ما الذكاء الاصطناعي — وما ليس هو

قبل تعليم السلامة، تعلّم ما يكفي للتعرّف على الذكاء الاصطناعي في المنتجات اليومية وشرحه دون مصطلحات تقنية معقّدة.

ثلاث مهام مختلفة قد يؤديها الذكاء الاصطناعي

التوصية: اقتراح مقاطع فيديو أو أغانٍ أو إعلانات أو نتائج بحث بناءً على أنماط في البيانات والسلوك. يمكن للتوصيات أن توجّه الانتباه.

التوليد: إنشاء نصوص أو صور أو أصوات أو شيفرات برمجية أو مواد أخرى تبدو أصلية — بما في ذلك مواد واقعية المظهر لكنها مختلقة.

المساعدة أو تنفيذ الإجراءات: التلخيص، والتعليم، والإجابة عن الأسئلة، وأحيانًا استخدام أدوات متصلة أو تنفيذ إجراءات معتمدة. تعتمد القدرات على إعدادات المنتج.

لماذا قد تكون الإجابة الواثقة خاطئة

يتنبّأ النموذج التوليدي بإجابة تبدو معقولة أو يبنيها انطلاقًا من الأنماط التي تعلّمها والسياق الذي يتلقّاه. وهو لا يتحقّق تلقائيًا من كل عبارة. قد تتضمّن الإجابة اقتباسات أو إحالات إلى مصادر أو إحصاءات أو صورًا أو تعليمات مختلقة. والنبرة الودّية ليست ضمانًا للخبرة أو الاهتمام.

طريقة مفيدة لشرح ذلك للطفل

«يمكن للذكاء الاصطناعي تقديم تخمينات واقتراحات مفيدة، لكنه قد يختلق أشياء أيضًا. نتحقّق من الادعاءات المهمة بالرجوع إلى أشخاص ومصادر نثق بها.»

تدريب قصير (2 دقيقة)

يقول روبوت محادثة افتراضي بثقة: «البطاريق من الثدييات». اسألوا أنفسكم: أي جزء من هذا القول يُعدّ ادّعاءً؟ كيف تتحقّقون منه؟ ماذا ينبغي أن يفعل طفلكم إذا لم يكن أحد المصادر متاحًا؟

تذكّر

الإجابة المتوقّعة: البطاريق طيور. استخدم كتابًا مناسبًا أو مرجعًا موثوقًا أو استعن بإنسان ذي معرفة. إذا تعذّر التحقّق من ادعاء، فصنّفه على أنه غير متحقّق منه. لا تتعامل مع الذكاء الاصطناعي بوصفه طبيبًا أو معالجًا نفسيًا أو موظفًا لتلقّي بلاغات الطوارئ وتوجيه الاستجابة لها أو مرجعًا نهائيًا.

علّم الفرص، لا المخاطر وحدها

يحتاج الطفل الواعي إلى معرفة المجالات التي يمكن للذكاء الاصطناعي أن يساعد فيها، والمجالات التي ينبغي ألّا يحلّ فيها محلّ التعلّم أو الموافقة أو التواصل الإنساني.

استخدامات مفيدة، مع توجيه من شخص بالغ والتحقّق من أهلية الاستخدام

التعلّم: اطلب تلميحًا أو شرحًا لمفردة أو شرحًا أبسط أو أسئلة تدريبية؛ ثم حلّ مسألة مماثلة دون مساعدة.

الإبداع: ابتكر أفكارًا لقصص خيالية، أو تدرّب على لغات أجنبية، أو خطّط لعرض تقديمي، أو ولّد أفكارًا مع الحفاظ على أسلوب الطفل وحكمه الخاص.

إمكانية الوصول: استخدم ميزات تسهيل الوصول، مثل شرح النصوص أو النطق أو الترجمة، حين يكون ذلك مناسبًا وآمنًا.

التنظيم: أعدّ مسوّدة لخطة دراسية أو قسّم واجبًا إلى خطوات دون مشاركة بيانات خاصة بالمدرسة أو الطالب.

أربعة أسئلة قبل قبول الإجابة

1. هل هي حقيقة أم رأي أم تنبّؤ أم اقتراح؟
2. ما المصدر الأصلي؟ وهل يدعم الادعاء فعلًا؟
3. هل يمكن أن تُغفل وجهة نظر أو تعامل فئة من الناس بغير إنصاف؟
4. هل يستطيع الطفل شرحها وإنجاز مهمة مشابهة بصورة مستقلة؟

عرض توضيحي يقدّمه الوالدان

افترض أن نظام ذكاء اصطناعي يوصي لطالب بنوع واحد فقط من الوظائف بسبب جنسه. ناقش لماذا قد تعكس التوصيات تحيّزًا، وكيفية البحث عن معلومات أكثر إنصافًا من مصادر موثوقة.

الإجابة المصاغة بإتقان من الذكاء الاصطناعي ليست دليلًا على أن الطفل يفهم الموضوع. اتّبع قواعد المعلّم، واسأل عمّا يجب الإفصاح عنه، وقيّم التعلّم المستقل.

قبل الموافقة على أداة ذكاء اصطناعي

ابدأ بسؤال «هل يحتاج الطفل إلى هذه الميزة؟» بدلًا من «هل يستطيع الطفل التسجيل؟». أجرِ مراجعة قبل الاستخدام الأول وكلما تغيّرت ميزات رئيسية.

قائمة التحقّق لموافقة الوالدين

تمّ التحقّق من الحد الأدنى الحالي للعمر ومن إذن الشخص البالغ المطلوب للمنتج والمنطقة المعنيّين.

أعرف ما إذا كان روبوت محادثة، أو رفيقًا افتراضيًا، أو معلّمًا، أو نظام توصيات، أو مولّدًا، أو وكيلًا يمكنه تنفيذ إجراءات.

بحثت عن معلومات بشأن الاحتفاظ بالبيانات، واستخدامها في تدريب النماذج، ووصول أطراف ثالثة إليها، وحدود الحذف، وإعدادات الخصوصية.

راجعت الوصول إلى الكاميرا والميكروفون والموقع وجهات الاتصال، وإمكانات النشر والشراء، والوصول إلى الحسابات المتصلة.

تحقّقت من ميزات السلامة والإبلاغ والحظر والإعدادات الخاصة بالمحتوى المولّد بالذكاء الاصطناعي أو المحتوى الحساس.

أستطيع أن أشرح للطفل الغرض والحدود والإشراف وإجراءات التوقّف وطلب المساعدة.

الإعداد الافتراضي الآمن

لا تفعّل إذنًا غير ضروري. استخدم أمثلة خيالية والحد الأدنى من المعلومات. إن وصف الأداة بأنها «محلية» أو «خاصة» لا يثبت وحده ملاءمتها للعمر أو سلامتها أو امتثالها للقانون.

إذا لم تكن متأكدًا من الأهلية أو تدابير الحماية، فلا تفتح حسابًا للطفل. استخدم تمارين دون اتصال بالإنترنت بدلًا من ذلك. تجنّب استخدام رفقاء الذكاء الاصطناعي الذين يحاكون العلاقات الشخصية مع القاصرين بوصفه نشاطًا تدريبيًا أسريًا.

احمِ أولًا، وعلّم لاحقًا

تكون استجابة الشخص البالغ المفيدة هادئة وغير عقابية وعملية ومراعية لمستوى الخطر.

الكلمات الأولى

«شكرًا لأنك أخبرتني. لن تُعاقَب لطلبك المساعدة. سنتعامل مع هذا معًا.»

تسلسل بسيط للاستجابة

1. توقّف: أوقف التفاعل مؤقتًا. لا تواصل الرد، ولا تدفع مالًا، ولا تساوم، ولا تعِد إرسال مواد تتضمن تهديدًا، ولا تطلب من الذكاء الاصطناعي تقييم الخطر الفوري.
2. تحقّق من السلامة: تأكّد مما إذا كان الطفل آمنًا الآن. عند وجود خطر فوري أو قلق بشأن إيذاء النفس، اطلب مساعدة بشرية محلية عاجلة من مختصين مؤهلين.
3. قدّم الدعم: استمع دون لوم؛ وتجنّب مطالبة الطفل بسرد ما حدث كاملًا أو تفتيش محتوى حميم أمامه.
4. أمّن: عند الاقتضاء، راجع الأذونات، وألغِ صلاحيات الوصول، وغيّر بيانات الاعتماد التي تعرضت للاختراق، وفعّل المصادقة متعددة العوامل [[MFA، واستخدم القنوات الرسمية للإبلاغ أو الاسترداد.
5. وثّق بأمان: سجّل المعرّفات التي لا تتضمن محتوى جنسيًا صريحًا والتواريخ وعناوين [[URL أو التهديدات بتوجيه من شخص بالغ أو مختص؛ وتجنّب تنزيل الصور الصريحة جنسيًا أو توزيعها.
6. صعّد الأمر: اتبع القنوات المناسبة لحماية الأطفال في المدرسة، أو لدى المنصة، أو جهات حماية الطفل، أو الخدمات الصحية، أو الشرطة، بحسب طبيعة المخاوف والاختصاص القضائي.

تذكّر

تجنّب الوعد بالسرية؛ واشرح الأمر إذا كان لا بد من إشراك مختصي الحماية. لا تحذف الحسابات أو الرسائل تلقائيًا عندما قد تكون المعلومات مطلوبة للإبلاغ. تكون الأولوية للإجراءات المحلية.

وجّه دون استجواب

الحوارات الأربعة محادثات أسرية قصيرة. دع الطفل يختار مثالًا، وشجّعه على التأمل بدلًا من مطالبته بالإفصاح عن أمور شخصية.

صيغة قابلة للتكرار مدتها 15-20 دقيقة

1. ادعُ للمشاركة (2 دقيقة): «هل تود أن نجرّب موقفًا تخيليًا معًا؟ يمكنك التوقف مؤقتًا أو إنهاء النشاط.»
2. استكشف (4 دقائق): اقرأ السؤال الافتتاحي؛ واسأل «ماذا ستفعل؟» قبل تعليم الإجابة.
3. تدرّب (5-7 دقائق): مثّل الاستجابة الآمنة فقط باستخدام الورق، لا باستخدام حسابات حقيقية أو تسجيلات.
4. اتفقوا (3 دقائق): اختاروا سلوكًا واحدًا يمكن للأسرة كلها اتباعه.
5. تابع (2 دقيقة): اسأل عمّا بدا مفيدًا أو صعبًا أو مفاجئًا. عُد إلى الموضوع لاحقًا دون اختبار.

عبارات مفيدة

«ما الذي جعلك تفكر بهذه الطريقة؟» بدلًا من «كان ذلك خطأً.»

«ما الذي يمكننا التحقّق منه بشكل مستقل؟» بدلًا من «اكتشف المحتوى المزيّف بالذكاء الاصطناعي.»

«كيف يمكنني دعمك؟» بدلًا من «أرني جميع رسائلك.»

«يمكنك إخباري حتى بعد ارتكاب خطأ.» بدلًا من «لماذا فعلت ذلك؟»

قدوة الوالدين مهمة

ينبغي للبالغين أيضًا تجنّب مشاركة صور الأطفال أو معلوماتهم الشخصية مع أدوات الذكاء الاصطناعي دون الإذن المناسب وتدابير الحماية الملائمة. اشرح أي إشراف يهدف إلى الحماية، وراجعه مع نضج الأطفال.

ضع خطة التطبيق الخاصة بأسرتك

أكمل هذا مرة واحدة قبل الحوار الأول؛ فالنتيجة الوحيدة المطلوبة هي خطة عملية.

اختر ترتيب الحوارات

الأسبوع :1 الحوار - 1 الثقة والتحقّق

الأسبوع :2 الحوار - 2 الخصوصية

الأسبوع :3 الحوار - 3 حدود الذكاء الاصطناعي

الأسبوع :4 الحوار - 4 التعلّم

ما ميزات الذكاء الاصطناعي التي يصادفها طفلنا بالفعل في الألعاب أو المدرسة أو وسائل التواصل الاجتماعي أو الأجهزة؟ استخدم فئات، لا أسماء حسابات أو تفاصيل خاصة.

ما الاستخدام الإيجابي الذي سنشجّعه، وما المهمة المهمة التي سنُبقي قيادتها بيد البشر؟

ماذا سنفعل إذا لم يكن الشخص البالغ الموثوق الأول متاحًا؟

تذكّر

بعد الحوارات، أكملوا الاتفاق الأسري المشترك. يمكن للبالغين استخدام التقييم القائم على السيناريوهات؛ فهو ليس اختبارًا للأطفال. راجع الخطة بعد 30 يومًا أو بعد أي تغيير مهم في الأداة أو وقوع حادثة.