No todos los clavos se golpean con el mismo martillo: la aplicación de IA a problemas ambientales puede no requerir el uso de LLMs ni GPUs...
 - Si sabéis enunciar problemas medioambientales, tenéis medio camino hecho. Recorred el resto hasta la meta con las botas adecuadas.

Eskerrik asko!

¡Muchas gracias!

Prof. Dr. Ibai Laña
ilana@mondragon.edu

Prof. Dr. Javier Del Ser
javier.delser@tecnalia.com
javier.delser@ehu.eus

9/9/2026