

دورة هندسة أمن وكلاء الذكاء الاصطناعي وضمانه

21 – 25 مارس 2027

جنيف

AGILE LEADERS
Training Center

دورة هندسة أمن وكلاء الذكاء الاصطناعي وضمانه

المرجع: 199_118049 **التاريخ:** 21 – 25 مارس 2027 **المكان:** جنيف **الرسوم:** Euro 6200

نظرة عامة

دورة هندسة أمن وكلاء الذكاء الاصطناعي وضمانه دورة متقدمة لمدة أسبوع لمهندسي أمن الذكاء الاصطناعي ومتخصصي أمن التطبيقات ومطوري الوكلاء والمهندسين وفرق الاختبار والمخاطر التقنية، ويعود المشاركون منها بخطة ضمان أمن وكلاء الذكاء الاصطناعي. يفحص المشاركون نمذجة التهديدات وحقن الأوامر وإساءة استخدام الأدوات والهوية وأقل الصلاحيات وتسميم الذاكرة والاسترجاع والثقة بين الوكلاء وتسرب البيانات والعزل والمراقبة والاستجابة والاختبار الهجومي. تقدم أجايل للتدريب دورة في ضمان أمن وكلاء الذكاء الاصطناعي.

الفئة المستهدفة

- فرق أمن الذكاء الاصطناعي المسؤولة عن نماذج تهديد الوكلاء وضوابطهم وأدلة الضمان
- فرق أمن التطبيقات المسؤولة عن مراجعة التصميم والاختبار وبوابات الإصدار
- فرق هندسة الوكلاء المسؤولة عن الأدوات والذاكرة والاسترجاع والتنسيق والنشر
- فرق الاختبار الأمني المسؤولة عن التمارين الهجومية واكتشاف الإخفاقات
- فرق مخاطر التقنية المسؤولة عن ملكية الضوابط والمراقبة وجاهزية الحوادث

تفترض الدورة قدرة المشاركين على تقييم بنية التطبيقات والضوابط الأمنية، ولا تتناول الثقافة التمهيدية للذكاء الاصطناعي أو بناء الوكلاء العام أو إدارة منصات الموردين.

الإدارات والقطاعات المستهدفة

تدعم الدورة أمن الذكاء الاصطناعي في البرمجيات والتمويل والرعاية الصحية والتجزئة والخدمات اللوجستية والخدمات العامة والمهنية.

- أمن التطبيقات وأمن المنتجات
- هندسة الذكاء الاصطناعي وهندسة الحلول
- الهوية والوصول وحماية البيانات
- عمليات الأمن والاستجابة للحوادث
- مخاطر التقنية في المؤسسات المنظمة ومقدمي الخدمات الرقمية

AGILE LEADERS
Training Center

أهداف التعلم

بنهاية هذه الدورة سيكون المشاركون قادرين على:

- تحليل أصول الوكيل وحدود الثقة ومسارات الهجوم
- تقييم حقن الأوامر وإساءة الأدوات ومخاطر التسميم
- تطبيق الهوية وأقل الصلاحيات والعزل وضوابط الموافقة
- تشخيص إشارات المراقبة وحوادث أمن الوكلاء
- بناء اختبارات هجومية وبوابات للتقييم الأمني
- بناء خطة ضمان أمن وكلاء الذكاء الاصطناعي

جدول الدورة

اليوم 1: نمذجة تهديدات الوكلاء

- جرد أصول الوكيل وإجراءاته
- مخطط حدود الثقة وتدفق البيانات
- شجرة مسارات الهجوم على الوكيل
- مصفوفة تصنيف الاستقلالية والأثر
- سجل تتبع المتطلبات الأمنية

اليوم 2: الحقن والأدوات وحدود البيانات

- خريطة اختبار حقن الأوامر المباشر وغير المباشر
- نمط عزل المحتوى غير الموثوق
- سيناريو إساءة الأدوات وتسلسل الإجراءات
- شبكة ضوابط تسرب البيانات واستخراجها
- حدود التأكيد البشري والمعاملة

AGILE LEADERS
Training Center

اليوم 3: الهوية والذاكرة والثقة بين الوكلاء

- بطاقة هوية الوكيل المخصصة وملكيتها
- مصفوفة أقل الصلاحيات للأدوات والموارد
- مسار الوصول المحدد زمنيا والإلغاء
- مجموعة اختبار تسميم الاسترجاع والذاكرة
- بروتوكول الثقة في التسليم بين الوكلاء

اليوم 4: المراقبة والاختبار والاستجابة

- سجل تدقيق إجراءات الوكيل والتفويض
- لوحة الإشارات الأمنية وكشف الشذوذ
- مكتبة حالات الاختبار الهجومي للوكلاء
- دليل الاحتواء ومفتاح الإيقاف
- دليل فرز حوادث الوكلاء والتعافي

اليوم 5: ممارسة الضمان الأمني والمشروع الختامي

- تمرين مقترح: بناء نموذج تهديد لوكيل
- تمرين مقترح: اختبار الحقن وإساءة الأدوات
- تمرين مقترح: تحدي ضوابط الهوية والذاكرة
- تمرين مقترح: محاكاة الكشف والاحتواء
- المشروع الختامي: خطة ضمان أمن وكلاء الذكاء الاصطناعي

التمارين التطبيقية

تستخدم الدورة أنشطة مقترحة تحول بنية الوكيل إلى ضوابط أمنية وأدلة ضمان قابلة للاختبار.

- نشاط مقترح: رسم الأصول وتدفقات البيانات وحدود الثقة والإجراءات والهويات والأدوات ومخازن الذاكرة ومستويات الأثر
- نشاط مقترح: تشغيل سيناريوهات الحقن وإساءة الأدوات وتسرب البيانات وتسميم الاسترجاع والذاكرة مقابل ضوابط محددة
- نشاط مقترح: اختبار أقل الصلاحيات وبوابات الموافقة وحدود العزل وسجلات التدقيق والإلغاء ومسارات الإيقاف
- نشاط مقترح: تجميع النتائج والمالكين وأولويات المعالجة واختبارات الانحدار وإشارات المراقبة وإجراءات الاستجابة

AGILE LEADERS
Training Center

الأسئلة الشائعة

لمن تناسب دورة هندسة أمن وكلاء الذكاء الاصطناعي وضمانه، وماذا تفترض؟

تناسب الدورة مهندسي الأمن ومتخصصي أمن التطبيقات ومطوري الوكلاء والمهندسين وفرق الاختبار والمخاطر. وتفترض خبرة في تقييم بنية التطبيقات والضوابط.

كيف يختلف تدريب أمن وكلاء الذكاء الاصطناعي عن تدريب تطوير الوكلاء العام؟

يركز تدريب أمن الوكلاء على مسارات الهجوم والهويات والصلاحيات والمحتوى غير الموثوق والتسميم والتسرب والاختبارات الهجومية والمراقبة والاستجابة بدلا من بناء خصائص الوكيل العامة.

ما التهديدات التي ينبغي أن يفحصها تقييم أمن وكلاء الذكاء الاصطناعي؟

يفحص التقييم حقن الأوامر واستدعاءات الأدوات غير الآمنة والصلاحيات الزائدة وغموض الهوية وتسرب البيانات وتسميم الاسترجاع والذاكرة والتسليم غير الآمن وضعف سجلات التدقيق والاستقلالية غير المنضبطة وتأخر الاحتواء.

كيف تختبر الفرق أمن وكلاء الذكاء الاصطناعي قبل النشر؟

تختبر الفرق المهام التمثيلية والهجومية والحقن المباشر وغير المباشر وسياسات الأدوات وحدود الصلاحيات وبوابات الموافقة ومعالجة البيانات وكتابة الذاكرة ورسائل الوكلاء والتسجيل والإلغاء والاحتواء والتعافي.

ما الأدلة التي تدعم ضمان أمن وكلاء الذكاء الاصطناعي؟

تشمل الأدلة نموذج التهديد وتتبع الضوابط وسجلات الهوية والصلاحيات ونتائج الاختبارات الهجومية وحالات الانحدار وسجلات الإجراءات وقواعد المراقبة وملكية المعالجة وأدلة الحوادث وقرارات الإصدار.

الخاتمة

يعود المشاركون بخطة ضمان أمن وكلاء الذكاء الاصطناعي التي تربط التهديدات وحدود الثقة والضوابط والاختبارات والمراقبة وإجراءات الاستجابة. تغير الخطة طريقة تقييم أنظمة الوكلاء وإصدارها. وتدعم القرارات الأمنية القابلة للتتبع وملكية الضوابط والتحقق الهجومي والاحتواء السريع واختبار الانحدار بعد التغييرات المؤثرة.