



حوكمة الذكاء الاصطناعي الوكيل: تصميم وكلاء الذكاء الاصطناعي المستقلين وضبطهم

19 - 23 Jul 2027

دبي - فندق خمس نجوم في وسط النعمال



حوكمة الذكاء الاصطناعي الوكيل: تصميم وكلاء الذكاء الاصطناعي المستقلين وضبطهم

المرجع: 1147_5988 التاريخ: 19 - 23 Jul 2027 الموقع: دبي - فندق خمس نجوم في وسط النعمال الرسوم: SAR 19500

العوق التقني: مهارس - أسلوب التطبيق: دراسة حالة

مقدمة

تنتقل المؤسسات من مساعدات الذكاء الاصطناعي التي تجيب عن الأسئلة إلى وكلاء ذكاء اصطناعي يخططون للمهام ويستدعون الأدوات وينفذون إجراءات على أنظمة العمل بصلاحيات مفوضة. ومن دون حدود واضحة قد يعتهد الوكيل دفعة مالية أو يكشف بيانات أو ينقل الأخطاء بين العمليات بسرعة تسبق انتباه أي شخص. تزود هذه الدورة من كور كونسبت المديرين بتمهيجات تصميم الوكلاء المستقلين وحوكمتهم وضبطهم وفق النظر المتهمة 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 والثر المحتمل
- اختيار بنى الوكلاء وأنماط الاستدلال وبروتوكولات التشغيل البيئي، ومنها Protocol Context Model، بها 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 Applications Agentic for 10 Top OWASP وإطار MAESTRO CSA والمواصفة ISO/IEC 23894
- إرساء اليات 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 المالكون لعمليات العمل التي تجرب فيها وكلاء الذكاء الاصطناعي أو يجري توسيع استخدامها
- مالكو 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 القانئة على الوكلاء

محاور الدورة

اليوم الأول: أساسيات الذكاء الاصطناعي الوكيل وبيئة حالات الاستخدام

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

اليوم الثاني: البنى والبروتوكولات ومعايير الإدارة

- نهج الاستدلال ReAct والتخطيط ثم التنفيذ Execute-and-Plan
- هيكل الوكيل المنفرد والمنسق والعمال والوكلاء المتعددين
- بروتوكول Agent2Agent و Protocol Context Model للتشغيل البيئي بين الندوات والوكلاء
- ضوابط نظام إدارة الذكاء الاصطناعي وفق IEC/ISO 42001 لنشر الوكلاء
- وظائف الحوكمة والتحديد والقياس والإدارة في RMF AI NIST 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 الاهتمام البشري وحدود التصعيد Loop-the-in-Human
- طبقات الحواجز الوقائية: مرشحات المدخلات والتحقق من المخرجات ومحركات السياسات
- حزمة اختبار تقييم الوكيل: نجاح المهمة ومسار التنفيذ ومقاييس السلامة

اليوم الرابع: تهديدات الوكلاء وأثرهم وضبطهم في أثناء التشغيل

- قائمة Applications Agentic for 10 Top OWASP: اختطاف الهدف وإساءة استخدام الندوات وتصميم الذاكرة
- نهج التهديدات متعددة الطبقات للنظم متعددة الوكلاء بإطار MAESTRO CSA
- حقن الأنوامر غير المباشر عبر المستوى المسترجع ومخرجات الندوات
- تقييم أثر القرارات المستقلة وفق IEC/ISO 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 IEC/ISO 23894
- إعداد مسودة خطة حوكمة الوكيل وضوابطه
- مجلس مراجعة التصميم والدفاع عن الخطة

المهارات المكتسبة

- تصميم بنية الوكلاء
- تصنيف مخاطر الاستقلالية
- تصميم صلاحيات الندوات
- نهج 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 ب خطة 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، وحتى 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 مديرين 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 استقلاليته، مروراً بالبنى والبروتوكولات ومعايير الإدارة التي تشكلهم، وصولاً إلى التهديدات وضوابط التشغيل التي تبقيهم ضمن الحدود. ويحول اليوم النخيل هذه المعارف إلى خطة حوكمة الوكيل وضوابطه، فيمتلك المشاركون أساساً منظماً للاعتماد الوكيل التالي الذي تقترح مؤسستهم نشره وتقييمه والإشراف عليه.