HUMANITAIRES JOBS · الأكاديمية الإنسانية

الميني MBA الذكاء الاصطناعي والعمل الإنساني: العمل مع Claude وGemini

مادة الدرس · الأسبوع 4 : أخلاقيات الذكاء الاصطناعي والمبادئ الإنسانية: عدم الإضرار، والتحيّزات، والشفافية

برنامج مدته 3 أشهر بالضبط (12 أسبوعًا، 144 ساعة) لاستخدام مساعدَي الذكاء الاصطناعي Claude (Anthropic) وGemini (Google) في العمل الإنساني، بطريقة مفيدة وآمنة ومسؤولة: طريقة العمل والحدود، وكتابة الطلبات، وحماية البيانات، والأخلاقيات والمبادئ الإنسانية، والتقارير، والمقترحات، وتحليل البيانات وMEAL، والترجمة والتواصل مع المجتمعات المحلية، والبحث والتحقق من المصادر، والتضليل والأمن، والإنتاجية والأتمتة، وسياسة الاستخدام والتعميم في منظمة. تمارين مصحّحة، وثلاثة امتحانات شهرية مع دراسات حالة، وشهادة قابلة للتحقق من الأكاديمية الإنسانية. برنامج مستقل، بلا شراكة مع Anthropic ولا مع Google.

https://humanitairejobs.org/ar/academy/mini-mba-ia-claude-/gemini/semaine-4

وثيقة مجانية من الأكاديمية الإنسانية. يمكن طباعتها ومشاركتها بحرية لأغراض التدريب دون تعديل.

الأسبوع 4 : أخلاقيات الذكاء الاصطناعي والمبادئ الإنسانية: عدم الإضرار، والتحيّزات، والشفافية

الأهداف

- أن تربط استخدام الذكاء الاصطناعي بالمبادئ الإنسانية (الإنسانية، والحياد، وعدم التحيّز، والاستقلال) وبمبدأ عدم الإضرار.
- أن ترصد تحيّزات إجابة الذكاء الاصطناعي وآثارها المحتملة على الأشخاص المتضررين.
- أن تطبّق قواعد واضحة للشفافية، والمسؤولية البشرية، واحترام حقوق المؤلف.

المبادئ الإنسانية في مواجهة الذكاء الاصطناعي

تنطبق المبادئ الإنسانية الأربعة على الأدوات الرقمية أيضًا:

- **الإنسانية**: يجب أن تخدم الأداة الأشخاص المتضررين، لا الكفاءة الداخلية وحدها. توفير الوقت في تقرير مفيد إذا عاد هذا الوقت إلى البرامج.
- **عدم التحيّز**: يجب ألا يُدخل الذكاء الاصطناعي تمييزًا في الوصول إلى المساعدة (مثلًا بترتيب الأسر وفق معايير متحيّزة).
- **الحياد**: قد يتبنّى النص المولَّد، من دون أن يُلحظ ذلك، مفردات أحد أطراف النزاع. راجع أسماء الأماكن والجماعات والأحداث.
- **الاستقلال**: الاعتماد الكامل على مزوّد واحد في وظيفة حيوية خطر؛ حافظ على القدرة على العمل من دون الأداة.

يفرض مبدأ **عدم الإضرار** (*do no harm*) طرح سؤال قبل كل استخدام جديد: من قد يتضرر إذا أخطأت الأداة، أو تسرّبت البيانات، أو أُسيء تفسير النتيجة؟

التحيّزات: مصدرها وكيفية رصدها

تتعلم النماذج من نصوص أُنتجت في معظمها بلغات قليلة وفي مناطق قليلة من العالم. ومن ذلك تنتج **تحيّزات**:

- **التحيّز اللغوي**: جودة أقل في اللغات قليلة الحضور على الإنترنت (لغات أفريقية وآسيوية وأصلية كثيرة)، وأحيانًا في بعض لهجات العربية أو متغيرات الفرنسية.
- **تحيّز المنظور**: قراءة الأزمات «من الشمال»، وحلول نمطية، وتجاهل الفاعلين المحليين والمنظمات غير الحكومية الوطنية.
- **الصور النمطية**: تمثيلات جندرية أو ثقافية («الطبيب» رجل، و«المستفيدة» سلبية) في النصوص، وخصوصًا في الصور المولَّدة.
- **تحيّز المجاملة**: الميل إلى موافقة من يطرح السؤال. إذا طلبت «أكّد أن مشروعنا ناجح»، فالأرجح أن المساعد سيؤكد ذلك.

لرصدها: اطلب الموقف المعاكس («ما الحجج ضد هذا الاستنتاج؟»)، واجعل النصوص الموجّهة للمجتمعات المحلية يراجعها أشخاص من تلك المجتمعات، وقارن بالمصادر المحلية.

القرارات المتعلقة بالأشخاص: الحدّ الذي لا يُتجاوز

القرارات التي تمسّ مباشرةً حقوق الشخص أو وصوله إلى المساعدة لا تُسند إلى مساعد ذكاء اصطناعي توليدي:

- اختيار المستفيدين أو استبعادهم؛
- تقييم مصداقية رواية (طلب لجوء، بلاغ، حالة حماية)؛
- الفرز المسبق للمرشحين لوظيفة أو رفضهم؛
- القرارات التأديبية أو قرارات التحقيق.

يمكن لهذه الأدوات أن تساعد في **الإعداد** (كتابة دليل مقابلة، وهيكلة معايير منشورة، وتلخيص سياسة)، لكن القرار يعود إلى أشخاص مخوّلين، وفق إجراء موثّق، مع إمكانية الطعن. وفي الاتحاد الأوروبي، تصنّف لائحة الذكاء الاصطناعي (AI Act، المعتمدة في 2024 والمطبّقة تدريجيًا) فضلًا عن ذلك بعض الاستخدامات مثل التوظيف أو الوصول إلى الخدمات الأساسية بأنها «عالية المخاطر»، مع التزامات مشدّدة.

> اختبار: لو علم الشخص المعني أن القرار اتخذه ذكاء اصطناعي، هل سيكون ذلك مقبولًا وهل يمكنه الطعن فيه؟ إذا كان الجواب لا، فلا تستخدم الذكاء الاصطناعي لاتخاذ القرار.

الشفافية والمسؤولية

الثقة هي رأس مال المنظمة الإنسانية. وبعض القواعد البسيطة تحميها:

- **المسؤولية البشرية**: لكل مستند مؤلف بشري مسؤول عن محتواه، سواء ساعده الذكاء الاصطناعي أم لا.
- **شفافية متناسبة**: أشر إلى استخدام الذكاء الاصطناعي حين يكون ملموسًا أو حين يتوقعه المتلقي بشكل معقول (تقييم، بحث، محتوى منشور، تبادل مع المجتمعات المحلية). أما لتدقيق إملائي فلا داعي لذلك.
- **احترام قواعد المانحين والشركاء**: يؤطّر بعض المانحين وبعض الدعوات لتقديم مقترحات استخدام الذكاء الاصطناعي في الطلبات أو يطلبون التصريح به؛ اقرأ الشروط (الأسبوع 6).
- **لا انتحال**: لا تولّد أبدًا صورة أو صوتًا أو اقتباسًا يُنسب زورًا إلى شخص حقيقي، ولا تقدّم أبدًا صورة مولَّدة على أنها صورة من الميدان.

تنطبق المبادئ التوجيهية الأخلاقية للتواصل في القطاع (الكرامة، والموافقة المستنيرة، وعدم استجداء الشفقة) تمامًا على المحتويات المنتَجة بالذكاء الاصطناعي.

حقوق المؤلف والانتحال والنزاهة

ثلاث نقاط تستحق الانتباه:

- **محتوى الآخرين**: أن تطلب من الذكاء الاصطناعي «إعادة كتابة» تقرير منظمة أخرى لتقديمه على أنه لك يبقى انتحالًا. اذكر مصادرك.
- **المحتوى المولَّد**: يختلف الوضع القانوني للنصوص والصور المولَّدة من بلد إلى آخر ويتطور. وتحدد شروط استخدام Claude وGemini حقوق المستخدم على النتائج؛ للاستخدام التجاري أو النشر المهم، اطّلع عليها.
- **النزاهة الأكاديمية والمهنية**: في امتحانات هذا الميني MBA، كما في اختبار توظيف، فإن استخدام الذكاء الاصطناعي للإجابة بدلًا منك غشّ. والهدف أن تتقن الكفاءات، لا الأداة وحدها.

بالنسبة للصور، تبقى قاعدة المنظمة سارية: لا صور للآخرين بلا حقوق، ولا صورة لطفل يمكن التعرف عليه بلا موافقة، والصورة المولَّدة يُشار إليها بوضوح على أنها كذلك.

التكلفة البيئية وعدالة الوصول

يستهلك الذكاء الاصطناعي الكهرباء والماء في مراكز البيانات. وتختلف التقديرات لكل طلب كثيرًا حسب الدراسات والنماذج؛ والأثر الإجمالي للقطاع يتزايد بقوة. لمنظمة ملتزمة بقضايا المناخ، عادات معقولة:

- استخدام الذكاء الاصطناعي حين يضيف قيمة حقيقية، لا بشكل تلقائي لمهام تافهة؛
- تفضيل مستند معدّ جيدًا على عشرات المحاولات المتتالية؛
- إدراج استخدام الأدوات الرقمية في السياسة البيئية.

عدالة الوصول: كثيرًا ما تعاني الفرق الميدانية من اتصال ضعيف وأجهزة قديمة ووصول أقل إلى الاشتراكات المدفوعة. والمنظمة التي تقصر الذكاء الاصطناعي على المقر والموظفين الدوليين توسّع الفجوات. والسياسة العادلة تكفل الوصول والتدريب واللغات للموظفين الوطنيين والشركاء المحليين.

تمرين: تحليل ثلاثة استخدامات من منظور أخلاقي

الحالة. تدرس منظمة غير حكومية ثلاثة استخدامات: أ) توليد صور «واقعية» لأطفال يعانون سوء التغذية لحملة جمع تبرعات، لعدم توفر صور مأذون بها؛ ب) أن تطلب من الذكاء الاصطناعي ترتيب 600 طلب لوظيفة سائق ورفض الـ 500 الأدنى تقييمًا تلقائيًا؛ ج) استخدام الذكاء الاصطناعي لاقتراح نسخة أولى من رسائل التوعية بالكوليرا بلغة فرنسية مبسّطة، تُختبر بعد ذلك مع مجموعات من المجتمع المحلي ويترجمها متحدثون أصليون.

المطلوب. لكل استخدام، قل هل هو مقبول ولماذا، واقترح بديلًا عند الحاجة.

التصحيح. أ) غير مقبول: الصورة المولَّدة المقدَّمة على أنها حقيقية تخدع الجمهور، وتمسّ الكرامة، وتعزّز استجداء الشفقة؛ البديل: صور بموافقة مستنيرة، أو رسوم توضيحية يُشار إليها بوضوح، أو رسائل تركّز على الحلول. ب) غير مقبول: قرار بشأن أشخاص من دون رقابة بشرية، مع خطر التحيّز والتمييز وغياب الطعن؛ البديل: معايير منشورة تطبّقها لجنة، ويمكن للذكاء الاصطناعي المساعدة في كتابة الشبكة أو اختبار عملي. ج) مقبول: الذكاء الاصطناعي يُعدّ مسودة أولى، ويصادق عليها المجتمع المحلي والمتحدثون الأصليون؛ إنه استخدام متناسب وشفاف ومتحقَّق منه.

للتذكر

- تنطبق المبادئ الإنسانية ومبدأ عدم الإضرار على استخدام الذكاء الاصطناعي.
- تطال التحيّزات اللغة والمنظور والصور النمطية والمجاملة؛ ابحث عن الرأي المخالف.
- لا يُسند أي قرار يمسّ حقوق الشخص أو وصوله إلى المساعدة إلى ذكاء اصطناعي توليدي.
- يبقى مؤلف بشري مسؤولًا؛ أشر إلى الاستخدام الملموس للذكاء الاصطناعي؛ ولا صورة مولَّدة تُقدَّم على أنها حقيقية أبدًا.
- فكّر في التكلفة البيئية وعدالة الوصول للموظفين الوطنيين والشركاء.

الاختبار

1. أن تطلب من الذكاء الاصطناعي «أكّد أن مشروعنا ناجح» يعرّضك خصوصًا لتحيّز:
   A. لغوي
   B. المجاملة
   C. التاريخ
   D. الصيغة

2. أيّ من هذه المهام لا يجوز إسنادها إلى ذكاء اصطناعي توليدي؟
   A. كتابة دليل مقابلة
   B. تلخيص سياسة موارد بشرية منشورة
   C. تقرير أي الأسر تُستبعد من المساعدة
   D. تصحيح إملاء تقرير

3. يمكن استخدام صورة مولَّدة لوضع ميداني:
   A. كصورة حقيقية إذا كانت واقعية
   B. أبدًا، في أي حال
   C. للمانحين فقط
   D. إذا أُشير إليها بوضوح على أنها مولَّدة واحترمت الكرامة

4. يدعو مبدأ الاستقلال إلى:
   A. الحفاظ على القدرة على العمل من دون الأداة في الوظائف الحيوية
   B. استخدام مزوّد واحد
   C. عدم استخدام الذكاء الاصطناعي أبدًا
   D. ترك الذكاء الاصطناعي يقرر

5. الإجابة عن امتحانات هذا الميني MBA بالذكاء الاصطناعي بدلًا منك هي:
   A. أمر مستحسن
   B. غشّ
   C. أمر إلزامي
   D. أمر لا يهم

أجوبة الاختبار : 1. B · 2. C · 3. D · 4. A · 5. B