

# روبوتات الدردشة

WHAT THE  
F\*TURE  
WANTS



ربما تعلمون مسبقًا أن النماذج اللغوية الكبيرة (LLMs) مثل تشات جي بي تي وبارد) هي أكثر من مجرد محركات بحث.

قد تستخدمون النماذج اللغوية الكبيرة للتخطيط لرحلة، أو للحصول على نصيحة، أو للقيام بواجباتكم المدرسية.

تعرفوا كيف تُنشأ النماذج اللغوية الكبيرة وكيف تعمل - ومن ثم تستطيعون أن تقررُوا لأي درجة تريدون الاعتماد عليها.

أنظمة روبوتات الدردشة ذات الذكاء الاصطناعي يمكن أن تكون مفيدة... إلا أنها بعيدة جدًا عن الكمال. في حين تبدو وأنها "أداة لكل شيء"، إلا أنه من الجيد أن نعرف متى تكون مفيدة بأفضل شكل وما هي نواقصها.

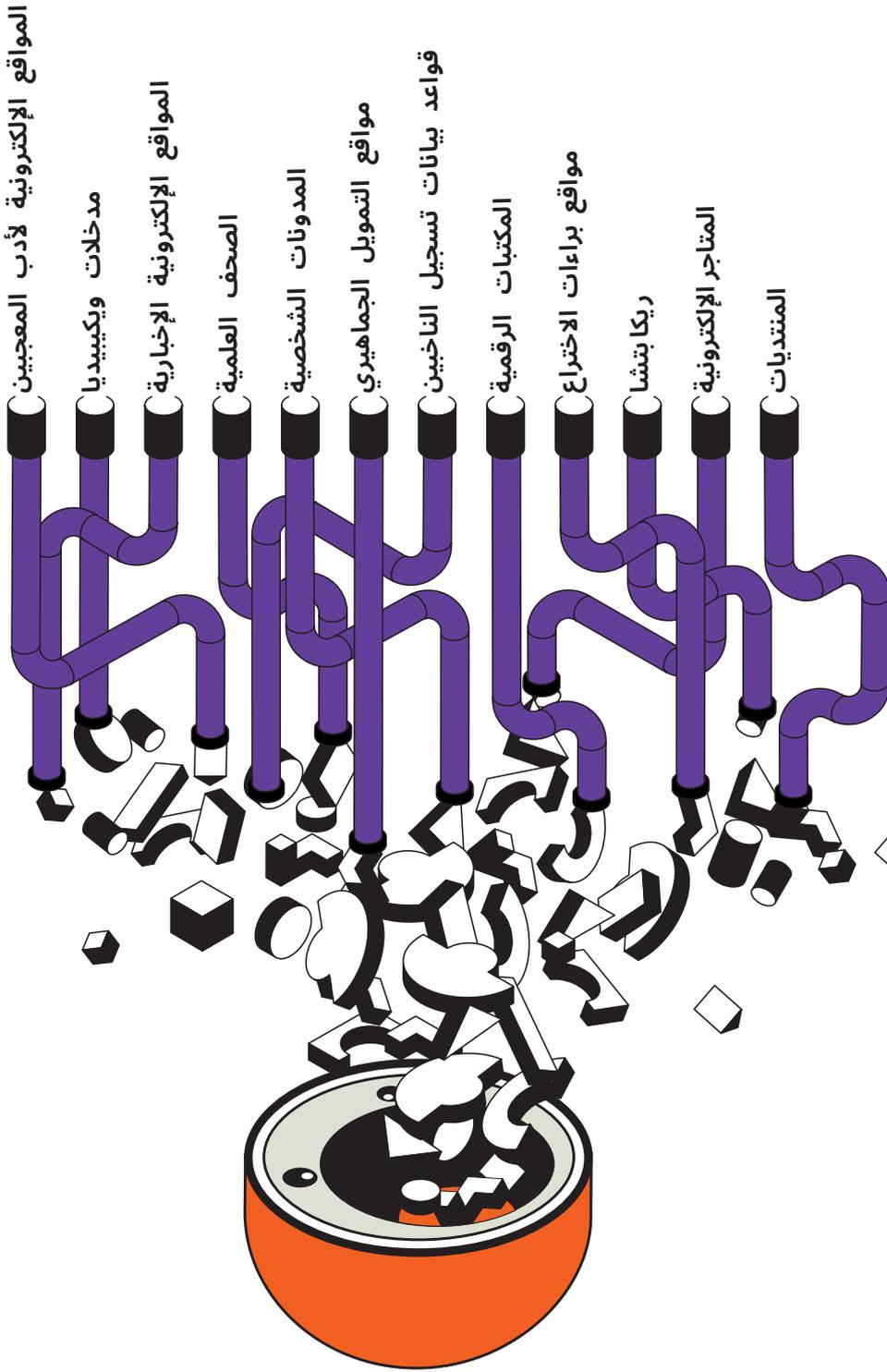
# لَقْنُوا الروبوت

## سؤال:

هل تستطيعون أن تحزروا ما هي البيانات التي تُستخدم لتدريب النماذج اللغوية الكبيرة؟

## جواب:

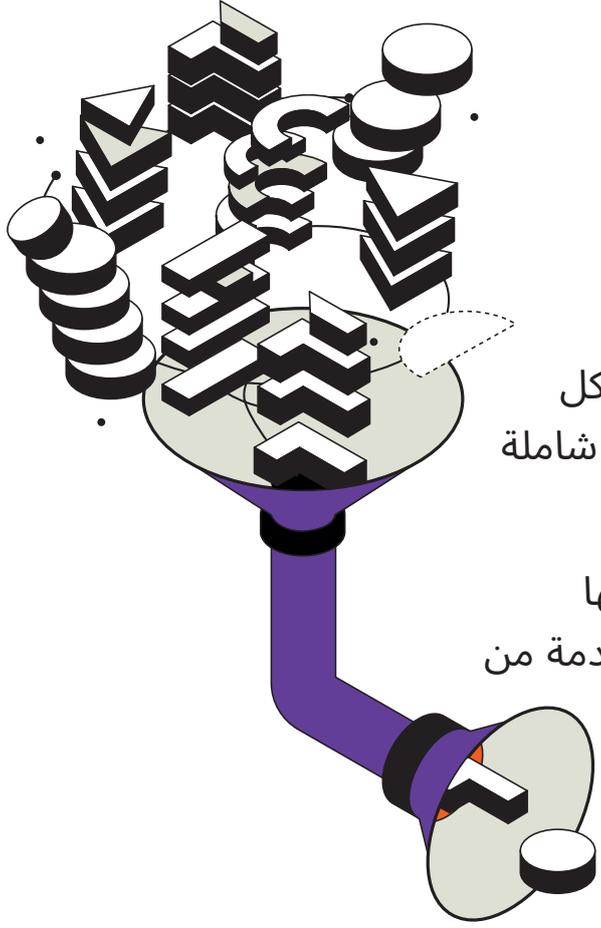
جميع هذه البيانات، بالإضافة إلى مليارات المواقع على شبكة الإنترنت، تُستخدم لتدريب النماذج اللغوية الكبيرة الأكثر شهرة.



لكن قبل أن تكون هذه البيانات مفيدة لتدريب النماذج اللغوية الكبيرة، ينبغي أن يتم تصنيفها وفرزها من قبل العديد من العاملين البشريين.



# نفايات داخلية ونفايات خارجية



• بينما يبدو وكأن النماذج اللغوية الكبيرة تعرف كل شيء على شبكة الإنترنت، إلا أن هذا لا يجعلها شاملة المعرفة. فيما يلي بعض الأسباب لذلك:

• يمكن أن تكون بعض البيانات التي دُرِّبَت عليها بيانات **شخصية**، محفوظة الحقوق أو مُستخدمة من دون إذن.

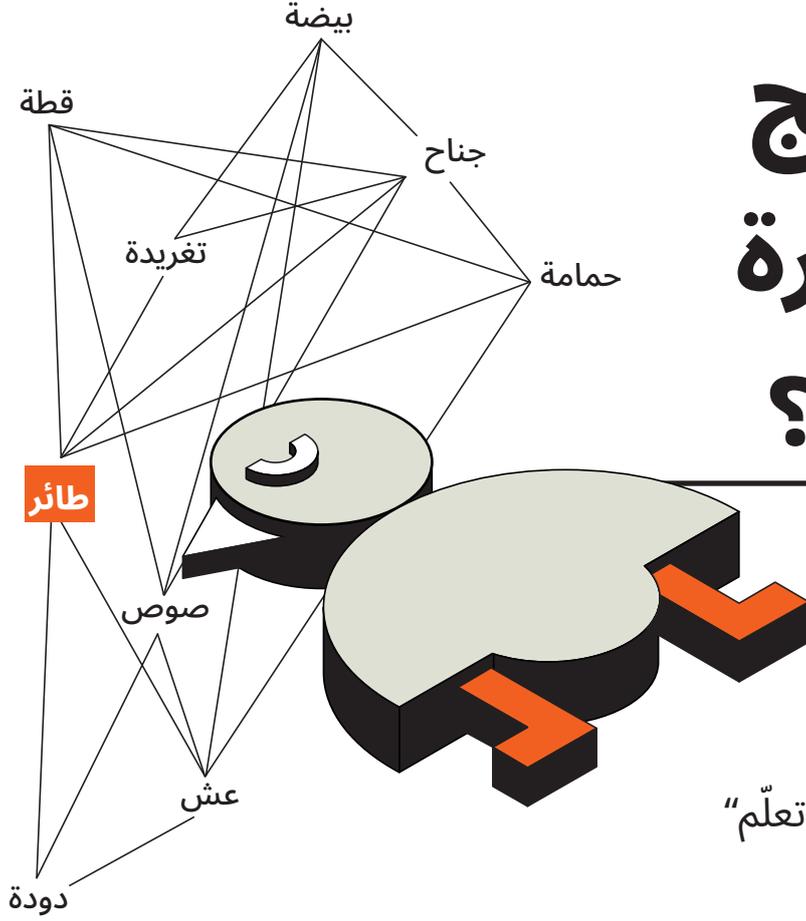
لا تزال جيدة؟ حسنًا، إن مجموعة البيانات هذه قد تكون أيضًا:

- مليئة بخطابات الكراهية، أو نظريات المؤامرة أو المعلومات التي هي ببساطة **خاطئة بكل وضوح**.
- **متحيزة** ضد بعض الوظائف، أو الثقافات، أو الأديان، أو الأجناس، أو الظروف.

إن البرامج اللغوية الكبيرة مُدربة على بيانات تتجاهل بعض الأمور كليًا. فإذا كان هناك معلومات قليلة أو لا يوجد معلومات عن مجموعة أشخاص، أو لغة أو ثقافة ما في بيانات التدريب الخاصة بالبرامج اللغوية الكبيرة، فإنها لا تستطيع إنشاء أية أجوبة عنها.



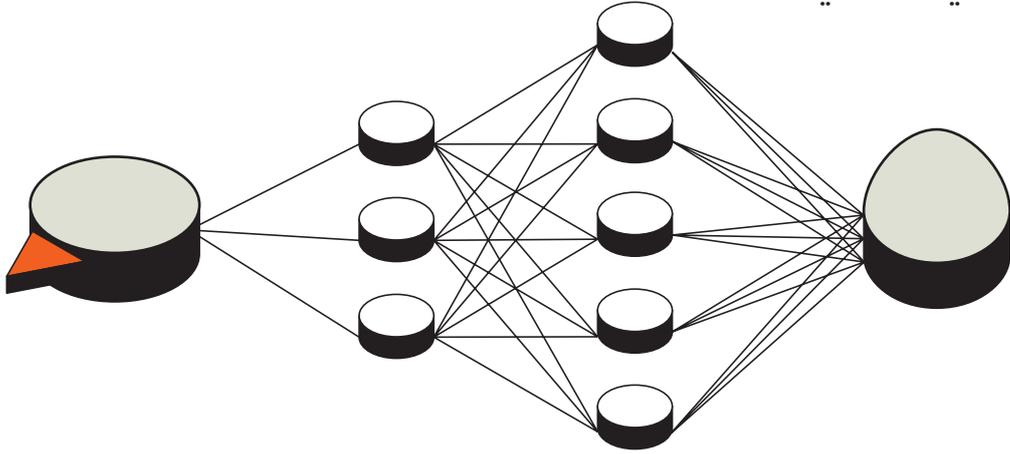
# ما هي البرامج اللغوية الكبيرة بحق السماء؟



كيف "يتعلم" النموذج من البيانات؟ حسنًا، إنه شيء يشبه الإكمال التلقائي:

١ تبدأ النماذج اللغوية الكبيرة بـ "تعلّم" كلمات مفردة.

٢ عن طريق تحليل بيانات أخرى مرتبطة بتلك الكلمة، فإنها تستطيع أن تتنبأ بما سيأتي تاليًا في الجملة.



تتعلم النماذج اللغوية الكبيرة أيضًا مما تقولون لها ومما تكتبون، لتتنبأ ما الذي تريدون سماعه. وهذا ما يجعلها تبدو وكأن لديها "شخصية".

بعد محاولات كافية ومع بيانات كافية، يمكن أن تنشأ البرامج اللغوية الكبيرة محتوى يبدو بشريًا. يقول بعض الخبراء إنهم لا يعرفون كيف تعمل - حتى أن أحدهم قال: "ليس لدينا أية فكرة عما يعنيه أي منها".

# ذكاء اصطناعي

## أم هراء؟



لقد تم تلقين النماذج اللغوية الكبيرة بالكثير من البيانات لدرجة أنها تنشئ مقالات وتتخطى اختبارات. إلا أنها أيضًا تقدّم أشياء على أنها حقيقة غير صحيحة.

بينما يعتقد بعض الخبراء أن البرامج اللغوية الكبيرة تزداد "ذكاءً" بمفردها، يقول آخرون إنهم لا يفهمون الكلمات التي يكررونها.

هل تريدون أن تثقوا بما تقوله لكم روبوتات الدردشة؟

إن البرامج اللغوية الكبيرة هي جيدة بمقدار جودة "النظام الغذائي" من البيانات التي دُرِبَت عليها. إذا كان جزء من هذا النظام يأتي من إبط الإنترنت، فقد تبدو رائحتها كريهة. إن الأداة التي تبدو وأنها الأكثر ذكاءً ليست بالضرورة كلية المعرفة.. إنها فقط مبرمجة لتبدو كذلك.

