JUPITER PRIME
Consulting • Training • Advisory
جوبيتر برايم للاستشارات والتدريب

كتيّب الدورة

دورة متقدمة في تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه

12 – 16 سبتمبر 2027

المنامة

info@jupiterprime.uk

jupiterprime.uk

مسجّلة في المملكة المتحدة والمملكة العربية السعودية

دورة متقدمة في تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه

المرجع: 37017_977 **التاريخ:** 12 – 16 سبتمبر 2027 **المكان:** المنامة **الرسوم:** GBP 5500

تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه دورة تدريبية مدتها أسبوع لمسؤولي الامتثال ومديري المخاطر وقادة الإشراف على الذكاء الاصطناعي، تتمحور حول مخرج واحد هو ملف أثر الذكاء الاصطناعي وحوادثه. فالأنظمة الآلية تقرر اليوم في الائتمان والمطالبات والتوظيف والوصول إلى الخدمات، وعلى الجهات أن تثبت أنها وزنت الضرر قبل الإطلاق وتصرفت بانضباط عند وقوع الخلل. تقدم جوبيتر برايم للتدريب هذه الدورة في تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه.

لماذا يهم تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه الآن

تضغط ثلاث فئات على هذا المجال اليوم. فالجهات الرقابية والمدققون يطلبون دليلا مكتوبا على اختبار العدالة والشفافية، والعملاء والموظفون يعترضون على النتائج الآلية أكثر من قبل، ومجالس الإدارة تريد معرفة الأنظمة القادرة على إحداث ضرر جسيم. الجهة التي تحسن ذلك تطلق أنظمتها بمساءلة واضحة وتحتوي الأعطال بسرعة، أما التي تسيء فتواجه الشكاوى وتعليق الخدمات والإضرار بالسمعة. وبين هذين الطرفين تتحدد ثقة المتعاملين بالقرارات الآلية. وتستند الدورة إلى ISO/IEC 42001 نظام إدارة الذكاء الاصطناعي وإطار NIST AI RMF إطار إدارة مخاطر الذكاء الاصطناعي.

المستفيدون من تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه

- المسؤولون عن الموافقة على أنظمة القرار الآلي الجديدة من جهة الامتثال
- المسؤولون عن تصنيف درجة تعرض الأنظمة وصيانة مكتبات الضوابط
- المسؤولون عن رئاسة لجان المراجعة الأخلاقية للذكاء الاصطناعي
- المسؤولون عن اختبار الأدلة التي تسند ضوابط الذكاء الاصطناعي في التدقيق الداخلي
- المسؤولون عن تنسيق الاستجابة للحوادث ضمن المرونة التشغيلية

يستفيد أكثر خط الدفاع الثاني حين يرسل معا أفرادا من الامتثال والمخاطر والتدقيق.

قدرات تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه

يحول المشاركون المبادئ الأخلاقية إلى تقييمات متدرجة ومسارات أدلة وأدلة استجابة.

- تصنيف أنظمة الذكاء الاصطناعي في فئات خطورة وفق الغرض المقصود والفئات المتأثرة — لجنة الموافقة
- تصميم نموذج تقييم الأثر الخوارزمي مع اختبارات العدالة والشفافية — مالك النظام
- تحديد سجلات المساءلة التي تسمي أصحاب القرار وصلاحيات التجاوز — لجنة المخاطر في مجلس الإدارة
- اختبار أدلة التدقيق مقابل أهداف ضوابط ISO/IEC 42001 — المدقق الداخلي
- بناء سلم شدة حوادث الذكاء الاصطناعي ومسار التصعيد — قائد المرونة التشغيلية
- إبلاغ نتائج ما بعد الحادث للمستخدمين المتأثرين والقيادات العليا — كبير مسؤولي الامتثال

مصطلحات تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه الأساسية

- تقييم الأثر الخوارزمي — مراجعة منظمة لمن قد يتضرر من نظام آلي وكيف، وأي ضمانات تحد من ذلك.
- فئة الخطورة — تصنيف يحدد عمق المراجعة التي يخضع لها النظام بحسب الضرر المحتمل منه.
- تفاوت النتائج — فرق قابل للقياس في النتائج بين مجموعات من الأشخاص يتأثرون بالقرار الآلي نفسه.
- بطاقة النموذج — وثيقة قصيرة تصف غرض النموذج وبيانات تدريبه وحدود أدائه ومستخدميه المقصودين.
- حادثة الذكاء الاصطناعي — واقعة يسبب فيها نظام ذكي ضررا أو معاملة غير عادلة أو خطأ جسيما أو يكاد.
- مراجعة السبب الجذري — تحقيق يتتبع الحادثة إلى البيانات أو النموذج أو الإجراء أو القرار البشري الذي يقف خلفها.

تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه يوما بيوم

اليوم 1: المبادئ الأخلاقية والمساءلة

- مبادئ العدالة والشفافية والمساءلة — تحويل القيم إلى اختبارات
- هيكل نظام الإدارة في ISO/IEC 42001 — موضع التقييمات ضمن الضوابط
- وظائف Govern Mapg Measureg Manageg في NIST AI RMF — إسناد الملكية
- سجلات المساءلة وصلاحيات التجاوز — تسمية أصحاب القرار
- ميثاق لجنة المراجعة الأخلاقية — صلاحيات التصعيد والاعتراض

اليوم 2: تصنيف الأثر وتقييمه

- تصنيف فئات الخطورة حسب الغرض والفئات المتأثرة — عمق المراجعة
- نموذج تقييم الأثر الخوارزمي — أسئلة ما قبل الإطلاق
- اختبار تفاوت النتائج — قبول النموذج أو رفضه
- متطلبات قابلية التفسير حسب الفئة — ما يجب إبلاغه للمستخدمين
- سجلات استشارة أصحاب المصلحة — إدراج آراء المتأثرين

اليوم 3: بناء أدلة التدقيق

- بطاقات النماذج وصحائف البيانات — توثيق الحدود والاستخدام المقصود
- اختبار الضوابط مقابل أهداف ISO/IEC 42001 — كفاية الأدلة
- سجلات المراقبة وتنبيهات الانحراف — إثبات الإشراف المستمر
- العناية الواجبة لنماذج الأطراف الخارجية — طلبات الأدلة من الموردين
- تقارير التأكيد للجان مجلس الإدارة — بيانات التعرض المتبقي

اليوم 4: معالجة حوادث الذكاء الاصطناعي

- سلم شدة حوادث الذكاء الاصطناعي — تفعيل التصعيد والإيقاف
- خيارات احتواء النماذج المعطوبة — التراجع أو الإبطاء أو التحول إلى العنصر البشري
- منهج مراجعة السبب الجذري — البيانات أو النموذج أو الإجراء أو القرار
- صياغة إشعار المستخدمين المتأثرين — مسارات الإنصاف والتظلم
- سجل الدروس وتحديث الضوابط — منع تكرار الأعطال

اليوم 5: ممارسة ملف أثر الذكاء الاصطناعي وحوادثه

- حالة نموذج فرز مطالبات التأمين — تطبيق تصنيف فئات الخطورة
- حالة كاشف احتيال تذاكر النقل العام — تطبيق اختبارات تفاوت النتائج
- محاكاة عطل نموذج حدود الائتمان — تطبيق خطوات الاحتواء والسبب الجذري
- جولة على أدلة التدقيق — تطبيق اختبارات ضوابط ISO/IEC 42001
- استكمال ملف أثر الذكاء الاصطناعي وحوادثه — عرض الفئات والأدلة وأدلة الاستجابة

تطبيق تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه بعد الدورة

- تصنيف كل نظام ذكاء اصطناعي عامل في وحدة أعمال واحدة إلى فئات خطورة وحصر تقييمات الأثر الناقصة.
- تنفيذ تمرين مكتبي على سلم شدة الحوادث مع فرق العمليات والاتصال.

أسئلة حول تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه

ما تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه ولماذا يهم الآن؟

هو ممارسة تقدير الضرر الذي قد يسببه نظام ذكي قبل إطلاقه، والاستجابة بانضباط حين يتعطل. ويهم الآن لأن المدققين والجهات الرقابية والمتأثرين يتوقعون دليلا مكتوبا على الأمرين معا.

من الأدوار الأكثر استفادة من تدريب تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه؟

يستفيد أكثر مسؤولو الامتثال ومديرو المخاطر وقادة الإشراف على الذكاء الاصطناعي، ومعهم المدققون الداخليون وفرق المرونة التشغيلية.

كيف يختلف تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه عن دورة استراتيجية حوكمة الذكاء الاصطناعي؟

يعمل على مستوى الأنظمة المفردة لا على مستوى سياسة المؤسسة. فدورة الاستراتيجية تضع المبادئ والهياكل، أما هنا فيصنف المشاركون الأنظمة ويختبرون النتائج ويجمعون أدلة التدقيق ويتدربون على الاستجابة للحوادث.

ماذا تعود به من تدريب تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه؟

تعود بملف مكتمل لأثر الذكاء الاصطناعي وحوادثه. يضم فئات الخطورة ونموذج التقييم وسجلات المساءلة وقائمة الأدلة ودليل الاستجابة للحوادث، جاهزة للتطبيق على أنظمتك.

تقييم أثر الذكاء الاصطناعي ومعالجة حوادثه: الطريق إلى الأمام

يحدد ملف أثر الذكاء الاصطناعي وحوادثه أولا الأنظمة التي تحتاج تقييما كاملا قبل إصدارها التالي. ويستخدمه قادة الامتثال والمخاطر والتدقيق لاعتماد الإطلاقات وإحاطة لجان مجلس الإدارة. وبعد أول حادثة فعلية أو حادثة وشيكة يستحق سلم الشدة وخيارات الاحتواء مراجعة جديدة، إذ تكشف أزمنة الاستجابة الفعلية أين يجب أن يتسارع التصعيد في المرات القادمة.