

IA para todas as idades

30 xullo

Axentes

Un axente é unha IA que usa un modelo base, pero con ferramentas, memoria e obxectivos para actuar como un especialista.

Os modelos base son os que xa coñecemos: ChatGPT, Copilot...

Axentes

— — —

Metafóricamente, se ChatGPT pretende funcionar como unha persona, un axente é ChatGPT “vestido” coa roupa dunha profesión concreta.

Avogado, científico, axente de viaxes, etc.

Non só devolve texto, senón que *actúa*.

Axentes

Un exemplo concreto: Operator de ChatGPT

É capaz de operar na web.

Se lle pides que organize unha viaxe, non só propón ideas, se non que pode mesmo comprar os billetes e facer as reservas de hotel.

Axentes

— — —

Característica	Modelo base (ChatGPT)	Axente de IA especializado
Responde preguntas	✓ Si	✓ Si
Planifica e actúa	✗ Non	✓ Si
Usa ferramentas	♦ Ás veces	✓ Sempre
Ten memoria	♦ Limitada ou nula	✓ Personalizada
Ten un rol claro	✗ Xeralista	✓ Especialista

Axentes

Como se crea un axente?

O modelo base recibe un *prompt* (e.g., “actúa como un axente de viaxes, comprando billetes, etc...”).

Despois hai que engadirlle funcionalidade (as “ferramentas”), memoria, etc.

Axentes

— — —

Instruccións específicas: “Actúa como experto en viaxes baratas”

Ferramentas conectadas: Pode mirar webs, usar follas de cálculo, enviar emails...

Memoria personalizada: Recordar os gustos e as viaxes pasadas do usuario.

Planificación de tarefas: Decide e executa de xeito autónomo

Axentes

Satya Nadella (presidente de Microsoft) declarou que, no futuro, os axentes de IA substituirán ás aplicacións software (Word, Excel, Acrobat, etc).



Aspectos éticos da IA

— — —

Como todas as novas tecnoloxías, a IA plantexa numerosos problemas éticos e morais.

Responsabilidade

A intelixencia artificial xa toma decisións en contextos importantes.

Pero cando algo sae mal... quen é responsable?

Responsabilidade

— — —

Responsabilidade legal: Quen responde perante a lei cando hai un dano?

Responsabilidade moral: Quen debería ser considerado culpable segundo a ética?

Responsabilidade técnica: Foi un erro do sistema, da programación ou do uso?

Uso de LLMs no ámbito xudicial

— — —

Xuíz estadounidense usa IA para redactar veredicto cheo de erros.

Avogados multados por usar IA nun escrito con citas falsas.

Anthropic culpa a Claude dos erros nun documento legal.

IA no ámbito xudicial

— — —

Algúns sistemas de IA son usados para avaliar risco de reincidencia (e.g. COMPAS nos EUA) e aconsellar decisións de sentenza

Se unha persoa recibe unha pena inxusta por recomendación dun algoritmo, quen é o responsable?

O xuíz?

O programador?

A empresa creadora?

O Estado?

IA no ámbito xudicial

Prexuizos nos datos: algoritmos que perpetúan discriminación racial ou social.

Falta de transparencia: moitos modelos son caixas negras.

Falsa confianza na "obxectividade" da máquina.

Conducción autónoma

— — —

Os coches autónomos xa circulan en fase de probas ou con uso limitado.

Un coche sen condutor atropela a un peón.

Posibles responsables:

0 fabricante do vehículo?

A empresa de software?

0 condutor que ía dentro, sen conducir?

0 Estado por permitir o uso?

Condución autónoma

— — —

Dilema do tranvía versión IA:

O coche debe salvar o pasaxeiro ou os peóns?

A programación decide por nós?

Responsabilidade difusa: hai moitos axentes implicados na cadea de decisión.

Preguntas

— — —

Deberían os xuíces ter prohibido usar recomendacións de IA?

A quen demandarías se un coche autónomo che causa un accidente?

Pode unha máquina ser responsable moral ou legalmente?

Consecuencias IA no mercado laboral

A IA está transformando rapidamente o mercado laboral.

Algúns sectores verán perdas de emprego, outros novos postos.

IA no mercado laboral

— — —

Alta automatización:

Transporte e loxística (vehículos autónomos)

Atención ao cliente (chatbots, asistentes virtuais)

Administración e tarefas repetitivas

Mayor demanda futura:

Desenvolvemento e mantemento de IA

Ética, regulación e supervisión de sistemas intelixentes

Educación e formación tecnolóxica

IA no mercado laboral

Tipos de impacto

Substitución: postos desaparecen completamente

Complementariedade: humanos + IA colaboran

Transformación: cambian as tarefas dentro do mesmo posto

Retos sociais

— — —

- Desigualdades: non todos os traballadores poden adaptarse ao mesmo ritmo
- Formación: é necesaria unha política pública clara de reciclaxe profesional
- Incerteza: dificultade para prever a evolución a medio prazo

Retos sociais

— — —

- A IA pode liberar ás persoas de tarefas rutinieras
- Pode mellorar a produtividade e crear novos nichos laborais
- Xorden novas profesións: auditor/a de algoritmos, facilitador/a de IA, deseñador/a de interaccións humanas-máquina