

الفلسفة وتعويدة الجي بي تي!

Philosophy and the Mantra of GPT!

دكتور / صلاح عثمان (أستاذ المنطق وفلسفة العلم – رئيس قسم الفلسفة – كلية الآداب
– جامعة المنوفية – جمهورية مصر العربية)

Salah Osman

(Menoufia University, Egypt)

salah.mohamed@art.menofia.edu.eg

DOI: [10.13140/RG.2.2.28579.32802](https://doi.org/10.13140/RG.2.2.28579.32802)

مقال منشور بموقع أكاديمية بالعقل نبدأ، القاهرة، بتاريخ ١٢ أبريل ٢٠٢٣
With Mind We Start, 2023, April 12.

لم نعد بحاجة إلى فانوس سحري نمسح عليه بأصابعنا لكي يخرج منه المارد القادر على خدمتنا وتلبية بعض أهم مطالبنا الحياتية، ولم نعد بحاجة إلى تعويذات نلج بها في عالم السحر والخيال؛ لقد خرج المارد بالفعل من قممه الحاسوبي؛ من جوف مختبرات البرمجة والذكاء الاصطناعي، بتعويذات (أكواد) رياضية رمزية سرعان ما تمكن من التهامها وهضمها، ليبيت قادراً على إنتاج تعويذات أخرى مماثلة، وربما أفضل منها! خرج «المحول التوليدي المدرب مسبقاً» Generative pre-trained transformer، المعروف اختصاراً باسم «جي بي تي» GPT، ملوحاً بإمكانات بحثية وخدمية وإنتاجية هائلة، ومهدداً بتصفية قطاعات بأسرها من المهن والوظائف، وبتجريف مهارات البحث العلمي لدى طلاب المدارس والجامعات!

المحولات التوليدية المدربة مسبقاً عائلة من نماذج اللغة المدربة عموماً على مجموعة كبيرة من البيانات النصية لإنشاء نص يشبه النص الإنساني، تم تطويرها من قبل مؤسسة بحثية غير ربحية تُعرف باسم «أوبن إيه آي» OpenAI، تأسست سنة ٢٠١٥ بالولايات المتحدة بهدف تعزيز وتطوير أنظمة الذكاء الاصطناعي. وقد تم بناء هذه النماذج باستخدام مجموعة من المحولات (أو الشبكات العصبية غير الترتيبية) التي يُمكن صقلها لمباشرة مهام معالجة اللغة الطبيعية (لغة الحياة اليومية)، مثل إنشاء نص وترجمته وتصنيفه. يشير «التدريب المسبق» في الاسم إلى عملية التدريب الأولية على مجموعة نصية كبيرة، حيث يتعلم النموذج التنبؤ بالكلمة التالية في المقطع مما يوفر أساساً متيناً للنموذج لأداء جيد في المهام النهائية بكميات محدودة من البيانات الخاصة بالمهمة.

النموذج الأخطر للمحولات التوليدية المُدربة مُسبقًا يُعرف باسم «تشات جي بي تي» ChatGPT، وهو برنامج للردشة تم إطلاق نسخته الأولى (GPT-3.5) في الثلاثين من نوفمبر سنة ٢٠٢٢، وسرعان ما جذب الانتباه لردوده التفصيلية وإجاباته الواضحة عبر العديد من مجالات المعرفة، وبعده قُدرت قيمة «أوبن إيه آي» بنحو ٢٩ مليار دولار أمريكي في سنة ٢٠٢٣. أما النسخة الأحدث من المُحولات (GPT-4) فقد تم إطلاقها في الرابع عشر من مارس الماضي (٢٠٢٣)، وهو نموذج لغوي ضخم متعدد الوسائط (يقبل النصوص والصور كمدخلات).

الشعور بالفزع - كما عبر عنه عددٌ من خبراء ومُطوري الذكاء الاصطناعي - يعكس مدى خطورة هذه التقنية، وهو ما يستدعي كالعادة تدخل الفلسفة، أخلاقياً على الأقل، وهو ما ناقشه في السطور التالية.

في منشور تحت عنوان «ماكينات الرد: لماذا يجب على الذكاء الاصطناعي أن يدفعنا إلى طرح أسئلة أفضل؟» Answer Machines: Why AI Should Make Us Ask Better Questions، نُشر بموقع «روفان» The Ruffian بتاريخ ٢٥ يناير ٢٠٢٣، ذهب الكاتب والمدون البريطاني «إيان ليسلي» Ian Leslie إلى أن ثمة موضوعًا واحدًا يجب أن يصبح أكثر أهمية في عصر الذكاء الاصطناعي، ألا وهو الفلسفة!

يُعلق «لوي تشو» Luyen Chou (المشتغل بتدريس الفلسفة وإنتاج وتوزيع التقنيات بجامعة كولومبيا ببروكلين - نيويورك) على منشور «إيان ليسلي» مازحًا بقوله: «بصفتي متخصصًا في الفلسفة، لم أجد بعد وظيفة مدفوعة الأجر في هذه الصناعة المذكورة، دعني أخبرك أن كلماتك قد أطربت أذني! وبعيدًا عن المزاح، كلما فكرت في كلٍ من القدرات المذهلة والقيود المهمة التي كشفت عنها تقنيات الذكاء الاصطناعي الجديدة، كلما انتفتت مع ليسلي»!

لا شك أن أحد القيود المفروضة على نماذج الذكاء الاصطناعي التوليدية مثل ChatGPT هو التحفظ في إصدار أحكام منطقية. ولعل هذا التحفظ في الحقيقة ليس تقييدًا للتكنولوجيا ذاتها بقدر ما هو نتيجة للقيود التي يفرضها المطورون البشريون، ذلك أن التركيز على البُعد الأخلاقي للعُنصر البشري من شأنه أن يهدئ المخاوف بشأن الدور الناشئ للذكاء الاصطناعي في عالمنا، ويدعم في الوقت ذاته الاعتراف بالفاعلية البشرية، وبملكيتنا جميعًا لابتكارنا التكنولوجية، بل ويدعم نموذجًا يؤدي فيه الذكاء الاصطناعي دورًا "استشاريًا" لا يزال يُلقي على عاتق البشر مسؤولية اتخاذ أي القرار النهائي.

مع ذلك، فإن ثمة شك في استمرار هذه القيود المفروضة ذاتيًا على اتخاذ القرار بشأن الذكاء الاصطناعي، فلقد خرج المارد من قممه بالفعل. وكما تعلمنا من ثوراتنا التكنولوجية السابقة، فإن الجهود المبذولة لتقييد القدرات الناشئة، مهما كانت حسنة النية، لم ولن تتسم بالدوام أبدًا، لأن

شخصًا ما في مكانٍ ما سوف يكون دومًا على استعداد للانطلاق إلى أبعد مدى تحت ضغط الميزة التنافسية، وحتى الجهات الأكثر تحفظًا سوف تجد من المستحيل مقاومة الرغبة في فعل الشيء ذاته!

خذ على سبيل المثال الجدل المثير للغاية حول استخدام الذكاء الاصطناعي المستقل في سباق التسلح بين القوى الكبرى؛ فعلى الرغم من تزايد القلق من قبل التقنيين وعلماء الأخلاق وواضعي السياسات، فقد رفضت دولٌ مثل الولايات المتحدة بحزم تقييد تقنيات الذكاء الاصطناعي، مثل الطائرات المُسيرة دون طيار، فيما يتعلق باتخاذ قرارات مستقلة في ساحة المعركة، بحجة أنهم بفعلهم هذا سيضعون دفاعاتهم الوطنية في مأزق ضد خصوم أقل دقة! ولكن إذا كان من المتوقع أن تتخذ طائرة مُسيرة دون طيار قرارًا نهائيًا بشأن ما إذا كانت ستطلق صاروخًا يمكن أن يقضي على إرهابي خطير، رغم ما قد يُسببه ذلك من إصابات في صفوف المدنيين الأبرياء، ألن يتحمل المُشغل البشري الذي يتخذ هذا القرار (برمجياً) عبء العواقب لبقية حياته؟ هل يمكن للذكاء الاصطناعي إصدار مثل هذه الأحكام بشعور متساوٍ من المسؤولية؟ وإذا لم يكن الأمر كذلك، فكيف يجب أن نبرمج تقنيات الذكاء الاصطناعي لاتخاذ القرار الصحيح في هذه اللحظات الصعبة؟

قد يرفض بعض الباحثين هذا النوع من السيناريو باعتباره حالة متطورة، بعيدة كل البعد عن دردشة الـ «جي بي تي» التي تقدم لنا مثلاً مقارنةً بين المهارات الشعرية للشاعر الإنجليزي «ويلفريد أوين» Wilfred Owen (١٨٩٣ - ١٩١٨) والشاعر الأمريكي «روبرت لي فروست» Robert Lee Frost (١٨٧٤ - ١٩٦٣)، لكن أعتقد أن هذا يخطئ الهدف، وهو ما نفضله في السطور التالية.

يبدأ مُعلم الفلسفة عادةً دروسه لطلابه بتوضيح الهدف من الفلسفة وصلتها الوثيقة بقضايا الحياة اليومية ومشكلاتها الكبرى. قد يبدأ مثلاً بإشراك الطلاب في التجربة الفكرية المعروفة باسم «معضلة القطار» أو «معضلة عربّة الترام» Trolley Car. بالنسبة لأولئك الذين ليسوا على دراية بهذه التجربة، فإن الأمر يسير على النحو التالي: أنت تقف عند مفترق طرق من السكك الحديدية في مواجهة قطار يندفع نحو هذا المفترق، وثمة طفل يجلس على المسار الرئيس للقطار، وبإمكانك سحب العتلة التي تتحكم في تحويل مسار القطار لكي تنقذ الطفل، فهل تفعل؟ هنا ترتفع أيدي الطلاب بنسبة مائة في المائة بالموافقة. حسنًا، ماذا إذن لو كان هناك خمسة أشخاص يجلسون على المسار الرئيس للقطار، بينما يجلس الطفل على المسار الثانوي الذي يمكن تحويل القطار إليه؛ هل تقف ساكنًا وتترك القطار يقتل خمسة أشخاص، أم تقوم بسحب العتلة لتحويل مسار القطار بحيث يقتل الطفل فقط؟ أيهما أكثر أخلاقية؟ هنا يُصبح الطلاب أقل ثقة، ويترددون في الإجابة، وقد لا يتفقون على إجابة واحدة على الإطلاق!

سرعان ما تكشف هذه التجربة الفكرية أنه في الغالبية العظمى من الحالات، لا توجد إجابة أخلاقية واضحة وصحيحة تمامًا، وهذا هو الهدف من الفلسفة، ومن الأخلاق كتخصص فرعي يندرج تحتها: مساعدتنا في التوصل إلى مجموعة من المبادئ الأساسية، وتطوير أدوات التفكير الناقد التي تمكننا من اتخاذ أفضل القرارات عند مواجهة مثل هذه المعضلات الأخلاقية.

ثمة موقع على الإنترنت يُقدم لك محاكاة لمعضلة القطار، مع عددٍ من السيناريوهات التي بإمكانك أن تختار أيًا منها، ليوضح لك في النهاية مدى اقتراب أو ابتعاد قراراتك عن قرارات الآخرين. قد تظن أن هذا الموقع (وغيره من المواقع المماثلة) قد استعان بباحثين متخصصين في فلسفة الأخلاق، وقد يكون هذا صحيحًا بشكلٍ جزئي، لكن المحاكاة في مجملها من تصميم باحثين متخصصين في الذكاء الاصطناعي، وهدفها «تكوين صورة جماعية للرأي البشري حول كيفية اتخاذ الآلات للقرارات عند مواجهة معضلات أخلاقية»، بقصد استخدام هذا النموذج لدعم عملية صنع القرار في السيارات ذاتية القيادة.

هذا أمر منطقي تمامًا، لأنه مع انتشار السيارات ذاتية القيادة الموجهة بالذكاء الاصطناعي، تتخذ السيارات بالفعل قرارات مستقلة بشأن ما يجب فعله عندما تسوء الأمور؛ مثلاً: ما الذي يمكن للسيارة أن تفعله في مواجهة القرار بين ضرب طفل في ممرٍ للمشاة مقابل رجل عجوز على كرسي متحرك؟ وماذا عن ضرب الطفل مقابل الانحراف والاصطدام بالحائط، مما قد يؤدي إلى إصابة ركاب السيارة أو قتلهم؟ وهل يجب على السيارات ذاتية القيادة حماية سلامة ركابها قبل كل شيء؟ أم يجب أن تكون الأولوية لتقليل الإصابات البشرية حتى ولو على حساب الركاب؟ هذه هي الأسئلة التي يطرحها صانعو السيارات لأول مرة في تاريخ الصناعة، وقد أنكر بعضهم هذه المشكلة، مشيرين إلى أن السيارات ذاتية القيادة يجب أن تقلل بشكلٍ كبير من عدد الحوادث على الطريق، ومجرد إبطاء السيارة عند مواجهة حادث محتمل يحل المشكلة. لكن الموجة الأخيرة من حوادث السيارات ذاتية القيادة الناجمة عن القرعة المفاجئة تدحض ذلك، وقد استدعت شركة تسلا في نوفمبر من سنة ٢٠٢٢ حوالي ١٢٠٠٠ سيارة بسبب احتمال وقوع حوادث نتيجة «الكبح المفاجئ» في أوضاع القيادة الذاتية بالكامل!

النقطة المهمة هنا هي أن حتمية اتخاذ القرارات من قبل الآلات تضع البشر مباشرةً وراء عجلة القيادة عندما يتعلق الأمر بتحديد القواعد والمقايضات المقبولة للقرارات الأخلاقية المعقدة. وكما يعلم أي دارس الفلسفة و/ أو الأخلاق، لا توجد خوارزمية يمكننا اللجوء إليها لاتخاذ هذه القرارات بشكلٍ موضوعي. ورغم كمّ التوقعات المُبشرة بشأن مستقبل الذكاء الاصطناعي التوليدي، لا يمكن للتكنولوجيا بحال من الأحوال أن تحل بشكلٍ مستقلٍ هذه الأنواع من المشكلات التي تحفل بها حياتنا!

ومع ذلك، لو عُدنا إلى المعضلة السابقة، لا يمكن للعربة ذاتية القيادة أن تُدعن لمن يتعامل معها بشريًا، ولا يمكن أن تنتظر الإجماع أو موافقة الأغلبية على القرار، بل يجب أن تعمل على أساس مجموعة محددة مسبقًا من القواعد، وفي ظل هذه الظروف، حتى التناقص عن العمل يغدو قرارًا له عواقبه!

نستطيع القول إذن أن مباحث الفلسفة والأخلاق والتفكير النقدي ستؤدي دورًا كبيرًا في عالم يؤدي فيه الذكاء الاصطناعي دورًا كبيرًا بالمثل في الأنشطة البشرية؛ سواء أكانا بإزاء قرارٍ بإطلاق صاروخ، أو انحراف سيارة، أو حتى استخدام اختصارات مثل ق. م (قبل الميلاد) في ورقة بحث تاريخية! وبعبارة أخرى، الذكاء الاصطناعي كان وسيظل في المستقبل المنظور يعتمد على قواعد العمل التي يوفرها البشر. وقد يبدو صادمًا أن شركات مثل جنرال موتورز وتسلا تُوظف الآن متخصصين في مجالي الفلسفة والأخلاق!

دعنا نتحدث عن كتابة أكواد البرمجة Coding. في سنة ٢٠١٨، كان ثمة ما يُقدر بنحو ١,٥ مليون وظيفة شاغرة في مجال التكنولوجيا في الولايات المتحدة وحدها. كانت الغالبية العظمى منها وظائف كتابة أكواد على مستوى المبتدئين، وقد سارعت بعض شركات البرمجة مثل «تريولوجي» Trilogy، والمؤسسات التعليمية مثل «فلاتيرون» Flatiron والمواقع التفاعلية مثل «كودكاديمي» Codecademy، بالإضافة إلى الكليات والجامعات التقليدية، إلى محاولة ملء الفراغ من خلال تقديم دورات تدريبية مكثفة لمساعدة الطلاب غير التقنيين على اكتساب المهارات التي يحتاجونها للعمل في مجال التكنولوجيا الرقمية، مع التركيز على مهارات كتابة الأكواد ولغات البرمجة والخوارزميات وعلوم البيانات.

لكن «الكود» ذاته يتم الآن - وبشكلٍ متزايد - إنشاؤه باستخدام أدوات الذكاء الاصطناعي. خذ مثلاً برمجيات تشغيل أنظمة القيادة الذاتية الخاصة بشركة تسلا، والتي يرجع فضل إنجاز الجزء الأكبر منها إلى عالم الحاسوب السلوفاكي - الكندي «أندرج كارباثي» Andrej Karpathy (من مواليد سنة ١٩٨٦)، المعروف على نطاق واسع باعتباره رائدًا بارزًا في علوم الحاسوب والذكاء الاصطناعي، وأيضًا أحد الأعضاء المؤسسين لشركة «أوبن إيه آي» OpenAI (الشركة المنتجة لبرنامج «شات جي بي تي»); ففي تغريدة له بتاريخ ٣٠ ديسمبر ٢٠٢٢، ذكر «كارباثي» أن ٨٠٪ من أكواده يتم إنتاجها (وبدقة تصل إلى ٨٠٪) بواسطة «جت هاب كوبيلوت» Github Copilot (وهو روبوت ذكاء اصطناعي مُولد، يقوم بالتكملة التلقائية للكود بطريقة ذكية، ويمكنه إنشاء دوال كاملة وليس مجرد تكملة للكلمات أو الجمل). وعلى حد تعبير «كارباثي»: «أنا لا أقوم بكتابة الكود بالفعل، أنا مجرد مُحرر، أحث الروبوت على كتابة الكود!» أو كما قال بتعبيرٍ أكثر بلاغة في تغريدة بتاريخ ٢٣ يناير ٢٠٢٣: «إن اللغة الإنجليزية هي

اللغة الأكثر رواجًا في البرمجة الجديدة»، بمعنى أنك لست في حاجة إلى تعلم لغات البرمجة لكي تُنشئ كودًا، فقط اطلب ذلك من الروبوت باللغة الإنجليزية!

الشيء الأهم في هذا الصدد هو أن قيام المُبرمج البشري أو الروبوت بكتابة كود مُعين يستلزم ما هو أكثر من مجرد الأمر البسيط؛ كأن تقول له مثلاً: أكتب لي كودًا بلغة «بايثون» Python (إحدى لغات البرمجة) لكي تظهر على الصفحة الرئيسية لموقعي جُملة «مرحبًا بالعالم» Hello World؛ هنا لا بد من أن يكون المُبرمج أو الروبوت مُلمًا مثلاً بماهية الموقع أو المدونة أو التدوينة التي أعمل على إنشائها، كما يستلزم فهمًا عميقًا لما هو ضروري لتحرير النص أو حذفه. هذه هي المفاهيم المُسبقة التي يمكن أن نتوقعها من مُبرمج يتمتع بمستوى تدريبي مميز، لكنه حتى الآن ليس شيئًا نتوقع أن يفهمه الحاسوب، وإن كان في طور التطوير!

نعم، يستطيع «جي بي تي» أن يُنشئ لك كودًا مُعقدًا أو ورقة بحثية متخصصة في أقل من دقيقتين، كما يستطيع أن يكمل كافة المهام في دورة تدريبية نموذجية لأكثر من ٢٠٠ ساعة في أقل من ساعة! ويعني ذلك ببساطة أن العالم لن يكون في حاجة إلى مبرمجين بشريين على الإطلاق في غضون خمس سنوات على الأكثر (باستثناء الفئة ذات المستوى الأعلى)! ليس ذلك فحسب، بل ولن يكون في حاجة إلى كُتاب للنصوص، أو وكلاء للسفر، أو مترجمين، أو محاسبين، أو مدرسين، أو مُحللي بيانات، أو حتى متخصصين في الدعم الفني؛ فكل هذه المهن مُهددة بالتقلص، وربما بالتلاشي أمام قطار الذكاء الاصطناعي المنطلق بلا هوادة! ومع ذلك ستزداد قيمة وأهمية اتخاذ القرار الأخلاقي، والأهداف الاستراتيجية، وبالتأكيد: الرؤي الفلسفية المستقبلية والتفكير الناقد، لأن مخرجات هذه العناصر ستكون الوقود الحيوي للاقتصاد المدفوع بالذكاء الاصطناعي، وكذا لتطوير الذكاء الاصطناعي ذاته وتحسينه بمرور الوقت.

على أن هذا لا يعني أن المدارس والجامعات والمؤسسات التعليمية والتدريبية لم يعد لها هدف، لكنها يجب أن تُعيد حساباتها وتعمل على الارتقاء في سلسلة القيمة؛ فالبرامج التي تركز على إدارة المنتجات والتسويق، وتصميم واجهات الدماغ والحاسوب، وعلوم البيانات، ومبيعات التكنولوجيا، وحتى القيادة التنفيذية، سيكون لها قيمة كبيرة في اقتصاد يحركه الذكاء الاصطناعي؛ وسيظل الطلاب بحاجة إلى إتقان المهارات الفنية والبحثية لتعزيز قدراتهم على التفكير النقدي واستخدام تقنيات الذكاء الاصطناعي بطرق هادفة.

لا شك أن تقنيات الذكاء الاصطناعي التوليدي يمكن أن تُساعدنا في إيجاد حلول لقضايانا المُلحة، فقط إذا قمنا - كجنس بشري - بتزويدها بوضوح بالمشكلات التي نحاول حلها بالفعل، والمواقف التي يكون الحل مقبولًا من خلالها. لا شك أيضًا أن «جي بي تي» يمكنه بسرعة حل مُعضلة القطار، إذا أخبرناه أن المشكلة التي نحاول حلها مثلاً هي المساواة العرقية، أو تعظيم الناتج المحلي الإجمالي، أو طول العمر، أو الانسجام المجتمعي. لكن ربما لا تتوافر حتى الآن

الرغبة الجدية في اتخاذ مثل هذه الأنواع من القرارات الأخلاقية كجزء من خارطة طريق «جي بي تي»!

References:

Boudette, N. E., Metz, C., & Ewing, J. (2022, June 15). *Tesla Autopilot and Other Driver-Assist Systems Linked to Hundreds of Crashes*. The New York Times. Retrieved April 13, 2023, from <https://www.nytimes.com/2022/06/15/business/self-driving-car-nhtsa-crash-data.html>

Chou, L. (2023, February 9). *Is Philosophy The New Coding in a Post-GPT World?* LinkedIn. Retrieved April 13, 2023, from <https://www.linkedin.com/pulse/philosophy-new-coding-post-gpt-world-luyen-chou>

Karpathy, A. (2022, December 30). *Nice read on reverse engineering of github copilot. COPILOT has dramatically accelerated my coding, it's hard to imagine going back to "Manual coding". still learning to use it but it already writes ~80% of my code, ~80% accuracy. I don't even really code, I prompt. & edit.* <https://t.co/kvqtoex9qj>. Twitter. Retrieved April 13, 2023, from <https://twitter.com/karpathy/status/1608895189078380544>

Karpathy, A. (2023, January 24). *The hottest new programming language is English*. Twitter. Retrieved April 13, 2023, from <https://twitter.com/karpathy/status/1617979122625712128>

Leslie, I. (2023, January 25). *Answer Machines*. Answer Machines - by Ian Leslie. Retrieved April 13, 2023, from <https://ianleslie.substack.com/p/answer-machines>

Moral machine. Moral Machine. (n.d.). Retrieved April 13, 2023, from <https://www.moralmachine.net/>

Spangler, T. (2017, November 23). *Morality, Ethics of a Self-Driving Car: Who Decides Who Lives, Dies?* Detroit Free Press. Retrieved April 13, 2023, from <https://www.freep.com/story/money/cars/2017/11/21/self-driving-cars-ethics/804805001/>

توثيق المقال بنظام APA:

عثمان، صلاح (١٢ أبريل ٢٠٢٣). «الفلسفة وتعويدة الجي بي تي». أكاديمية بالعقل نبداً، القاهرة. تم الاسترداد

بتاريخ ١٣ أبريل ٢٠٢٣ من:

<https://mashroo3na.com/إصدارات/مقالات/الفلسفة-وتعويدة-الجي-بي-تي/>

APA Citation:

Osman, S. (عثمان، ص) (2023, April 12). Philosophy and the Mantra of GPT! (الفلسفة وتعويدة الجي بي تي). Retrieved April 13, 2023, from <https://mashroo3na.com/إصدارات/مقالات/الفلسفة-وتعويدة-الجي-بي-تي/>
