جامعة البلقاء التطبيقية
Al-Balqa Applied University
حوضي من عدن الى عمان البلقاء

كلية السلط التقنية – قسم
الهندسة والذكاء الإصطناعي

قضايا أخلاقية وتطبيقات عملية في الذكاء الاصطناعي

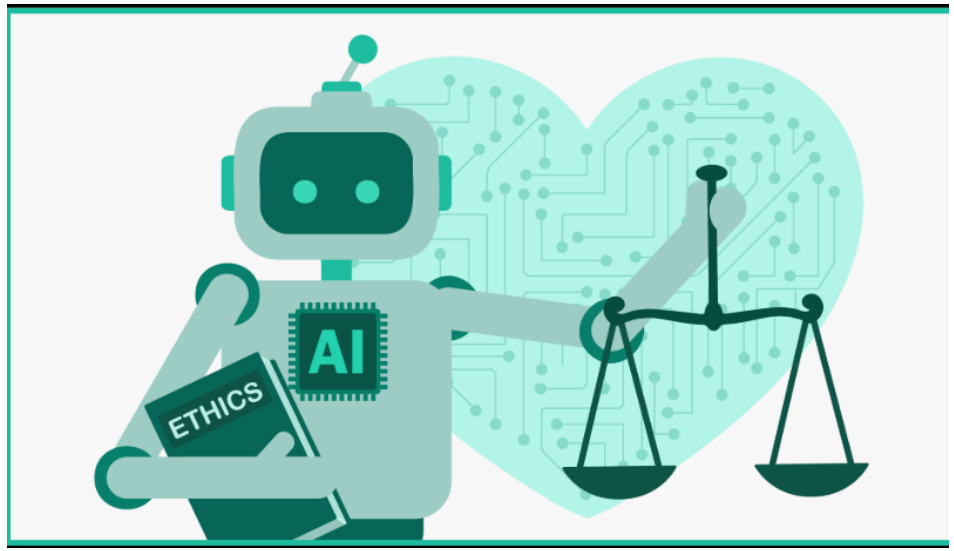
الموضوعات

- أخلاقيات الذكاء الاصطناعي والذكاء الاصطناعي المسؤول
- الذكاء الاصطناعي في الأعمال والمجتمع
- التفاعل بين الذكاء الاصطناعي والانسان
- أمان وسلامة الذكاء الإصطناعي

أخلاقيات الذكاء الإصطناعي والذكاء الإصطناعي المسؤول

أخلاقيات الذكاء الاصطناعي والذكاء الاصطناعي المسؤول

- يشهد العالم توسّعًا كبيرًا في استخدام أنظمة الذكاء الاصطناعي في القطاعات المختلفة: الأعمال، الصحة، التعليم، الأمن، المدن الذكية وغيرها. هذا التوسع يجلب **فرصًا كبيرة** لتحسين الكفاءة وجودة الحياة، لكنه في الوقت نفسه يطرح **قضايا أخلاقية وقانونية ومجتمعية**، مثل:



- من يضمن عدم استغلال البيانات الشخصية؟
- ماذا لو كان النموذج متحيّزًا ضد فئة معيّنة؟
- من يتحمّل المسؤولية عند الخطأ:
- المطوّر أم الشركة أم النظام نفسه؟

أخلاقيات الذكاء الاصطناعي والذكاء الاصطناعي المسؤول

- لهذا بدأت منظمات دولية مثل **اليونسكو (UNESCO)** و**منظمة التعاون الاقتصادي والتنمية (OECD)** و**المعهد الوطني للمعايير والتكنولوجيا (NIST)** بوضع أطر ومعايير لأخلاقيات الذكاء الاصطناعي وإدارة مخاطره.

مفاهيم أساسية

١. أخلاقيات الذكاء الاصطناعي (AI Ethics)

مجال يُعنى بالمبادئ والقيم التي يجب أن تلتزم بها أنظمة الذكاء الاصطناعي، بحيث تحترم حقوق الإنسان، والعدالة، والشفافية، والخصوصية، وتقلّل الأضرار على الأفراد والمجتمع.

٢. الذكاء الاصطناعي المسؤول (Responsible AI)

هو تطبيق عملي لأخلاقيات الذكاء الاصطناعي داخل المؤسسات: سياسات، إجراءات، تقييم مخاطر، واختبارات مستمرة لضمان أن أنظمة الذكاء الاصطناعي آمنة، عادلة، وموثوقة طوال دورة حياتها.

مفاهيم أساسية

٣. أُطر عالمية مرجعية (Global Frameworks)

- توصية اليونسكو لأخلاقيات الذكاء الاصطناعي التي تضع قيَمًا.
  - مثل: احترام حقوق الإنسان، الاستدامة البيئية، الإدماج، والعدالة
- مبادئ منظمة التعاون الاقتصادي والتنمية: التي تشجّع على ذكاء اصطناعي مبتكر وموثوق يحترم حقوق الإنسان والقيم الديمقراطية
- إطار إدارة مخاطر الذكاء الاصطناعي من NIST (NIST AI RMF 1.0) الذي يقدّم خطوات منهجية لتحديد مخاطر الذكاء الاصطناعي وقياسها وإدارتها.

لماذا نحتاج إلى أخلاقيات في الذكاء الاصطناعي؟

- تبنّت منظمات دولية مثل **UNESCO** و **OECD** و **NIST** أطرًا ومعايير لأخلاقيات ومخاطر الذكاء الاصطناعي بسبب:
  - ✓ انتشار أنظمة الذكاء الاصطناعي في: الأعمال، الصحة، التعليم، المدن الذكية، الأمن.
  - ✓ وجود مخاطر حقيقية مثل:التحيّز ، انتهاك الخصوصية، قرارات غير مفهومة، واستغلال البيانات.
  - ✓ قرارات هذه الأنظمة قد تؤثّر على: التوظيف، القروض البنكية، الرعاية الصحية، فرص التعليم.

المبادئ الأخلاقية الرئيسة

١. **الخصوصية والمراقبة (Privacy & Surveillance)**

- تتعلّق بحماية بيانات الأفراد، وحقهم في عدم التتبّع والمراقبة الدائمة.
- في سياق الذكاء الاصطناعي، تُجمع كميات كبيرة من البيانات (Big Data) من الهواتف، الكاميرات، وسائل التواصل، إنترنت الأشياء مما يخلق مخاطر:
  ١. التتبّع الدائم لمواقع الأشخاص وسلوكهم.
  ٢. بيع البيانات لأطراف أخرى دون موافقة صريحة.
  ٣. ربط بيانات من مصادر مختلفة للوصول لصورة شديدة التفصيل عن حياة الفرد.
- الأطر الدولية مثل اليونسكو و OECD تؤكد على ضرورة **تقليل البيانات** والحصول على موافقة واضحة، وتطبيق تشفير وإخفاء هوية حيث أمكن.

المبادئ الأخلاقية الرئيسة

➢ **مثال تطبيقي:**

تطبيق صحي يجمع بيانات نبض القلب والنوم من الساعة الذكية.

– الاستخدام الإيجابي: تحسين الصحة وتقديم تنبيهات مبكّرة.

– التحدّي الأخلاقي: هل يمكن للشركة أن تبيع هذه البيانات لشركات تأمين لتحديد أسعار أعلى لبعض المستخدمين؟

المبادئ الأخلاقية الرئيسة

٢. العدالة والإنصاف (Bias & Fairness)

- العدالة تعني ألّا يميّز النظام بين الأفراد أو المجموعات لأسباب غير مبرّرة (الجنس، العِرق، الدين...)
- التحيّز قد ينتج من بيانات تاريخية منحازة أو من طريقة تصميم النموذج.

➢ أطر OECD و UNESCO تضع **العدالة وعدم التمييز** ضمن القيم الأساسية للذكاء الاصطناعي

مثال:

➢ نظام توظيف آلي تم تدريبه على بيانات سابقة من شركة كانت تفضّل الرجال على النساء في الوظائف التقنية.

➢ فينعكس ذلك في النموذج الذي يعطي نقاطًا أقل للمتقدّمات.

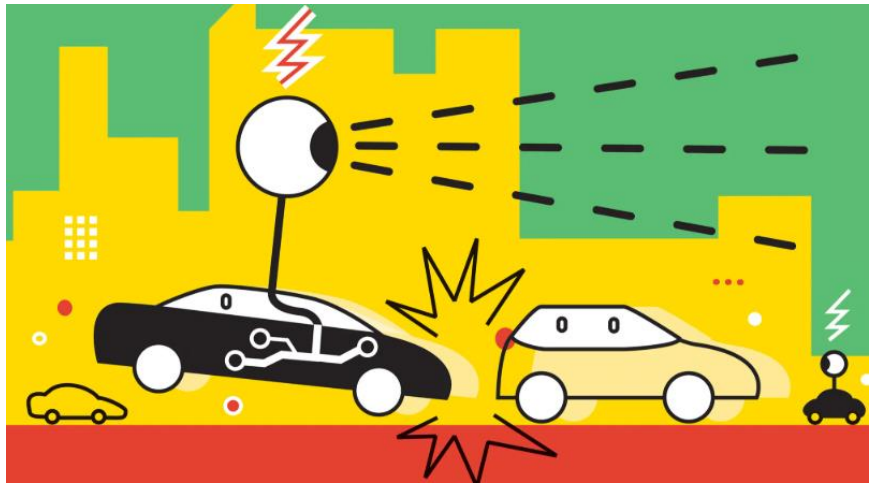
المبادئ الأخلاقية الرئيسة

٣. المُسائلة والمحاسبة (Accountability)

- وجود أنظمة ذكية لا يلغي مسؤولية البشر والمؤسسات، ويطرح أسئلة مثل: من المسؤول قانونيًا عند حدوث ضرر؟.
- الأطر الدولية تشدّد على: وجود شخص/جهة واضحة مسؤولة عن القرارات المدعومة بالذكاء الاصطناعي، وضرورة توثيق قرارات التصميم والاختبارات لاستخدامها عند التحقيق أو التقاضي.

مثال:

- سيارة ذاتية القيادة تسبّبت في حادث: لا يمكن القول "النظام أخطأ" فقط؛ يجب تحديد مسؤولية الشركة المصنّعة، والمبرمجين، وربما السائق المشرف.



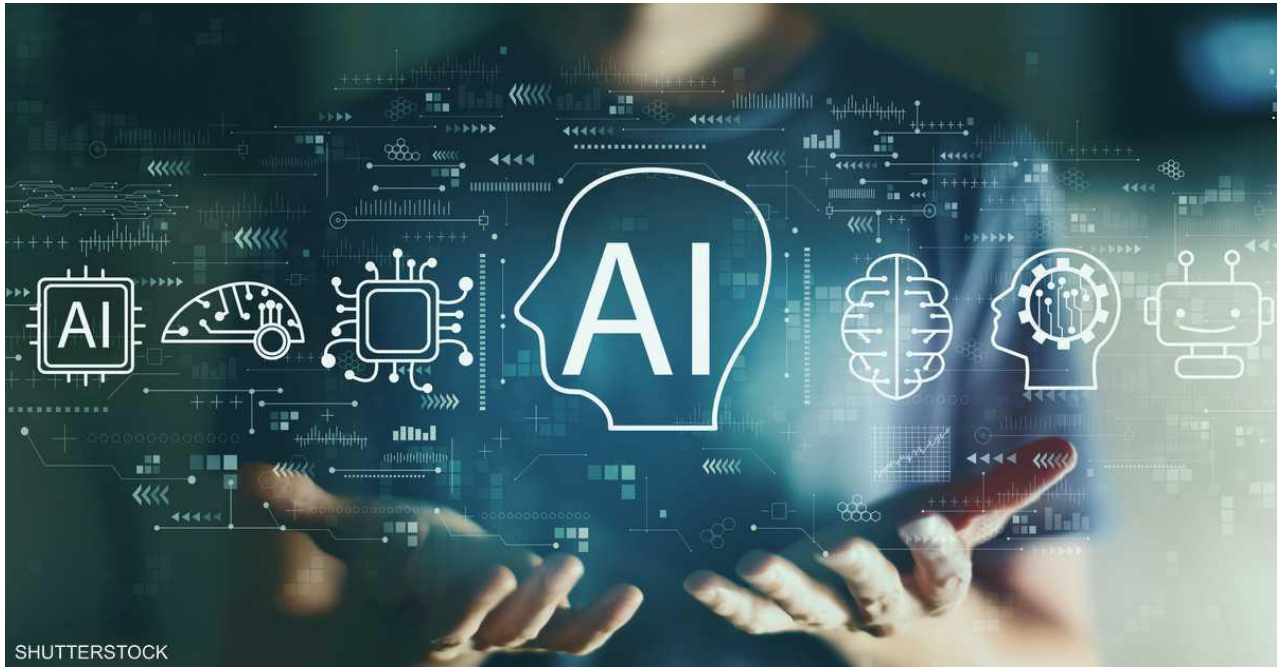
المبادئ الأخلاقية الرئيسة

٤. السلامة والموثوقية (Safety & Reliability)

- السلامة: ألا يسبّب النظام أذى غير مقبول للأشخاص أو البيئة.
- الموثوقية: أن يعمل النظام وفق التوقّعات في ظروف مختلفة، مع قدرته على التعامل مع الأخطاء والضجيج في البيانات.

مثال:

نظام تشخيص طبي يعتمد على صور الأشعة يجب أن يُختبر على مجموعات بيانات من مستشفيات متعددة ومن فئات سكانية مختلفة قبل استخدامه في قرارات علاجية.



الذكاء الاصطناعي في الأعمال والمجتمع

الذكاء الاصطناعي في الأعمال والمجتمع

أمثلة لتطبيقات عملية للذكاء الاصطناعي في الأعمال والمجتمع:

١. الذكاء الاصطناعي في الأعمال (AI in Business)

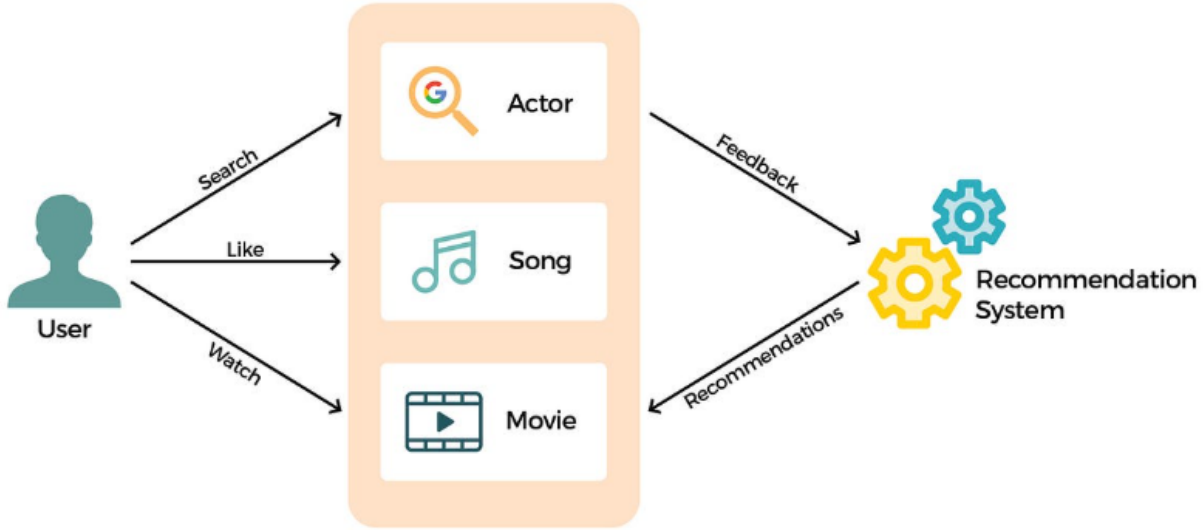
A. خدمة العملاء – المحادثات الآلية (Chatbots)

- تستخدم البنوك وشركات الاتصالات والمتاجر الإلكترونية روبوتات محادثة للإجابة عن الأسئلة المتكررة (FAQs) وحجز المواعيد وتتبّع الشحنات.
- تعطي هذه الأنظمة ردودًا فورية على مدار الساعة، وتقلّل الضغط على الموظفين.

الذكاء الاصطناعي في الأعمال والمجتمع

B. أنظمة التوصية (Recommendation Systems)

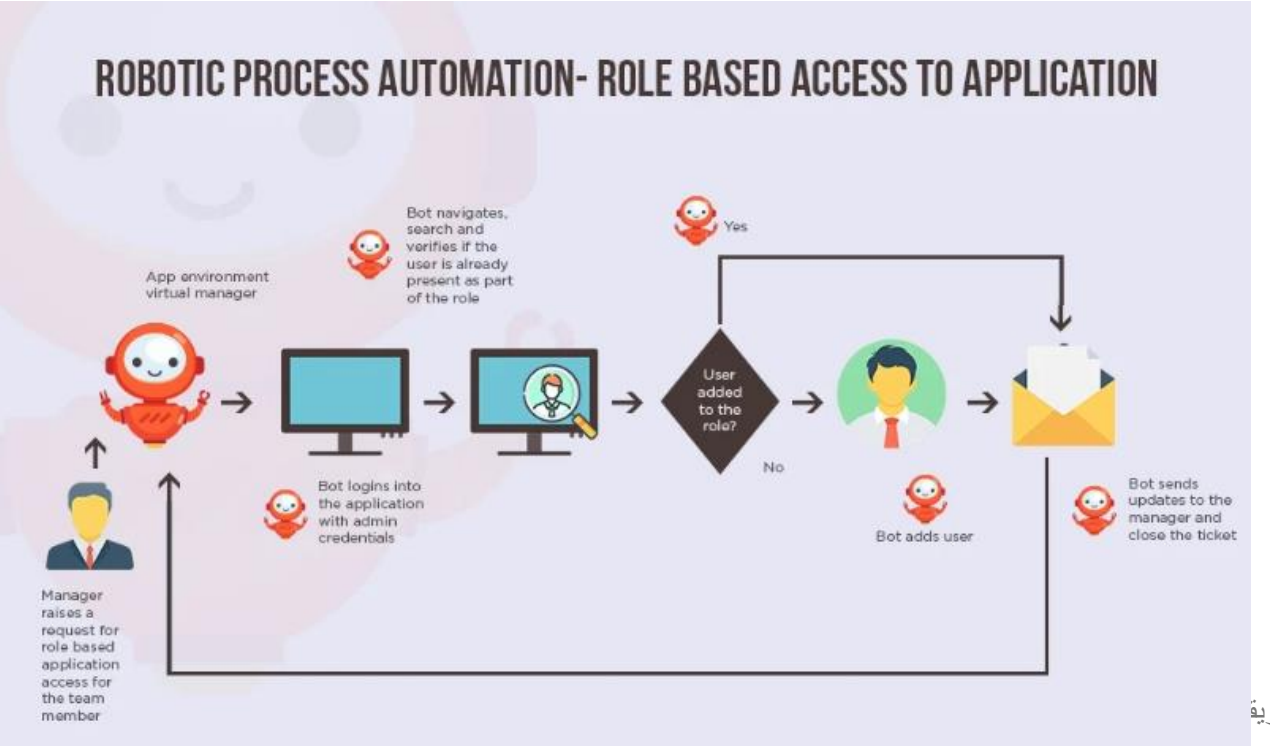
- مثال: مواقع مثل Amazon و Netflix تحلل سلوك المستخدم (ما يشاهده، ما يشتريه، ما يقيمه) لتقترح منتجات أو أفلامًا مناسبة.
- البُعد الأخلاقي:
  - يمكن أن تدفع المستخدم لاستهلاك أكثر مما يحتاج.
  - قد تحصره في "فقاعة" من المحتوى المتشابه (Filter Bubble)



الذكاء الاصطناعي في الأعمال والمجتمع

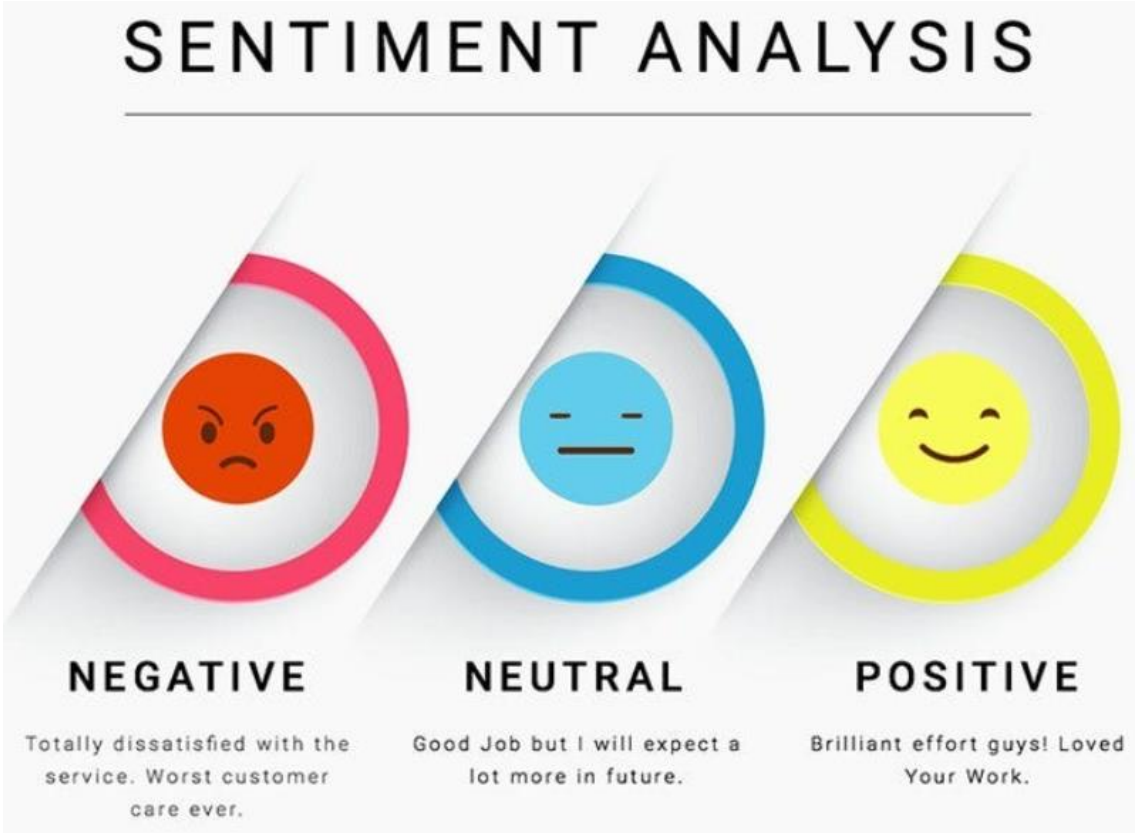
C. الأتمتة والروبوتات البرمجية (RPA & Automation)

- استخدام الروبوتات البرمجية لتنفيذ مهام مكتبية متكررة: إدخال بيانات، توليد تقارير، إرسال رسائل.
- هذا يزيد الكفاءة لكنه قد يستبدل بعض الوظائف الروتينية، ويحتاج إلى إعادة تأهيل الموظفين.



مثال على استخدام الروبوتات البرمجية (RPA) لأتمتة عملية منح الصلاحيات داخل الأنظمة، حيث يقوم الروبوت البرمجي بتنفيذ خطوات متكررة وفق قواعد محددة، مما يقلّل الأخطاء البشرية ويزيد الكفاءة، مع الالتزام بسياسات الأمان والمسؤولية.

الذكاء الاصطناعي في الأعمال والمجتمع



D. تحليل المشاعر والآراء (Sentiment Analysis)

- تحليل تعليقات العملاء في شبكات التواصل لاستخراج انطباعهم عن منتج أو حملة تسويقية.
- مثال:
  - شركة تقرأ آلاف التغريدات لمعرفة هل الحملة الجديدة "إيجابية" أم "سلبية"، وتعدّل استراتيجيتها بناء على ذلك.

الذكاء الاصطناعي في الأعمال والمجتمع

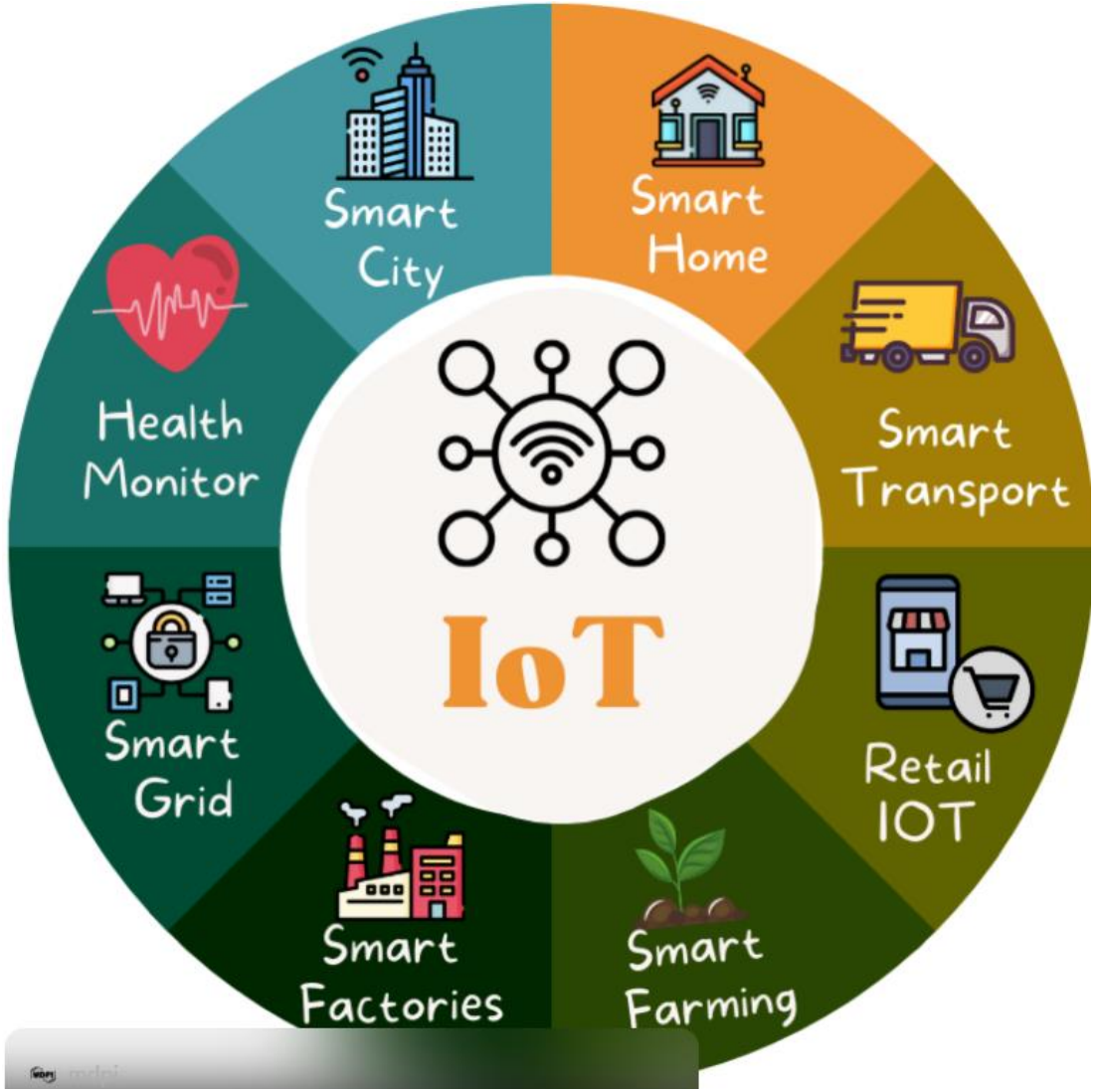
٢. الذكاء الاصطناعي في المجتمع (AI in Society)

A. القطاع الصحي

- أنظمة تشخيص تعتمد على صور MRI وCT.
- أنظمة توقّع انتشار الأمراض أو اكتظاظ أقسام الطوارئ.

✓ هذه التطبيقات يمكن أن تحسّن جودة الرعاية لكنها تثير أسئلة عن الخصوصية ودقّة النماذج

الذكاء الاصطناعي في الأعمال والمجتمع



B. المدن الذكية (Smart Cities)

- استخدام حسّاسات وإنترنت الأشياء (IoT)لإدارة إشارات المرور، استهلاك الطاقة، مراقبة التلوّث.

- مثال:
  - نظام يضبط أضواء الشوارع بناءً على وجود المشاة أو السيارات لتقليل استهلاك الكهرباء.

الذكاء الاصطناعي في الأعمال والمجتمع

C. الأثر المجتمعي والتوظيف (Societal Impact & Employment)

- تأثير الذكاء الاصطناعي على سوق العمل، مثل:

✓ **استبدال بعض الوظائف الروتينية (مُدخل البيانات، خدمة العملاء)**

✓ **خلق وظائف جديدة** في تطوير النماذج، إدارة البيانات، تحليل البيانات والنتائج.

✓ **الحاجة لإعادة تأهيل:** تدريب العاملين لاكتساب مهارات رقمية جديدة.

➢ OECD تشير في تقاريرها إلى أن الذكاء الاصطناعي قد لا يزيل عددًا هائلًا من الوظائف دفعة واحدة، لكنه سيغيّر محتوى الوظيفة والمهارات المطلوبة بشكل كبير.

التفاعل بين الذكاء الاصطناعي والإنسان
(Human–AI Interaction)

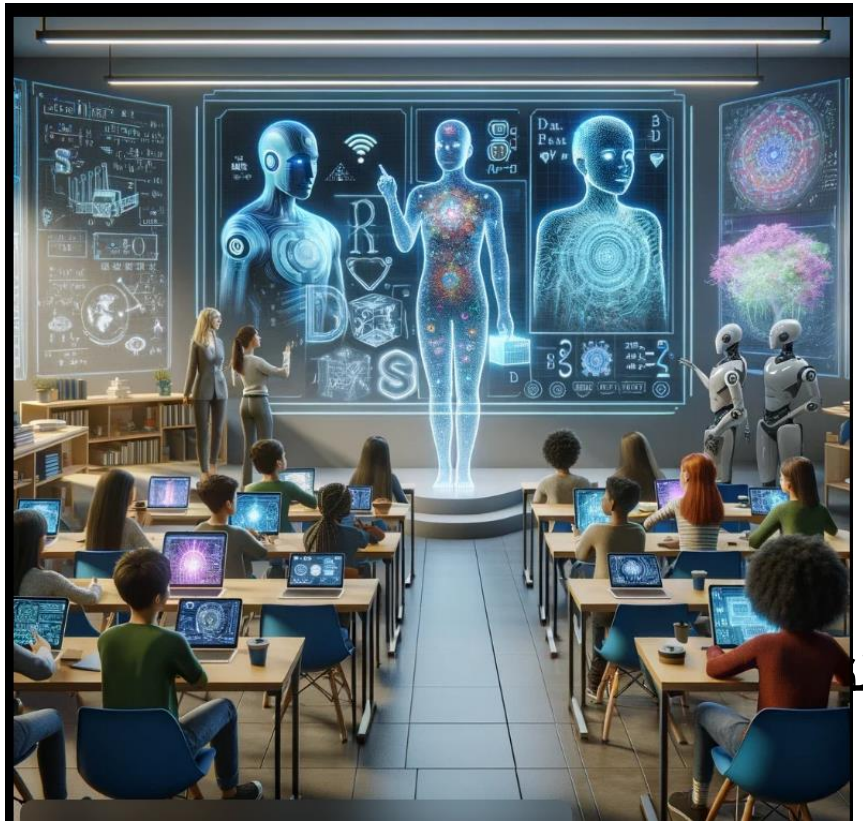
- **ما هو التفاعل بين الذكاء اصطناعي والإنسان؟**

هو مجال يدرس كيف يتعاون الإنسان مع أنظمة الذكاء الاصطناعي: كيف يفهم الإنسان مخرجات النظام، وكيف يشرح النظام نفسه للمستخدم، وكيف يتشاركان المسؤولية.

- **تكامل القدرات بين الإنسان والآلة**

مقارنة نقاط قوة الانسان (مثل الفهم السياقي، التعاطف ...) ونقاط قوة الذكاء الاصطناعي (مثل سرعة المعالج، تحليل كميات ضخمة من البيانات ...)

الإنسان .VS الذكاء الاصطناعي



➢ **نقاط قوة الإنسان:**

- فهم السياق والخلفية الثقافية
- التعاطف والتواصل
- الإبداع والحكم الأخلاقي

➢ **نقاط قوة الذكاء الإصطناعي:**

- السرعة والحساب (Computation)
- التعامل مع بيانات ضخمة (Big Data)
- التعرّف على الأنماط الدقيقة التي لا يستطيع الانسان ملاحظ
  بسهولة (Pattern Recognition)

➢ الهدف ليس أن يستبدل أحدهما الآخر، بل أن **يتكاملا**.

أدوار الذكاء الاصطناعي بالنسبة للإنسان

- فيما يلي أدوارًا يمكن أن يلعبها الذكاء الاصطناعي في العلاقة مع الإنسان:

A. **أداة (Tool):** مثل محرر صور يستخدم الذكاء الاصطناعي لتحسين الصورة، لكن المستخدم يقرر متى وكيف يطبّق التعديلات.

B. **مساعد (Assistant) :**مثل مساعد صوتي مثل Google Assistant يقترح تذكيرًا أو يضبط المنبّه، لكن القرار النهائي للمستخدم.

C. **شريك - متعاون (Partner/Collaborator):**مثل نظام تحليل بيانات يعمل مع فريق تسويق؛ الفريق يطرح الأسئلة والنظام يقدم تحليلات تُستخدم في اتخاذ القرار.

D. **مشرف أو مراقب (Overseer) :**مثل نظام يراقب خطوط الإنتاج ويطلق إنذارًا عند اقتراب حدوث عطل، ليتمكّن الإنسان من التدخل مبكرًا.

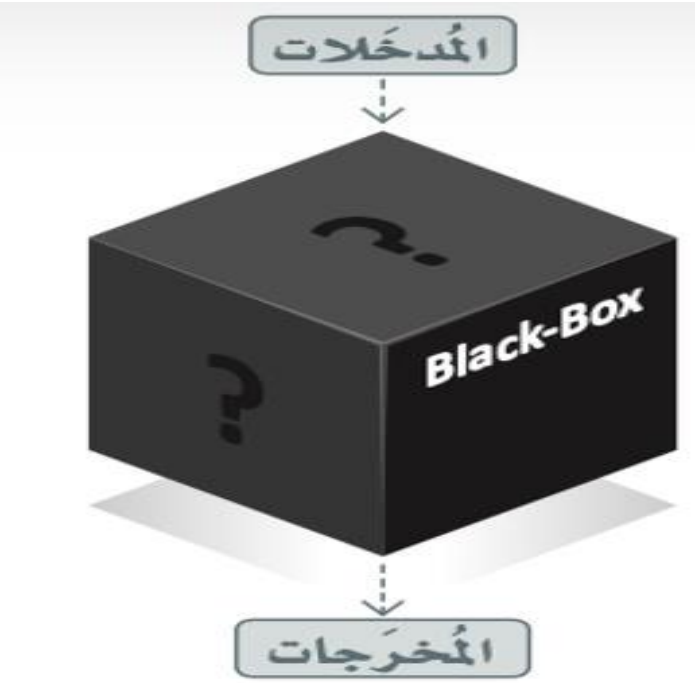
التحديات والمخاطر في التفاعل

- فيما يلي عدداً من التحديات:

١. الإفراط في الاعتماد (Over-reliance)

٢. تحيّز الأتمتة (Automation Bias)

يميل بعض المستخدمين لتصديق النظام الآلي حتى عندما يخطئ، لمجرد أنه "ذكي".



٣. صندوق أسود (Black Box) :

بعض النماذج – خصوصًا العميقة – صعب تفسيرها، مما يقلّل الثقة ويصعّب المساءلة.

التحديات والمخاطر في التفاعل

٤. قضايا الخصوصية والأمان:

خاصة مع الأنظمة التي تجمع بيانات حساسة أو تسجّل المحادثات والتفاعلات.

٥. التأثير على المهارات البشرية:

إذا اعتمدنا على الأنظمة في كل شيء، قد تضعف بعض المهارات لدى البشر (الحساب، الكتابة، البحث...).

٦. التسريح من الوظائف

هناك مخاوِف بشأن تسريح البشر من وظائفهم وتأثيره على سبل عيش العاملين، فعلى الرغم من أن الذكاء الاصطناعي يمكِنه أن يؤدي إلى تحسين الكفاءة والإنتاجية، إلا أنه يمكِن أن يؤدي أيضًا إلى فقدان البشر لوظائفهم وتزايد عدم المساواة في الدخل؛ مما قد يكون له عواقب اجتماعية واقتصادية سلبية.

إرشادات عملية لتصميم تفاعل مسؤول

- استنادًا إلى "إرشادات التفاعل بين الإنسان وذكاء اصطناعي" من مايكروسوفت:
  - توضيح ما الذي يمكن للنظام فعله، وما حدوده، منذ البداية.
  - توضيح أن النظام قد يخطئ، وتشجيع المستخدم على المراجعة.
  - جعل التفسيرات بسيطة ومرتبطة بسياق المستخدم.
  - توفير واجهة اعتراض أو تصحيح للمخرجات (Feedback)

الإرشادات العالمية لأخلاقيات الذكاء الاصطناعي
(International AI Ethics Guidelines)

الأمن السيبراني
أمان وسلامة الذكاء الإصطناعي

أمن وسلامة الذكاء الاصطناعي

- يُقصد بأمن وسلامة الذكاء الاصطناعي مجموعة الإجراءات التي تهدف إلى ضمان عمل أنظمة الذكاء الاصطناعي بشكل آمن وموثوق، ومنع إساءة استخدامها أو التلاعب بنتائجها.

- تعتمد أنظمة الذكاء الاصطناعي على كميات كبيرة من البيانات، مما يجعلها عرضة لمخاطر تتعلق بخصوصية البيانات، مثل جمع البيانات دون علم المستخدم أو تسريبها.

- قد تتعرض نماذج الذكاء الاصطناعي لهجمات تؤثر على دقة مخرجاتها، مثل تزويد النظام ببيانات خاطئة أو مضللة تؤدي إلى قرارات غير صحيحة.

- من مخاطر الذكاء الاصطناعي أيضًا التحيّز في النتائج، أو حصر المستخدم داخل فقاعة من المحتوى المتشابه (Filter Bubble)، مما يقلل من تنوع المعلومات ووجهات النظر.

أمن وسلامة الذكاء الاصطناعي

➢ أمن الذكاء الاصطناعي

يتعلق أمن الذكاء الاصطناعي بالمخاطر المحتملة من الاختراق أو إساءة الاستخدام، مثل:الهجمات السيبرانية المتقدمة، التي تشمل:

١. هجمات التصيد الاحتيالي
٢. وسرقة البيانات
٣. التلاعب بالأنظمة الذكية
٤. التدخل في عمليات اتخاذ القرار.

أمن الذكاء الاصطناعي

لتحقيق أمن الذكاء الإصطناعي يُعتمد على مفهوم **ثالوث الاستخبارات المركزية (CIA Triad)** وهو إطار أساسي في **الأمن السيبراني**، ويُستخدم لضمان حماية المعلومات والأنظمة.

يتكوّن من ثلاثة عناصر رئيسية:

١. **السرّية (Confidentiality)**

ضمان أن المعلومات لا يطّلع عليها إلا الأشخاص المصرّح لهم.

مثال: تشفير بيانات المستخدمين بحيث لا يستطيع قراءتها أي شخص غير مخوّل.

٢. **السلامة / التكاملية (Integrity)**

ضمان عدم تعديل البيانات أو التلاعب بها دون إذن.

مثال: منع تغيير نتائج نظام ذكاء اصطناعي أو بيانات التدريب بشكل غير مصرح به.

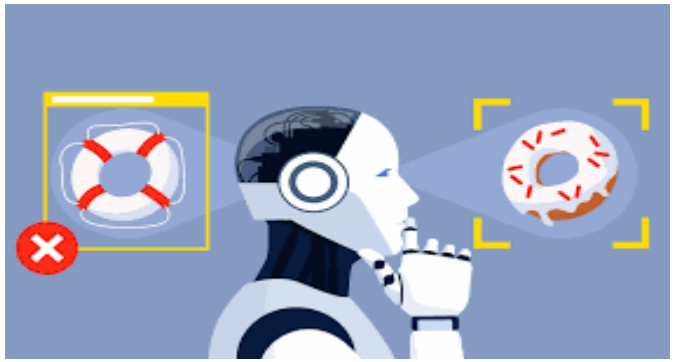
٣. **التوافرية (Availability)**

ضمان بقاء الأنظمة والبيانات متاحة للمستخدمين المصرّح لهم عند الحاجة.

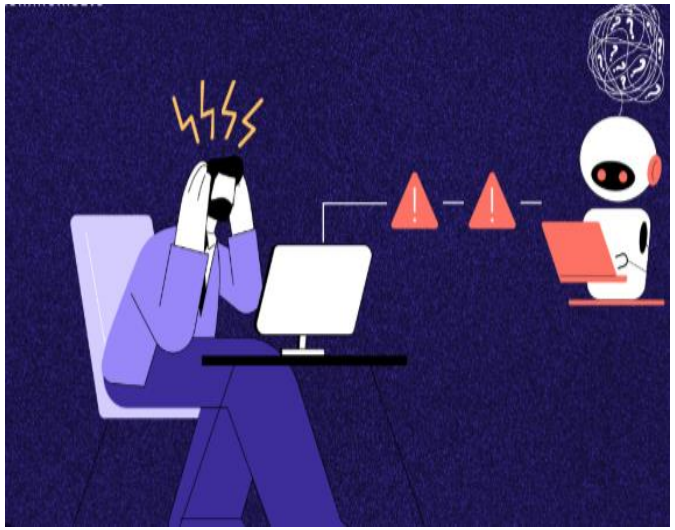
مثال: حماية نظام ذكي من هجمات تعطيل الخدمة حتى لا يتوقف عن العمل.

سلامة الذكاء الاصطناعي

- سلامة الذكاء الاصطناعي تُشير إلى مجموعة المبادئ والإجراءات التي تهدف إلى ضمان أن تعمل أنظمة الذكاء الاصطناعي **بشكل صحيح وآمن**، وأن تكون نتائجها **موثوقة** ولا تُسبب ضررًا للأفراد أو المجتمع.



- تركّز سلامة الذكاء الاصطناعي على تقليل المخاطر الناتجة عن:
  - الأخطاء البرمجية.
  - القرارات غير المتوقعة للأنظمة الذكية.
  - إساءة الاستخدام أو الاعتماد المفرط على الأنظمة الذكية.



الفرق بين السلامة والأمن

- **السلامة (Safety) :** منع الضرر الناتج عن أخطاء أو سلوك غير متوقّع للنظام. مثلا: منع السيارة ذاتية القيادة من اتخاذ قرار خاطىء يسبب حادثا
- **الأمن (Security):** حماية النظام من الهجمات والاختراقات الخارجية.
  مثلا: منع اختراق السيارة ذاتية القيادة.

سلامة الذكاء الاصطناعي

أهداف سلامة الذكاء الاصطناعي

- حماية الإنسان وحقوقه الأساسية.
- ضمان دقة مخرجات الأنظمة الذكية.
- تقليل الأضرار المحتملة الناتجة عن استخدام الذكاء الاصطناعي.
- تعزيز الثقة في تطبيقات الذكاء الاصطناعي.

مجالات سلامة الذكاء الاصطناعي

- **سلامة البيانات:** التأكد من جودة البيانات وخلوّها من الأخطاء أو التحيّز.
- **سلامة النماذج:** ضمان أن النموذج يتصرف كما هو متوقَّع ولا يعطي نتائج مضللة.
- **سلامة الاستخدام:** منع استخدام الذكاء الاصطناعي في سياقات ضارة أو غير أخلاقية.
- **الرقابة البشرية:** بقاء الإنسان ضمن دائرة اتخاذ القرار وعدم ترك القرار للنظام وحده.

الأمن السيبراني

- الأمن السيبراني هو مجموعة التقنيات والإجراءات التي تهدف إلى حماية الأنظمة والشبكات والبيانات من الهجمات الرقمية والاختراقات.

- يُعد الأمن السيبراني عنصرًا أساسيًا في حماية تطبيقات الذكاء الاصطناعي، نظرًا لاعتمادها الكبير على البيانات والاتصال بالشبكات.

- يركز الأمن السيبراني على تحقيق ثلاث ركائز أساسية:
  سرية البيانات، سلامتها، وتوافرها.

- تتضمن التهديدات السيبرانية:
  ١. هجمات الاختراق
  ٢. سرقة البيانات
  ٣. نشر البرمجيات الخبيثة
  ٤. تعطيل الأنظمة.

الأمثلة العملية لتطبيقات الذكاء الاصطناعي في الأمن السيبراني

الكشف عن البرمجيات الضارة (Malware Detection)

- يمكن للذكاء الاصطناعي اكتشاف البرمجيات الضارة من خلال تحليل أنماط البيانات وتحديد النشاط المشبوه في الأنظمة والشبكات.
- مثال: تقوم أنظمة الكشف المبنية على الذكاء الاصطناعي بتمييز البرامج الضارة الجديدة من الملفات، أو رسائل البريد الإلكتروني، أو الروابط المشبوهة بدقة وسرعة متناهية.

كشف اختراق الشبكة (Network (Intrusion Detection

- تساعد تقنيات الذكاء الاصطناعي في اكتشاف محاولات الاختراق من خلال تحليل حركة البيانات، وتحديد الأنشطة التي قد تشير إلى وقوع هجوم.
- مثال: يتم استخدام التعلُّم العميق على الذكاء الاصطناعي في محاولات اختراق الشبكة، كما يمكنه التنبؤ بالهجمات قبل أن تتسبب في أضرار كارثية.

تحليل المعلومات الاستباقية Threat Intelligence) (Analysis

- تحليل المعلومات الاستباقية يعني ان النظام ينتبه للخطر قبل ان يحدث، وليس بعد وقوعه.
- يستخدم الذكاء الاصطناعي لتحليل كميات كبيرة من البيانات الأمنية من مصادر مختلفة، يهدف اكتشاف التهديدات الناشئة والتنبؤ بها قبل وقوع الهجمات، مما يساعد على اتخاذ اجراءات وقائية وتقليل الأضرار المحتملة.
- مثال:يمكن للذكاء الاصطناعي التعرف على وجود برمجيات ضارة تنشر بسرعة عبر الإنترنت وتوفير الإنذارات اللازمة قبل انتشارها بشكل أوسع

الأمثلة العملية لتطبيقات الذكاء الاصطناعي في الأمن السيبراني

تحليل سلوك المُستخدِم (User Behavior Analysis)

- يمكن استخدام الذكاء الاصطناعي لتحليل سلوك المستخدمين، مما يساعد في تقليل مخاطر الأمن السيبراني المحتملة.
- مثال: يمكن لأنظمة الذكاء الاصطناعي التعرف على الأنشطة المشبوهة مثل محاولات الوصول إلى البيانات الحساسة خلال ساعات عمل غير اعتيادية أو من مواقع غير معتادة.

كشف الاحتيال (Fraud Detection)

- يمكن للذكاء الاصطناعي اكتشاف الأنشطة الغير العادية التي تشير إلى الاحتيال، مثل اختراق حسابات الاتصالات أو انتحال الشخصية.
- مثال: يتم تدريب النظام القائم على الذكاء الاصطناعي ليشمل علاقة التفاعل أثناء عمليات الدفع عبر الإنترنت، ويقارن أنماط الاستخدام الاعتيادية للمستخدم مع أي احتمالات احتيال ممكنة.

كشف الاحتيال يركز على هل هذا التصرف احتيالي أم لا

تحليل سلوك المستخدم يركز على كيف يتصرف المستخدم

التكامل بين أمن الذكاء الاصطناعي والأمن السيبراني

- تطلب الاستخدام الآمن للذكاء الاصطناعي الجمع بين تقنيات الذكاء الاصطناعي الآمن وتطبيق مبادئ الأمن السيبراني.
- يشمل ذلك وضع سياسات وضوابط واضحة، ومراقبة أداء الأنظمة بشكل مستمر، وتحديثها لمواجهة التهديدات الجديدة.
- يُعد رفع الوعي الأمني لدى المستخدمين والمطورين خطوة أساسية للحد من المخاطر المرتبطة بالذكاء الاصطناعي والهجمات السيبرانية.