



حوكمة الذكاء الاصطناعي الوكيل: تصوير وكلاء الذكاء الاصطناعي المستقلين وضبطهم

9 13 أغسطس 2027

لندن - فندق أربع نجوم في وسط لندن



حوكمة الذكاء الاصطناعي الوكيل: تصميم وكلاء الذكاء الاصطناعي المستقلين وضبطهم

المرجع: 5991_1147 التاريخ: 9 13 أغسطس 2027 المكان: لندن - فندق أربيع نجوم في وسط لندن الرسوم: SAR 23000

مقدمة

تنتقل المؤسسات من مساعدات الذكاء الاصطناعي التي تجيب عن الأسئلة إلى وكلاء ذكاء اصطناعي يخططون للهامر ويستدعون الأدوات وينفذون إجراءات على أنظمة العمل بصلاحيات مفوضة. ومن دون حدود واضحة قد يعتهد الوكيل دفعة هالية أو يكشف بيانات أو ينقل الأخطاء بين العمليات بسرعة تسبق انتباه أي شخص. تزود هذه الدورة من كور كونسبت المديرين بمنهجيات تصميم الوكلاء المستقلين وحوكمتهم وضبطهم وفق النظر الممتحدة لإدارة الذكاء الاصطناعي ومخاطره وأمنه. ويطبق المشاركون كل منهجية على حالات من قطاعات متعددة، ويخرجون بخطة حوكمة الوكيل وضوابطه لحالة استخدام في مؤسساتهم.

أهداف الدورة

- التمييز بين أنظمة الذكاء الاصطناعي الوكيل وأدوات الذكاء الاصطناعي التوليدي التقليدية، وتصنيف حالات الاستخدام المقترحة حسب درجة الاستقلالية وقابلية التراجع والنثر المحتمل
- اختيار بنى الوكلاء وأنماط الاستدلال وبروتوكولات التشغيل البيئي، ومنها Model Context Protocol، بما يناسب عملية العمل المستهدفة
- تحديد أهداف الوكيل وصلاحياته على الأدوات ونقاط الاعتقاد وشروط الإيقاف وفق مبدأ الحد الأدنى من الصلاحيات
- تقييم تهديدات الوكلاء ومخاطرهم باستخدام OWASP Top 10 for Agentic Applications وإطار CSA MAESTRO والمواصفة ISO/IEC 23894
- إرساء البات التقييم والمراقبة في أثناء التشغيل والتدخل التي تبقي الوكلاء ضمن حدود الصلاحيات المفوضة
- إعداد خطة حوكمة الوكيل وضوابطه بصيغة جاهزة للعرض على لجنة حوكمة الذكاء الاصطناعي

الفئة المستهدفة

- المديرون الهالكون لعمليات العمل التي تجرب فيها وكلاء الذكاء الاصطناعي أو يجري توسيع استخدامها
- هالكو منتجات وحلول الذكاء الاصطناعي المسؤولون عن تحديد سلوك الوكيل ومعايير قبوله
- مديرو البنية المؤسسية وبنية الحلول المسؤولون عن اعتماد ربط الوكلاء بالأنظمة الأساسية
- مديرو الحوكمة والمخاطر والامتثال في الذكاء الاصطناعي المشرفون على نشر الأنظمة المستقلة
- مديرو أمن المعلومات المسؤولون عن الهوية والوصول وتغطية التهديدات لوكلاء الذكاء الاصطناعي
- مديرو التحول الرقمي الذين يعدون دراسات الجدوى للنتيجة القائمة على الوكلاء

محاور الدورة

اليوم الأول: أساسيات الذكاء الاصطناعي الوكيل وبيئة حالات الاستخدام

- مكونات الوكيل: النموذج والمخطط والذاكرة والندوات وبيئة التشغيل
- طيف الاستقلالية من المساعد إلى الوكيل المستقل بالكامل
- تطبيق مفاهيم المواصفة ISO/IEC 22989 على أنظمة الوكلاء
- مصفوفة فرز حالات الاستخدام: القيمة وقابلية التراجع ونطاق الضرر
- حصر الوكلاء الحاليين واكتشاف الوكلاء غير المعتمدين

اليوم الثاني: البنى والبروتوكولات ومعايير الإدارة

- نهج الاستدلال ReAct والتخطيط ثم التنفيذ Plan-and-Execute
- مبادئ الوكيل المنفرد والمنسق والعمال والوكلاء المتعددين
- بروتوكولات Model Context Protocol وAgent2Agent للتشغيل البيئي بين الندوات والوكلاء
- ضوابط نظام إدارة الذكاء الاصطناعي وفق ISO/IEC 42001 لنشر الوكلاء
- وظائف الحوكمة والتحديد والقياس والإدارة في NIST AI RMF وتطبيقها على الوكلاء

اليوم الثالث: تصميم وكلاء خاضعين للضبط

- نموذج توصيف الوكيل: الأهداف والصلاحيات والندوات وشروط الإيقاف
- الوصول إلى الندوات بالحد الأدنى من الصلاحيات وتصميم بيانات اعتماد محددة النطاق
- نقاط الاعتقاد البشري وحدود التصعيد Human-in-the-Loop
- طبقات الحواجز الوقائية: مرشحات المدخلات والتحقق من المخرجات ومحررات السياسات
- حزمة اختبار تقييم الوكيل: نجاح المهمة ومسار التنفيذ ومقاييس السلامة

اليوم الرابع: تهديدات الوكلاء وأثرهم وضبطهم في أثناء التشغيل

- قائمة OWASP Top 10 for Agentic Applications: اختطاف الهدف وإساءة استخدام الندوات وتسليم الذاكرة
- نهج التهديدات متعددة الطبقات للنظم متعددة الوكلاء بإطار CSA MAESTRO
- حقن الأنوامر غير المباشر عبر المحتوى المسترجع ومخرجات الندوات
- تقييم أثر القرارات المستقلة وفق ISO/IEC 42005
- المراقبة في أثناء التشغيل وسجلات التدقيق وإجراءات الإيقاف الطارئ

اليوم الخامس: دراسات الحالة وخطة حوكمة الوكيل وضوابطه

- دراسة حالة وكيل المشتريات: حدود الاعتماد والوصول إلى بيانات الموردين
- دراسة حالة وكيل خدمة العملاء: صلاحية رد المبالغ والتحويل إلى الموظف
- بناء سجل مخاطر الوكيل المتوافق مع ISO/IEC 23894
- إعداد مسودة خطة حوكمة الوكيل وضوابطه
- مجلس مراجعة التصميم والدفاع عن الخطة

المهارات المكتسبة

- تصميم بنية الوكلاء
- تصنيف مخاطر الاستقلالية
- تصميم صلاحيات الندوات
- نهضة تهديدات الذكاء الاصطناعي الوكيل
- تصميم الإشراف البشري
- تقييم الوكلاء واختبارهم
- ضوابط نظام إدارة الذكاء الاصطناعي
- مراقبة الوكلاء في أثناء التشغيل

لهذا تحضر هذه الدورة

- تعود بخطة حوكمة وضوابط لحالة استخدام فعلية، بعد أن راجعها الزملاء بصفتهم مجلس مراجعة التصميم
- توجه الفرق التقنية والموردين بدقة إلى ما يسمح للوكيل بفعله، والندوات التي يستدعيها، ومتى يجب أن يتخذ الإنسان القرار
- تتعرف على أنماط الإخفاق التي تحول تجربة ناجحة إلى وكيل خارج السيطرة في بيئة الإنتاج
- تقارن مناهج حوكمة الوكلاء مع مديريين من قطاعات المال والحكومة والرعاية الصحية والصناعة

الخاتمة

لا يحقق الذكاء الاصطناعي الوكيل قيمته إلا حين يعمل كل وكيل ضمن صلاحيات صممت واختبرت وروقت بعناية. تنتقل هذه الدورة من مكونات الوكلاء ومستويات استقلالياتهم، مروراً بالبنية والبروتوكولات ومعايير الإدارة التي تشكلهم، وصولاً إلى التهديدات وضوابط التشغيل التي تقيهم ضمن الحدود. ويحول اليوم الأخير هذه المعارف إلى خطة حوكمة الوكيل وضوابطه، فيمتلك المشاركون أساساً منظماً للاعتماد الوكيل التالي الذي تقترح مؤسستهم نشره وتقيده والإشراف عليه.