

مقالة رأي وتحليل

التاريخ التحريري: ٣١ أكتوبر ٢٠٢٦

التوقيع المتأخر: من يملك قرارًا اتخذته الآلة؟

> في صيف 2020 بدت درجة الامتحان كأنها إجابة حسابية عن سؤال تعليمي. لكن البحث عمّن «قرر» تلك الدرجة قاد إلى سلسلة من البشر، يملك كل واحد منهم جزءًا من السلطة ولا يملك أحدهم وحده الصورة كلها. هذه ليست حكاية عن حاسوب أخطأ فحسب، بل عن مؤسسة قسمت القرار ثم نسيت أن تجمع المسؤولية.

الدرجة التي لم يكتبها طالب في ورقة امتحان

في 13 أغسطس 2020 أُعلنت في إنجلترا نتائج المستوى المتقدم، المعروف باسم **A-level**، وهو امتحان تؤثر درجاته في انتقال آلاف الطلاب إلى الجامعة. غير أن الامتحانات لم تُعقد ذلك الصيف. كان وباء كوفيد-19 قد دفع الحكومة البريطانية، في 18 مارس، إلى إلغاء الاختبارات المقررة، ووُضعت آلية تمنح الدرجات اعتمادًا على تقديرات المدارس وترتيبها للطلاب ومعالجة إحصائية تهدف إلى ضبط توزيع النتائج قياسًا بالأعوام السابقة.12

الدرجة التي ظهرت للطالب كانت محددة لا تقبل نصف حرف. أما الطريق إليها فكان موزعًا: معلمون قدموا تقديراتهم وترتيب طلابهم، وجهة تنظيمية صممت طريقة المعايرة، ومجالس امتحانات أصدرت النتائج وفق متطلبات رسمية، ووزارة حددت مقصد السياسة العامة. لم تكن المسألة أن روبوتًا جلس إلى مقعد المصحح. لم يوجد أصلًا امتحان مكتوب يصححه. كانت الدولة تحاول إنشاء بديل لواقعة تعليمية لم تحدث.

بعد أربعة أيام فقط، في 17 أغسطس، تبدلت القاعدة. أصبحت النتيجة النهائية، في الحالات ذات الصلة، هي الأعلى بين تقدير المدرسة والدرجة المحسوبة. وفي تحليل نشرته هيئة تنظيم الامتحانات **أوفكوال (Ofqual)** لاحقًا لعينة محددة من طلاب إنجلترا بعمر 18 عامًا، شملت **457,420 تسجيلًا لمادة امتحانية** تعود إلى **246,110 طالبًا**، تطابقت الدرجة المحسوبة

مع تقدير المدرسة في نحو **59% من التسجيلات**، وكانت أدنى منه في **39%**، وأعلى منه في نحو **2%**. هذه نسب *تسجيلات المواد* لا نسب الطلاب؛ لذلك لم يكن من الدقة القول إن 39% من الطلاب خُفضت درجاتهم.3

لا يثبت اختلاف الدرجة عن تقدير المعلم أن المعلم كان مصيباً أو أن الحاسوب ظلم الطالب في كل حالة. لكنه يثبت شيئاً آخر: أن المنظومة استطاعت أن تصدر حكمًا ذا أثر شخصي كبير، ثم تغير أساس إصدار ذلك الحكم بقرار مؤسسي خلال أيام. فمن كان صاحب القرار الأول؟ من تحمّل تبعته؟ ومن كانت له سلطة منعه قبل أن يصل إلى الطالب؟

في 2 سبتمبر 2020 كتب رئيس مجلس أوفكوال، روجر تايلور (Roger Taylor)، إلى لجنة التعليم في البرلمان البريطاني معلنًا، في ترجمة موجزة لعبارته، أن الهيئة **تقبل نصيبها من المسؤولية كاملًا**. وأوضح أن قرار اعتماد تقديرات المعلمين بعد توحيدها إحصائيًا صدر في إطار توجيه وزاري، بينما تقع مسؤولية التنفيذ على الهيئة. ثم وصل إلى استنتاج لافت: لم يكن تحسين الخوارزمية وحده كافيًا لجعل النتائج مقبولة؛ فحدود البيانات وطبيعة العملية نفسها كانتا جزءًا من المشكلة.4

من السهل قراءة الحادثة باعتبارها «فضيحة خوارزمية» والانصراف. لكن هذا الاسم يعفي البشر على نحو عجيب. فالمعادلة لا تصدر توجيهًا وزاريًا، ولا تختار أن تكون العدالة مساوية لتشابه توزيعات الدرجات بين سنتين، ولا تقرر مقدار الضرر المقبول إذا اختلفت النتيجة الفردية عن توقع الطالب. تلك قرارات بشرية سبقت الحساب.

الآلة قد تنفذ الاختيار. لكنها لا تزيل التاريخ البشري للاختيار.

أربع سلطات اختبأت في كلمة واحدة

تقول المؤسسات عادةً إن «القرار اتُّخذ» وكأن للفعل مالكًا واحدًا واضحًا. لكن القرار المؤسسي ليس نقطة واحدة. إنه انتقال بين حقوق مختلفة: حق اقتراح الفعل، وحق اعتماد معيار المفاضلة، وحق التصديق على النتيجة، وحق وقفها أو تصحيحها.

في بحثهما «**الفصل بين الملكية والسيطرة**» (**Separation of Ownership and Control**) المنشور عام 1983، فصل الاقتصاديان يوجين فاما (Eugene F. Fama) ومايكل جنسن (Michael C. Jensen) بين **إدارة القرار** (**Decision Management**)، وتشمل اقتراح الخيارات وتنفيذها، و**الرقابة على القرار** (**Decision Control**)، وتشمل اعتماد الخيارات ومراقبة الأداء. كان موضوعهما بنية المؤسسات وتعارض مصالح المديرين والمالكين، لا الذكاء الاصطناعي. لكن الفصل يساعد على رؤية ما تخفيه واجهة حديثة تعرض النتيجة بكبسة زر.5

خذ مثالًا مصرفيًا **افتراضيًا**: نموذج يتوقع احتمال تعثر طالب ائتمان. الفريق التقني يُنتج التقدير؛ فريق المخاطر يختار الحد الذي يفصل بين القبول والرفض؛ الإدارة تعتمد حجم الخسائر المقبولة؛ نظام التشغيل يرسل الإشعار؛ موظف خدمة العملاء يشرح النتيجة؛ وكيان المصرف المرخّص هو الطرف الذي يرتبط بعلاقة نظامية وتعاقدية مع العميل. من صاحب قرار الرفض؟

يمكن إعطاء جوابين صحيحين على سؤالين مختلفين. **صاحب المسؤولية القانونية الخارجية** تحدده الأنظمة والعلاقة التعاقدية والوقائع، ولا تزول مسؤوليته لمجرد شراء نموذج من مورّد. أما **صاحب القدرة السببية الداخلية** فقد يكون من اعتمد متغيراً، أو حدد عتبة، أو أمر بالعمل دون مراجعة. ليس أحد الوصفين بديلًا عن الآخر. التحقيق الجاد يحتاجهما معًا.

الطبقة	السؤال الذي تجيب عنه	السلطة اللازمة	ما لا يكفي لإثباتها
تحديد الغاية	ما المصلحة التي يُسمح للنظام بتحسينها؟	اعتماد الهدف والقيود والحقوق غير القابلة للمقايضة	شراء «نموذج عالي الدقة»
بناء الوسيلة	ما البيانات والافتراضات والتكاليف التي تحول الغاية إلى حساب؟	اختبار النموذج والبدائل وحدود صلاحيته	عرض درجات الثقة للمستخدم
نفاذ النتيجة	متى تصبح التوصية قرارًا مؤثرًا؟	اعتماد العتبات والصلاحيات وحدود الأتمتة	توقيع إلكتروني آلي باسم قسم
التصحيح والتعويض	من يستطيع وقف الأثر أو إبطاله أو إعادة تقييمه؟	صلاحية نافذة ووقت وموارد ومسار تظلّم	بريد اعتراض لا يملك صاحبه تغيير شيء

هذه ليست قائمة جديدة للمسؤوليات القانونية في بلد بعينه، بل خريطة تشغيلية مقترحة لتعقبها. الفشل يحدث حين تُفصل السلطات الأربع دون وصلات واضحة: من يراجع صحة النموذج لا يملك تغيير الغاية؛ ومن يوقع على الغاية لا يفهم ما تفعله المعادلة؛ ومن يواجه العميل لا يملك إلا قراءة السبب الذي صيغ له سلفًا.

صاحب السلطة المكتوبة ليس دائمًا صاحب السلطة الواقعية

في عام 1997 نشر فيليب أغيون (Philippe Aghion) وجان تيرول (Jean Tirole) بحثًا بعنوان «**السلطة الرسمية والسلطة الفعلية داخل المؤسسات**» (**Formal and Real Authority in Organizations**). الفكرة التي تهمنا هنا بسيطة وعميقة: امتلاك الحق الرسمي في الحسم لا يعني امتلاك القدرة الفعلية على اختيار البديل، خصوصًا حين يمتلك المرؤوس المعلومات والتفاصيل التي لا يستطيع الرئيس استيعابها في الوقت المناسب.6

يتضاعف هذا الانفصال مع النماذج المعقدة. قد يكون مدير المخاطر مخولًا اعتماد السياسة، لكنه لا يعرف أثر إزالة أحد المتغيرات على آلاف الحالات. وقد يعرف عالم البيانات أثر المتغير، لكنه غير مخول تقرير ما إذا كان الأثر مقبولًا من الناحية الائتمانية أو القانونية. وفي الاجتماع، تبدو عبارة «اعتمدنا النموذج» حاسمة. لكنها قد تغطي ثلاثة أفعال متباينة: مراجعة العرض، وفهم المخاطر، وقبول التبعات. ليس كل من فعل الأول قد فعل الثاني أو الثالث.

ثم تظهر مشكلة الوكيل والأصيل (**Principal–Agent Problem**). يفوض مجلس الإدارة التنفيذ إلى الإدارة، وتفوض الإدارة مهمة فنية إلى فريق، وقد يفوض الفريق استضافة النظام أو تشغيله إلى مورّد. كل حلقة تحاول تحقيق هدف مشروع من زاويتها: تسليم المشروع في موعده، أو خفض التعثر، أو تقليل زمن الإجابة، أو رفع معدل القبول. لكن الاتفاق على أهداف محلية لا يضمن أن يُتخذ القرار النهائي وفق المصلحة التي من أجلها أُنشئت المؤسسة.

في هذه السلسلة قد يصبح صاحب الصلاحية آخر من يعلم بما تسمح به الصلاحية. ليس الأمر مؤامرة ولا تقاعسًا بالضرورة؛ إنها نتيجة متوقعة لتفاوت المعرفة والوقت وأدوات التأثير. وما لم تُصمّم المؤسسة عمدًا منافذ تعيد السلطة الفعلية إلى من يتحمل المسؤولية، فإن «المراجعة البشرية» قد تتحول إلى توقيع يصل بعدما يصبح الرجوع مكلفًا.

عندما تكون المعادلة صحيحة والسؤال خاطئًا

لنعد إلى إنجلترا. لم يكن استقرار توزيع درجات سنة 2020 مقارنة بالسنوات السابقة هدفًا عبثيًا. لو جرى اعتماد تقديرات المدارس دون معايرة، أمكن أن تختلف سخاوة التقديرات من مدرسة لأخرى، وأن تتضخم الدرجات، وأن تتأثر المقارنة بين دفعات الطلاب. هذه مخاوف أعلنتها أوفكوال في سجلاتها. وفي المقابل، حوّل توحيد التوزيع الإحصائي بعض الفروق بين المدارس إلى أثر على درجة طالب محدد لم تُتح له فرصة إثبات مستواه في اختبار فعلي.14

لذلك كان من الممكن أن يكون النموذج منسجمًا مع غايته المحددة، بينما تكون الغاية نفسها غير كافية لتأسيس قرار مقبول. الفرق بين **صحة الحساب داخل النموذج ومشروعية استخدامه للحكم على فرد** ليس فرقًا في سرعة المعالج. إنه اختلاف في السؤال الذي وُجه إلى المعالج أولًا.

وهنا تُظهر الواقعة خطرًا يصعب رؤيته من داخل فرق التقنية: قد يقيّم المصمم نجاحه بمدى انضباط توزيع النتائج؛ وتقيس الحكومة نجاحها باستمرار العملية التعليمية؛ وتقيس الجامعة أهلية المتقدم للقبول؛ ويقيس الطالب عدالة فرصته في إظهار ما تعلمه. هذه ليست وحدات قياس لنفس الشيء، بل مصالح وحقوق قد تتعارض. لا تستطيع دالة هدف واحدة أن تصالحها تلقائيًا؛ لا بد من قرار سياسي أو مؤسسي صريح بشأن ترتيبها وحدود التضحية بها.

بعد الأزمة، نشر **مكتب تنظيم الإحصاءات البريطاني (Office for Statistics Regulation)** في 2 مارس 2021 مراجعة تستخلص دروس التجربة. ومن أهمها ضرورة تمييز مسؤولية **الجهة التي تطلب النموذج وتحدد الغرض منه** عن مسؤولية **من يطوره**، وفحص ملاءمة النموذج للغرض وقابليته للقبول العام، وتحديد أساس الاعتراض وتوقيته. لم تكن التوصية أن تتوقف الدولة عن استخدام النماذج، بل أن تتوقف عن اختزال قراراتها العامة في ما يبدو تمرينًا تقنيًا.7

ما الذي يقرره النموذج فعلًا في طلب قرض؟

لنفترض، لأغراض الشرح فقط، أن مصرفًا يستخدم نموذجًا يتنبأ باحتمال سداد القرض، نرمز له بـ $p(x)$، حيث x مجموعة المعلومات المتاحة عن الطلب. ولنفترض أن B هو العائد الإضافي إذا سُدّد القرض، وL الخسارة إذا تعثر، وC كلفة تقديم الخدمة. في نموذج مبسط ومحايد للمخاطر قد تكون قاعدة القبول:

اقبل الطلب إذا كان

$$p(x)B - (1 - p(x))L - C > 0.$$

هذه ليست قاعدة مصرفية حقيقية ولا توصية ائتمانية. القروض الفعلية أعقد، والقرار مقيد بالسيولة ورأس المال والأنظمة والمساواة في المعاملة ومخاطر النموذج وحقوق العميل. المعادلة هنا كاشفة لشيء واحد: قد ينتج النموذج **التوقع**، لكنه لا يحدد بالضرورة قيمة B وL وC، ولا ما إذا كانت المؤسسة تقبل صياغة الأهداف بهذه الطريقة. حتى اختيار «الربح المتوقع» بوصفه غاية وحيدة هو قرار يمكن الاعتراض عليه.

ثم إن عبارة «يتنبأ بالسداد» تخفي مسألة معرفية. التنبؤ ليس وعدًا، واحتمال السداد ليس هو القدرة الأخلاقية على الوفاء، ولا يساوي بالضرورة الأهلية القانونية للحصول على التمويل. إذا أُدخلت في النموذج بيانات غير ممثلة أو تاريخ ائتماني تشكل تحت معايير قديمة، فإن الرقم قد يحمل آثار قرارات سابقة. لا يعفينا اكتشاف هذه الحدود من اتخاذ القرار؛ لكنه يغير ما يُسمح لنا بادعائه عن أساسه.

لنتخيل أن احتمال السداد المقدّر لعميل هو 0.79، وأن مديرًا اختار عتبة 0.80. لا يستطيع محلل البيانات إبلاغ العميل بأن 0.79 «أمر النموذج» فحسب؛ فقد كان هناك قرار بشري جعل الفارق بين 0.79 و0.80 حدًا ذا أثر. وإذا تغيرت تكلفة رأس المال أو ميزانية المخاطر، قد تتحرك العتبة دون أن يتغير النموذج ولا بيانات العميل. هكذا تُنتج المؤسسة أحيانًا قرارات جديدة بمجرد تغيير زر في السياسة، ثم تنسبها خطأً إلى «ذكاء» النظام.

يهم هذا التمييز طالب القرض الذي يريد سببًا يمكن فحصه وتصحيحه، ويهم موظف المخاطر الذي قد يُطلب منه تفسير اختيار لم يشارك في تصميمه، ويهم المساهم الذي يحتاج إلى معرفة من غيّر شهية المخاطرة، ويهم الجهة المشرفة التي لا يكفيها أن تجد اسم نموذج في سجل النماذج.

لماذا لا تصلح عبارة «لا نفهم الصندوق الأسود» اعتذارًا؟

تُبيّن قواعد الائتمان الأمريكية فرقًا مفيدًا، **ضمن نطاقها القضائي وحده.** فالمادة §1002.9 من اللائحة B (Regulation B)، في صيغتها المنشورة الحالية، تنظم إبلاغ طالب الائتمان بالقرار السلبي وبيان أسبابه المحددة وفق شروط وإجراءات تتباين باختلاف نوع الطلب. وفي تفسيرها الرسمي تُعد عبارات مثل «لم تحقق درجة التأهيل في نظام التصنيف» غير كافية وحدها؛ المطلوب أن تتصل الأسباب بالعوامل التي استُخدمت بالفعل.8

كانت هيئة حماية المستهلك المالي الأمريكية (Consumer Financial Protection Bureau, CFPB) قد أصدرت في 2022 تعميمًا يؤكد عدم كفاية غموض الخوارزمية عذرًا لعدم بيان الأسباب. لكن التعميم نفسه، مع تعميم ذي صلة من 2023، **سُحب في 12 مايو 2025.** لذلك لا يصح نسب موقفه التفسيري السابق إلى سياسة الهيئة الراهنة؛ الأهم هنا هو التمييز بين وثيقة مسحوبة ونص اللائحة الجاري، لا إصدار فتوى حول التزامات مصرف خارج الولايات المتحدة.9

قد يقول أحدهم: نعرف سبب الرفض لأن لدينا تقنية تفسر النموذج، مثل أساليب إسناد أهمية المتغيرات. لكن تفسير التوقع الحسابي ليس بالضرورة تفسير القرار المؤسسي. فالأول يخبرنا أي مدخلات أسهمت في الرقم وفق طريقة تحليل محددة؛ والثاني يجب أن يبين لماذا اختار المصرف هذا الرقم معيارًا، وكيف عُرضت بيانات العميل، وأي سياسة جعلت النتيجة نافذة. بل قد تعطي طرق التفسير التقريبي تفسيرات متباينة لنفس النموذج، ولا تمثل دائمًا علاقة سببية بين كل متغير والنتيجة الواقعية.

في الجهة المقابلة، قد يكشف النظام عن أسبابه الحسابية جميعًا ويظل القرار غير مقبول. من قال إن امتلاك بيانات تاريخية يخول استخدامها في كل غرض؟ ومن أقر أن خطأً معينًا في الرفض يمكن أن يتحمله العميل وحده؟ الشفافية نافذة مهمة، لكنها لا تُحوّل تلقائيًا قرارًا غير مشروع إلى قرار مشروع. والمعرفة بأن الباب أُغلق لا تعني امتلاك مفتاح فتحه.

خطأ المؤسسة حين توظّف «إنسانًا للاعتماد»

تَعِد حلول كثيرة بإضافة خطوة مراجعة بشرية (Human-in-the-Loop) إلى منظومة القرار. إنها أداة مفيدة في ظروفها، لكن الاسم لا يقول شيئًا عن صلاحية الشخص الذي يجلس داخل الحلقة. هل يرى الوقائع ذات الصلة؟ هل يعرف حدود النموذج؟ هل يستطيع مخالفة التوصية دون عقوبة ضمنية؟ هل يملك وقتًا كافيًا؟ وهل تتوقف النتيجة عن النفاذ إلى أن ينتهي من التقييم؟

لا تتحقق الرقابة بأن تضيف المؤسسة مربعًا صغيرًا مكتوبًا عليه «اعتمد». قد يصبح الموظف، مع مئات التوصيات المتشابهة، أكثر شبهًا بآلة ختم منها بصاحب حكم. والمفارقة الإدارية هنا لطيفة في صورتها ومكلفة في آثارها: تُدخل المؤسسة الإنسان إلى النظام لتثبت أنه لا يزال موجودًا، ثم تصمم عمله بحيث لا يستطيع الاختلاف معه.

لهذا يعالج **إطار إدارة مخاطر الذكاء الاصطناعي** (AI Risk Management Framework) الصادر عن المعهد الأمريكي للمعايير والتقنية (NIST) في 2023 المسألة بلغة المسؤوليات الفعلية؛ فقسم **GOVERN 2.1** يطلب توثيق الأدوار وخطوط التواصل، بينما يسند **GOVERN 2.3** مسؤولية القرارات المتعلقة بمخاطر تطوير النظم ونشرها إلى القيادة التنفيذية. الإطار **طوعي**، وليس قانونًا يمنح الأشخاص سلطات لم يكونوا يملكونها.10

وبينما ينص **قانون الذكاء الاصطناعي الأوروبي** (EU AI Act) في المادة 14 على أن تكون الرقابة البشرية على الأنظمة عالية المخاطر قادرة، بحسب السياق، على فهم قدرات النظام وحدوده، وتجاهل مخرجاته أو عكسها، والتدخل لإيقافه، فإن تطبيق واجبات الأنظمة عالية المخاطر يجري على مراحل؛ وبحسب المفوضية الأوروبية بعد تعديلات 2026، تبدأ أحكام عدد من استعمالات الملحق الثالث في **2 ديسمبر 2027**، وأحكام الأنظمة المدمجة في منتجات منظمة في **2 أغسطس 2028**. من الخطأ نقل النص إلى بنك سعودي كأنه التزام أوروبي نافذ عليه الآن، أو افتراض أن كل بند ينطبق على كل نموذج.11

المغزى التصميمي، بعيدًا عن الحدود الجغرافية، أن الرقابة البشرية ليست وصفًا لموقع الموظف على مخطط التدفق؛ إنها **قدرة عملية على الاعتراض قبل أن يصبح الاعتراض بلا قيمة.**

السعودية: من توزيع المسؤوليات إلى اختبارها

تضع **المبادئ الرئيسة للحوكمة في المؤسسات المالية الخاضعة لرقابة وإشراف البنك المركزي السعودي**، الصادرة في 30 يونيو 2021 والمحدّثة في نسخة ثانية بتاريخ 8 مايو 2024، أساسًا تنظيميًا قائمًا لتحديد دور مجلس الإدارة والإدارة التنفيذية واللجان والرقابة وإدارة المخاطر. وهي سارية بحسب سجل قواعد البنك المركزي. لكن هذه المبادئ ليست، في حد ذاتها،

«لائحة خاصة بتوقيع قرارات نماذج الذكاء الاصطناعي»؛ وإذا أردنا تطبيقها على الأنظمة الخوارزمية فعلينا تحويل معانيها إلى صلاحيات وإجراءات قابلة للاختبار، دون اختراع التزام نصي غير موجود.12

قد يسأل مجلس إدارة مصرف سعودي، في تجربة مراجعة داخلية مقترحة، عن «نسبة القرارات التي اتخذها الذكاء الاصطناعي». يبدو السؤال عصريًا، لكنه ناقص. السؤال الأدق: في أي مرحلة يتحول خرج النموذج إلى أثر يغيّر حقًا أو منفعة؟ من وافق على تلك المرحلة؟ وما الخسارة التي يمكن أن تقع قبل تدخل الموظف المخوّل؟ وإذا طلب العميل تصحيح معلومة أو طعن في حساب، هل يملك من يستقبل طلبه سلطة تغيير النتيجة أم سلطة إرساله إلى صندوق بريد آخر؟

إذا كانت لجنة المخاطر تجيز نموذجًا معينًا ثم تقوم إدارة المنتجات بتغيير عتبة القبول من دون اعتماد مناسب، فاختبار الدقة الإحصائية لن يكتشف بذاته هذه المشكلة في **حقوق القرار (Decision Rights)**. وإذا كان مورّد خارجي يستطيع تغيير إصدار النموذج في البيئة التشغيلية دون ضوابط تعاقدية وفنية ملائمة، فإن اسم المصرف على رسالة الرفض لا يشرح وحده أين انتقلت السيطرة الفعلية. وإذا تعلّم النظام من تفاعلات جديدة، فقد يتغير سلوكه بينما تبقى الورقة الموقعة على الإصدار السابق معلّقة في مجلد الامتثال، مطمئنة تمامًا إلى العالم الذي لم يعد موجودًا.

تنبه دراسة لمعهد الاستقرار المالي التابع لبنك التسويات الدولية **(Financial Stability Institute, BIS)** منشورة في ديسمبر 2024 إلى تحديات استخدام الذكاء الاصطناعي في القطاع المالي، من بينها الحوكمة، وإدارة النماذج، والاعتماد على الأطراف الثالثة. لا تعطي الدراسة حكمًا على كل مشروع أو جهة في المملكة، لكنها تُظهر أن انقسام المعرفة بين البنك والمورّد ليس سؤالًا محليًا هامشيًا.13

سلطة التصحيح النافذة: مفهوم مقترح لا شهادة جاهزة

أقترح هنا تعبير «**سلطة التصحيح النافذة**» **(Executable Correction Authority)** بوصفه **أداة تحليلية مقترحة في هذه المقالة**، لا نظرية أكاديمية معترفًا بها ولا معيارًا تنظيميًا قائمًا. المقصود هو اجتماع ثلاثة شروط في قرار ذي أثر: شخص محدد يتحمل مهمة المراجعة، ومعلومات كافية لفهم أساس القرار وحدوده، وصلاحية عملية وقانونية لتغيير النتيجة أو وقف نفاذها في الزمن الذي لا يزال فيه التصحيح ذا معنى.

يمكن تحويل الفكرة إلى اختبار مؤسسي من دون ابتكار مؤشر مقدس جديد. تختار جهة العمل عينة معلنة من القرارات المؤثرة التي صدرت فعلًا، وتعيد تنفيذ سيناريو اعتراض مشروع في بيئة آمنة، وتسأل لكل قرار: هل أمكن الوصول إلى المسؤول المخوّل؟ هل استطاع إعادة بناء الأسباب الجوهرية دون التخمين؟ هل استطاع، عند ظهور خطأ مثبت، إنفاذ تصحيح خلال مهلة معلنة للمستفيد؟ ثم تُحسب نسبة الحالات التي اجتازت **الشروط الثلاثة جميعًا** إلى إجمالي الحالات المختبرة، مع نشر التوزيع حسب نوع القرار ودرجة خطورته بدل الاكتفاء بمتوسط يخفي الأشد ضررًا.

هذا المقياس التجريبي **لا يثبت سلامة القرار الأصلي**، ولا يُنشئ حقًا قانونيًا في التراجع عن إجراء لا يقبل الرجوع، وقد يتحسن ظاهريًا إذا اختيرت عينات سهلة أو وُضعت مهلة متساهلة. إنه يكشف جانبًا واحدًا من المساءلة: قدرة المؤسسة على جعل التصحيح نافذًا. وفرضية المقالة القابلة للاختبار أن الأنظمة التي توضح هذا الحق مبكرًا وتختبره دوريًا ستقل فيها **القرارات**

الخاطئة التي تستمر بعد اكتشاف الخطأ، مقارنة بأنظمة لديها مسار شكوى اسمي فقط. يلزم اختبار ميداني مضبوط قبل التعامل مع الفرضية باعتبارها نتيجة علمية.

وهناك أربعة أسئلة عملية أكثر صدقًا من سؤال «هل لدينا إنسان في الحلقة؟». هل يستطيع صاحب الصلاحية رؤية البيانات الحاسمة؟ هل يستطيع الاعتراض على *السياسة* كما يعترض على *النتيجة*؟ هل يوقف قرارًا عاجلًا قبل تنفيذ أثر يصعب التراجع عنه؟ وهل يستطيع المستفيد نفسه معرفة طريق الاعتراض دون معرفة اللغة الداخلية لفرق النماذج؟ يدعو الدليل المشترك الصادر عن **مكتب مفوض المعلومات البريطاني** (ICO) و**معهد آلان تورنغ** (The Alan Turing Institute) إلى تفسير القرارات التي تدعمها أنظمة الذكاء الاصطناعي للأشخاص المتأثرين بها؛ والتفسير هنا جزء من العلاقة المؤسسية، لا عرض للرياضيات فقط.14

الاعتراض الأقوى: البشر ليسوا أكثر عدلًا لمجرد أنهم بشر

ليس البديل الحتمي لهذه المسائل أن يعود كل ملف إلى مكتب موظف يقرر وفق خبرته وحدسه. قد يكون البشر متناقضين ومتعبين، وتختلف أحكامهم باختلاف المزاج والخبرة وضغط العمل. ويمكن لنموذج مُختبر جيدًا، يستخدم معايير متسقة، أن يقلل أنواعًا من التفاوت ويرصد أخطاءً تفوت المراجعة اليدوية. كما قد تتسبب المراجعة البشرية لكل معاملة في تأخر يضر المحتاجين للخدمة. هذه حجة جدية، لا دفاعًا علاقاتيًا عن الأتمتة.

كذلك لا تثبت واقعة نتائج 2020 أن كل القرارات المبنية على الإحصاء غير صالحة. تقرير مكتب تنظيم الإحصاءات أقر صراحةً بأن العاملين في الجهات المنظمة سعوا بحسن نية إلى مخرجات دقيقة، وأن المهمة تعرضت لقيود زمنية وبيانية استثنائية، وأن التراجع عن استخدام النماذج الإحصائية عمومًا قد يفوت منفعة عامة حقيقية.7 والمشكلة الأخرى أن القرار البشري قد لا يكون قابلًا للتفسير أصلًا إذا لم توثق أسبابه.

لكن الاعتراض لا يهدم الأطروحة؛ إنه يصحح طريق تطبيقها. فالمساءلة لا تتطلب أن يكرر البشر كل عملية حسابية. يمكن توزيع الرقابة بحسب المخاطر: تفويض آلي واضح للحالات البسيطة والقابلة للعكس؛ مراجعة للحدود الحساسة؛ إيقاف إلزامي عند تجاوز التفويض؛ وعملية مستقلة تختبر عينة من القرارات، بما فيها الحالات التي أقرها الموظف دون تعديل. هذه ليست وصفة موحدة لجميع القطاعات، بل طريقة لإبقاء السلطة متناسبة مع أثر القرار.

يؤيد **توجيه مكتب مفوض المعلومات البريطاني** بشأن حوكمة الذكاء الاصطناعي ضرورة معرفة من يحدد أغراض معالجة البيانات ووسائلها ومن يعمل بمعالجة البيانات نيابة عن غيره؛ وقد تختلف الصفة القانونية بحسب المرحلة والعلاقة. لا يجوز أن تُختزل كل مسؤولية في عقد يصف المورّد بأنه «معالج» إذا كانت السيطرة الفعلية والاختصاص القانوني مختلفين. وهذه قاعدة بشأن حماية البيانات في نطاقها، وليست حكمًا شاملًا على المسؤولية المدنية أو المصرفية عن جميع مخرجات الذكاء الاصطناعي.15

الاختبار الحاسم إذن ليس مَن يضع إصبعه على زر الموافقة، بل مَن يستطيع أن يقول «هذه القاعدة غير مقبولة» وأن تترتب على اعتراضه نتيجة مؤسسية نافذة. وما لم توجد تلك القدرة فلن ينقذ كثرة التوقيعات منظومة القرار. قد يتحول التوقيع، في أسوأ أحواله، إلى مراسم لتوزيع الملامة بعد توزيع السلطة.

قبل النتيجة التالية

أعادت إنجلترا إصدار درجات ذلك الصيف على أساس مختلف. لم يعد بالإمكان إرجاع الطلاب إلى قاعات الامتحان يوم 13 أغسطس، ولم تكن إعادة الحساب مماثلة لاستعادة الأيام التي عاش فيها بعضهم أثر النتيجة الأولى. لهذا كان التراجع مهمًا ومحدودًا في وقت واحد. فالقرار ليس رقمًا في قاعدة بيانات فحسب؛ إنه فعل يدخل حياة الناس في ساعة محددة، ثم قد يترك أثرًا يواصل سيره حتى بعد تعديله.

الأثر الأعمق لتجربة أوفكوال ليس أنها حذّرتنا من «الخوارزمية الظالمة» بعبارة سهلة. لقد أظهرت أن الحكم قد يصبح شديد الوضوح على الورق، شديد الغموض في خريطة الصلاحيات. عندما تتقدم النماذج القادرة على كتابة التوصية وتحديد المخاطر وإرسال القرار، سوف يقل عدد اللحظات التي يتوقف عندها سير العمل طالبًا رأيًا بشريًا صريحًا. عندها يجب ألا يكون سؤال رئيس مجلس الإدارة «أي نموذج اتخذ القرار؟»، بل «أي قرار مني جعل هذا النموذج قادرًا على إحداث هذا الأثر، ومن يملك إيقافه إذا ثبت أننا أخطأنا؟».

كانت الدرجة المطبوعة في أغسطس 2020 تحمل حروفًا قليلة. خلف كل حرف سلسلة طويلة من القرارات. وإذا تعلمت المؤسسة الحديثة شيئًا منها، فلن تكتفي بأن تجد توقيعًا في آخر السلسلة. ستبحث عمن كان يملك القلم في بدايتها.

المراجع

1. Awarding GCSE, AS & A levels in summer 2020: interim report، **أوفكوال**، 13 أغسطس 2020، تقرير تنظيمي أولي أصلي؛ يصف آلية الدرجات والسياق وقيود التقديرات، وتفاصيله تخص إنجلترا.
2. Requirements for the calculation of results in summer 2020، **أوفكوال**، صدر 7 يوليو 2020 وعدّل حتى 20 أغسطس؛ متطلبات أصلية لمجالس الامتحانات، مع ضرورة الانتباه إلى أن الصفحة تعرض نسخًا تلت تغيير السياسة.
3. Grading gaps in summer 2020: who was affected by differences between centre assessment grades and calculated grades?، **أوفكوال**، تحليل رسمي لاحق لنتائج 2020؛ يشمل 457,420 تسجيل مادة من 246,110 طالبًا من فئة عمرية ونطاق محددين، ولا يقيس صدق كل تقدير فردي ولا كامل سكان المملكة المتحدة.
4. Written statement from Chair of Ofqual to the Education Select Committee، **روجر تايلور، أوفكوال**، 2 سبتمبر 2020؛ شهادة/تصريح أولي يحدد توجيه الوزير ومسؤولية الهيئة ويعرض دفاعها ونقدها الذاتي.
5. Eugene F. Fama & Michael C. Jensen، Separation of Ownership and Control، *The Journal of Law and Economics*، 1 يونيو 1983؛ بحث أكاديمي مؤسس للفصل بين إدارة القرار ومراقبته، لا دراسة تجريبية لنظم الذكاء الاصطناعي.
6. Philippe Aghion & Jean Tirole، Formal and Real Authority in Organizations، *Journal of Political Economy*، 1997؛ بحث نظري في اختلاف السلطة المعلنة عن التأثير الفعلي نتيجة توزيع المعرفة.
7. Office for Statistics Regulation، Ensuring statistical models command public confidence، 2 مارس 2021؛ مراجعة رسمية مستقلة نسبيًا لدروس نظم درجات 2020 في بلدان المملكة المتحدة، وليست حكمًا قضائيًا.
8. Consumer Financial Protection Bureau، 12 CFR §1002.9: Notifications، نسخة اللائحة الأمريكية الحالية لدى التحقق في 31 أكتوبر 2026؛ نص تنظيمي أصلي يحدد شروط إبلاغ أسباب القرار الائتماني السلبي، ضمن نطاق تطبيقه فقط.
9. Consumer Financial Protection Circular 2022-03، مايو/يونيو 2022، بيان تفسيري تاريخي؛ انظر سجل الهيئة الأصلي Withdrawn Guidance الذي يوثق سحبه وتعميم 2023-03 في 12 مايو 2025. لا يعرضان هنا بوصفهما توجيهين جاريين.
10. AI Risk Management Framework 1.0: Core, GOVERN، **المعهد الوطني الأمريكي للمعايير والتقنية (NIST)**، 2023؛ إطار طوعي لأدوار القيادة وإدارة مخاطر النماذج.

11. Article 14: Human Oversight، **المفوضية الأوروبية**، نص تشريعي بحسب النسخة المجمّعة 27 يوليو 2026؛ وللجدول الزمني المحدّث AI Act: Regulatory framework، إرشاد رسمي في 2026؛ يميز بين صدور القاعدة وبدء تطبيق مراحلها.

12. Key Principles of Governance in Financial Institutions Subject to Saudi Central Bank's Oversight and Supervision، **البنك المركزي السعودي**، 30 يونيو 2021، النسخة الثانية 8 مايو 2024، مصنفة «سارية» في سجل قواعده؛ مصدر تنظيمي أصلي للحوكمة المالية لا نص مخصوص بالتوقيع على قرارات الذكاء الاصطناعي.

13. Financial Stability Institute / BIS، Regulating AI in the financial sector: recent developments and main challenges، 12 ديسمبر 2024، دراسة مؤسسية مقارنة للمخاطر التنظيمية والتقنية؛ لا تقيس بنوكًا سعودية بعينها.

14. ICO، Explaining decisions made with AI **عبر** The Alan Turing Institute **و**GOV.UK، 27 يناير 2025؛ إرشاد عملي لتفسير القرار للمستفيد، لا تجربة ميدانية لقياس فعالية التفسيرات.

15. What are the accountability and governance implications of AI?، **مكتب مفوض المعلومات البريطاني (ICO)**، إرشاد رسمي متاح في تاريخ المراجعة؛ يشرح أدوار المتحكم والمعالج في **حماية البيانات**، وليس كل صور المسؤولية القانونية.