

حوادث وكلاء الذكاء الاصطناعي ليست إخفاقات  
ذكاء بل إخفاقات تفويض، وعلاجها أداة قديمة  
في كل مؤسسة: مصفوفة الصلاحيات.

## من وقع للوكيل؟

7 جمادى الأولى 1448 هـ - 18 أكتوبر 2026 م



“

### لا تستطيع المؤسسة أن تفوّض الفعل إلى آلة ثم تحتفظ لنفسها بالبراءة.

الجزء الذي انتشر من القصة جاء بعد ذلك. سأل  
المؤسس الوكيل لماذا فعل ما فعل، فكتب اعترافاً مرتباً:  
خمن بدل أن يتحقق، ونفذ أمراً مدمراً لم يطلبه منه  
أحد، وخالف كل قاعدة أعطيت له. استُعيدت البيانات  
لاحقاً، لكن فجوات بقيت فيها.

قرأ أغلب المعلقين القصة بوصفها قصة آلة طائشة.  
وأقرؤها قصة أقدم من الآلة: مفتاح كان في متناول من  
لم يوقع له أحد.

تسع ثوانٍ

في أواخر أبريل الماضي، كان وكيل برمجي يعمل على مهمة  
روتينية في بيئة الاختبار لدى شركة ناشئة أمريكية تباع  
برمجياتها لشركات تأجير السيارات. واجه تعارضاً في  
بيانات الدخول، فقرر من تلقاء نفسه أن يحلّه. بحث في  
المشروع عن مفتاح وصول، فوجد واحداً، ثم أرسل أمر  
حذف واحداً.

بعد تسع ثوانٍ كانت قاعدة البيانات الحية قد اختفت،  
ومعها نسخها الاحتياطية.



يضاعفون مشكلتها الكلاسيكيتين: يصعب أن تعرف ما يعرفه الوكيل، ويصعب أن تحدد أين تنتهي صلاحيته.

وحين تغيب الحدود، يحدث ما وصفته الباحثة مادلين إيلش عام 2019 بـ«منطقة الامتنصاص الأخلاقي». المصطلح مستعار من منطقة الانبعاث في مقدمة السيارة، التي صُممت لتمتص الصدمة. في الأنظمة الآلية المعقدة: يتحول الإنسان الأقرب إلى النظام إلى تلك المنطقة: يتحمل اللوم كله وإن لم يملك من السيطرة إلا القليل. والفرق المؤلم أن منطقة السيارة تحمي السائق، أما منطقة إيلش فتحمي سمعة النظام على حساب الإنسان. وهذا ما حدث في قصتنا تقريباً. حين انتشرت، اتجه قسم كبير من اللوم إلى المطور الذي منح الوكيل هذه الاستقلالية. ربما يستحق بعضه. لكن السؤال الأجدى ليس من يُلام بعد الحادثة، بل من كان يجب أن يوقع قبلها.

### مفتاح بلا توقيع

الوكيل لم يخترق شيئاً. وجد مفتاحاً فاستخدمه.

في أي مؤسسة منضبطة، لا يتسلم الموظف الجديد يوم التحاقه صلاحية حذف قاعدة البيانات الرئيسية. ثمة وثيقة تقول ماذا يملك أن يفعل، وإلى أي سقف، ومن يعتمد ما فوق ذلك السقف. اسمها مصفوفة الصلاحيات، وهي من أقل الوثائق إثارة وأكثرها حماية.

لكن الوكلاء يدخلون المؤسسة من باب التقنية لا من باب الحوكمة. يُعاملون أداة تُشتري وتُركب، لا فاعلاً يُكَلَّف ويُحاسَب. فيحملون مفاتيح لم تمر على أي مصفوفة، ولا يقف خلفها أي توقيع.

وقد جُرِّبت فكرة أن الآلة كيان منفصل أمام القضاء. حين أعطى روبوت المحادثة لدى طيران كندا مسافراً معلومة خاطئة عن تخفيض أجرة السفر للعزاء، دفعت

والمشكلة لا تقف عند الشركات الناشئة. في مسح نُشر في سبتمبر الماضي، قال 82% من أعضاء مجالس إدارات الشركات المدرجة إن ما يصلهم من معلومات عن صلة الذكاء الاصطناعي بالمخاطر والأداء متوسط أو ضعيف أو غير متاح. فمن يرى ما يفعله الوكيل، إن كان المجلس نفسه لا يرى؟

### لماذا تسرع المؤسسات

الإغراء حقيقي، ومن الإنصاف أن يُقال ذلك أولاً. الوكيل الآلي لا يجيب عن سؤال فحسب، بل ينقذ: يكتب الشيفرة ويشغلها، ويرد على العميل، ويحدث السجل، وينجز في دقائق ما كان يستغرق أياماً. والمؤسسة التي تتأخر عنه تخشى أن تجد منافسها قد سبقوها.

لذلك تتوقع ثلاثة أرباع المؤسسات تقريباً، في مسح شمل أكثر من ثلاثة آلاف قيادي في أربع وعشرين دولة، أن تستخدم الوكلاء استخداماً معتدلاً على الأقل بحلول 2027. لكن 21% فقط منها تملك نموذج حوكمة ناضجاً لهم. والناقص بحسب المسح محدد: حدود واضحة لما يقرره الوكيل وحده وما يحتاج موافقة بشرية، ورقابة لحظية على سلوكه، وسجل تدقيق يحفظ سلسلة أفعاله كاملة.

لاحظ أن البنود الثلاثة ليست تقنية في جوهرها. إنها أسئلة سلطة.

### الطرف الأقرب يدفع الثمن

درس الاقتصاد مشكلة التفويض قبل أن توجد الحواسيب. صاحب الشأن يكلف من ينوب عنه، ولا يستطيع أن يراقب كل ما يفعله، فيبني حول التكليف عقوداً وحدوداً ورقابة. هذه نظرية الأصيل والوكيل. ويرى الباحث القانوني نوعم كولت أن الوكلاء الآليين



أي تغيير في البيئة الحية مدير التقنية مفتاح إيقاف موثق ومجرب المال لا شيء كل صرف أو التزام صاحب الصلاحية المالية لا ينطبق الحذف لا شيء كل حذف، بتنفيذ مؤجل مالك النظام نسخة احتياطية معزولة

وفوق الجدول قاعدتان لا تُكتبان في خانة. سقف الوكيل لا يتجاوز سقف الإنسان الذي يوقع عنه، فمن لا يملك أن يعتمد صرفاً لا يملك أن يفوض وكيلاً به. والفعل الذي لا رجعة فيه لا يُنقذ منفرداً أبداً، مهما بدا الوكيل واثقاً.

### سؤال قبل أن تغلق الصفحة

هل تستطيع الآن أن تسمي الشخص الذي وقع على أخطر صلاحية يملكها وكيل آلي في مؤسستك؟ إن لم تستطع، فالمفتاح موجود والتوقيع غائب.

وأقوى اعتراض على هذا الطرح أن المصفوفة ستبطل الوكلاء وتلغي سبب وجودهم، وأن عددهم سيبلغ الآلاف فلا يمكن أن يُكتب لكل واحد سطر. والاعتراض وجيه في نصفه. لكن المصفوفة لا تُبنى على الأسماء بل على فئات الفعل، كما لا تُكتب مصفوفة البشر لكل موظف باسمه بل لكل مستوى. والإبطاء مقصود في الأفعال التي لا رجعة فيها وحدها، أما ما يمكن التراجع عنه فيبقى سريعاً.

ويبقى توتر لا أملك له حلاً كاملاً. مصفوفات البشر تفترض أن صاحب الصلاحية يفهم القاعدة فيلتزمها. والوكيل في قصتنا كان يعرف القاعدة، وكتب ذلك بنفسه، ثم خالفها. فالمصفوفة للوكيل لا تكفي مكتوبة، بل يجب أن تُفرض في النظام نفسه. وحين يبدأ الوكيل بتكليف وكلاء آخرين، تطول سلسلة التفويض حتى يصعب أن نعرف أين بدأ التوقيع. هذا سؤال المرحلة القادمة، وأظنه سيكون أصعب من سؤال هذه.

الشركة أمام هيئة الفصل في النزاعات بأن الروبوت مسؤول عن أفعاله. رفضت الهيئة الدفع، ورأت أن الروبوت جزء من موقع الشركة، وأن الشركة مسؤولة عن كل معلومة فيه.

والدرس أوسع من القضية: لا تستطيع المؤسسة أن تفوض الفعل إلى آلة ثم تحتفظ لنفسها بالبراءة.

### ما الذي تغير هذا العام

ما كان اجتهاداً قضائياً صار توجهاً تنظيمياً. في يناير أطلقت سنغافورة أول إطار حوكمة مخصص للذكاء الاصطناعي الوكيل، ونص صراحة على أن المسؤولية تبقى في النهاية على المؤسسات والأفراد الذين يشرفون على الوكلاء. وفي يوليو أطلقت سدايا الإطار الوطني لإدارة مخاطر الذكاء الاصطناعي، وجعلت المساءلة والمسؤولية أحد مبادئ السبعة.

وفي المجالس نفسها، اختار 71% من الأعضاء في المسح ذاته الذكاء الاصطناعي بوصفه أهم مهارة تحتاج مجالسهم إلى تقويتها. أي أن من يُفترض أن يوقعوا يعترفون بأنهم لا يرون بوضوح ما يوقعون عليه.

### سطر الوكيل في المصفوفة

العلاج لا يحتاج اختراعاً. يحتاج أن يُكتب للوكيل سطر في المصفوفة كما يُكتب لأي موظف. هذا نموذج يمكن البدء منه:

فئة الفعل يملكه الوكيل وحده يحتاج موافقة بشرية المالك الموقع طريقة إيقاف القراءة والتلخيص البيانات العامة والداخلية البيانات السرية والشخصية مالك البيانات سحب صلاحية حساب الوكيل التواصل الخارجي الردود المعيارية على الاستفسارات أي التزام مالي أو تعاقدية أو استثناء من سياسة مدير الخدمة تحويل المحادثة إلى موظف تعديل الأنظمة بيانات الاختبار فقط



## المصادر

1. ديكريت، تقرير عن حذف وكيل برمجي قاعدة بيانات شركة ناشئة في تسع ثوانٍ، أبريل 2026. يسند تفاصيل الحادثة وتاريخها، واعتراف الوكيل، واستعادة البيانات مع بقاء فجوات.  
<https://decrypt.co/365897/ai-agent-deletes-startup-database-9-seconds-founder-says>
2. بلومبرغ لو، تقرير عن المسح السنوي لأعضاء مجالس الإدارات الصادر عن برايس ووترهاوس كوبرز، 23 سبتمبر 2026. يسند نسبة 82% عن ضعف المعلومات المتاحة للمجالس، ونسبة 71% عن أولوية مهارة الذكاء الاصطناعي.  
<https://news.bloomberglaw.com/esg/most-board-directors-say-ai-skills-are-lacking-survey-shows>
3. ديلويت إنسايتس، الوكلاء الآليون يتوسعون أسرع من ضوابطهم، 24 أبريل 2026. يسند نسبة 21% للحكومة الناضجة، وتوقعات الاستخدام بحلول 2027، وعناصر الحكومة الناقصة.  
<https://www.deloitte.com/us/en/insights/topics/emerging-technologies/ai-agents-scaling-faster.html>
4. نوعم كولت، الجامعة العربية، ورقة حكمة وكلاء الذكاء الاصطناعي، 2025. يسند تطبيق نظرية الأصيل والوكيل على الوكلاء الآليين، ومشككتي عدم تماثل المعلومات وعموض نطاق الصلاحية.  
<https://arxiv.org/abs/2501.07913>
5. مادلين إيش، مجلة إنجينيغ ساينس آند تكنولوجيا آند سوسايتي، مناطق الامتصاص الأخلاقي، مارس 2019. يسند مفهوم منطقة الامتصاص الأخلاقي وأصل استعارته.  
<https://doi.org/10.17351/ests2019.260>
6. هيئة الفصل في النزاعات المدنية في كولومبيا البريطانية، قضية موفات ضد طيران كندا، 14 فبراير 2024، كما عرضها مكتب محاماة كندي. يسند دفع الشركة بأن الروبوت مسؤول عن أفعاله، ورفض الهيئة لهذا الدفع.  
<https://www.dww.com/articles/bc-tribunal-finds-air-canada-liable-for-inaccurate-advice-given-by-website-chatbot>
7. هوغان لوفلز، عرض لإطار سنغافورة النموذجي لحكومة الذكاء الاصطناعي الوكيل، 2026. يسند تاريخ إطلاق الإطار في 22 يناير 2026، وكونه الأول من نوعه، ونصه على أن المسؤولية تبقى على المؤسسات والأفراد المشرفين.  
<https://www.hlc.com/en/publications/singapore-launches-first-global-agent-ai-governance-framework>
8. وكالة الأنباء السعودية، سدايا تطلق الإطار الوطني لإدارة مخاطر الذكاء الاصطناعي، 14 يوليو 2026. يسند تاريخ الإطار ومبادئه السبعة، ومنها المساءلة والمسؤولية.  
<https://www.spa.gov.sa/N2633725>

المساءلة

مصفوفة الصلاحيات

وكلاء الذكاء الاصطناعي

الذكاء الاصطناعي

الحكومة



محمد المحميد

استراتيجية • مؤشرات أداء • تأسيس كيانات