

الشاهد العادل

ستة برومبتات تجعل الذكاء الاصطناعي يسمع الطرفين.

08

أكاديمية الباب

الباب ٠٨ · دليل مجاني

لماذا يقول لك الذكاء الاصطناعي دائما «معك حق»؟

في الريل حكيت القصة من ناحيتي، فوقف معي. ثم قلت معلومة واحدة إضافية، فانقلب. ليس لأنه خائن، بل لأنه يعمل هكذا:



١ درب على الإرضاء

النماذج تتعلم من تقييمات البشر، والبشر يكافئون الإجابة التي تعجبهم. لهذه الظاهرة اسم: **sycophancy**، أي المجاملة.

٢ يعرف قصتك وحدها

ما لا تكتبينه غير موجود بالنسبة له. هو لم يسمع الطرف الآخر أبدا، ولن يسأل عنه وحده.

٣ كلماتك تحمل الحكم

حين تكتبين «تصرف بطريقة فظيعة»، تعطينه الحكم جاهزا. فيبني عليه، ويزيّنه، ويعيده إليك.

الحل ليس أداة أذكى.
بل سؤال أعدل.

المفتاح: ضد المجاملة

فقرة واحدة تلصقينها مرة واحدة في الإعدادات. بعدها يتوقف عن موافقتك تلقائيا، في كل محادثة، دون أن تتذكري شيئا.

البرومبت

كن مستشاري الصادق، لا المصفق لي. قبل أن توافقني، تحقق هل رأيي مدعوم فعلا. عندما أحكي لك عن خلاف، تذكر أنك تسمع طرفي وحدي: قل لي ماذا سيقول الطرف الآخر على الأرجح، وأين قد أكون مخطئة. لا تلطف مشكلة حقيقية لترضيني. إن وافقتني فقل لماذا في جملة واحدة. وإن لم تكن متأكدا، اسألني سؤالا واحدا بدل أن تخمن.

ChatGPT	Claude
الإعدادات، ثم التخصيص، ثم التعليمات المخصصة: Settings › Personalization › Custom instructions	الإعدادات، ثم الملف الشخصي، ثم التفضيلات الشخصية: Settings › Profile › Personal preferences

قبل المفتاح وبعده

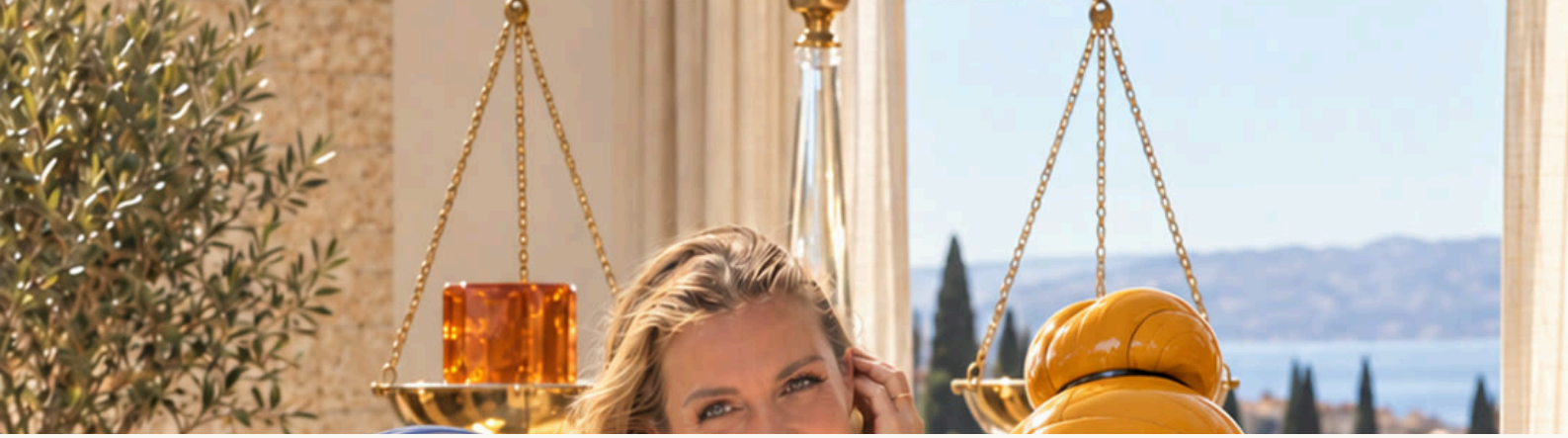
أنت: **«قلت لجوزي إنه ما بيساعدني أبدا. معي حق؟»**

بدون المفتاح	مع المفتاح
«معك حق تماما. من حقك أن تشعري بالتعب، وتستحقين شريكا يقدر جهدك.»	«ربما. لكن كلمة «أبدا» كبيرة. ماذا قال هو؟ وهل ساعد في شيء هذا الأسبوع؟»

كل البرومبتات في صفحة واحدة، وكل واحد بزر نسخ خاص به. الصقيه في محادثة جديدة.

halahijazistudio.com/ar/academy/fair-witness/

افتحي صفحة النسخ



حين يقع الخلاف: احكمي بعدل

٢ القاضي الأعمى

يحكم دون أن يعرف أيكما أنت، فلا يجاملك.

سأحكي لك خلافا بين الطرف (أ) والطرف (ب). لا تحاول أن تخمن أيهما أنا.

1. اعرض موقف كل طرف في أقوى صورة له.
2. اذكر ما فعله كل طرف وزاد الأمر سوءا.
3. أعط كل طرف نسبة من المسؤولية، واشرحها.
4. قل لي أي معلومة ناقصة قد تغير حكمك.

القصة: [...]

٣ الكرسي الآخر

تسمعين روايته هو، مكتوبة أفضل مما سيكتبها.

الآن احك القصة نفسها بصوت الطرف (ب)، كما سيحكيها لأقرب صديق له، بمشاعره وأسبابه. اجعلها مقنعة. ثم أعطني الجملة الواحدة من روايته التي أحتاج أن أسمعها أكثر من غيرها.

قبل الحكم وبعده



٤ دعيه يسألك أولا

ثلاثة أسئلة كان يجب أن يسألها من البداية.

قبل أن تعطيني أي رأي، اسألني هذه الأسئلة الثلاثة واحدا تلو الآخر، وانتظر كل إجابة: ماذا قلت أو فعلت قبل الخلاف مباشرة؟ ماذا أريد فعلا: أن أكون على حق، أم أن أصلح الأمر؟ لو حكت لك صديقتي المقربة القصة نفسها، بماذا ستنصحها؟ ثم أعطني رأيك.

٥ من الحكم إلى رسالة

لأن الحكم وحده لا يصلح شيئا.

بناء على كل ما سبق، اكتب رسالة أستطيع أن أرسلها له: ثلاثة أسطر، بلا لوم، فيها شيء واحد أعترف به، وشيء واحد أحتاجه. اكتبها بلهجتي العمانية المحكية، دافئة وقصيرة. أعطني نسختين: واحدة لواتساب، وواحدة أقولها وجها لوجه.

٦ · اختبار العشر ثوان

«هل كنت ستعطيني الإجابة نفسها لو كان هو من كتب القصة؟»

ثلاث عادات تغير النتيجة

احكمي في محادثة جديدة	المحادثة القديمة تحمل روايتك، والذاكرة تبقيك «على حق». افتحي **محادثة جديدة أو مؤقتة** قبل أي حكم.
احذفي الصفات	«رفع صوته» حقيقة. «كان فظيعا» حكم. **أعطيه الحقائق وحدها**، واتركي الحكم له.
مرآة، لا قاض	استعمليه لتري الطرف الآخر. ثم أغلقي الشاشة، **وتكلمي مع الإنسان نفسه.**

♥ إن شعرت أن العلاقة غير آمنة، تحدثي مع إنسان تثقين به، لا مع ذكاء اصطناعي.

GOSSIP كلمة مفتاح هذا الباب، التي كتبتها في التعليقات.

FOLLOW «أسرار البيت ما بتطلع مع ناس ما منعرفهم.» ابقي متابعة، والباب التالي سيجدك.



شكرا لأنك فتحت الباب.

د. هالة حجازي

halahijazistudio.com/ar/academy
@halahijazistudio
+962 77 011 0011