



إدارة البيانات الرئيسية وجودة البيانات وإدارة البيانات الوصفية



إدارة البيانات الرئيسية وجودة البيانات وإدارة البيانات الوصفية

العمق التقني: ممارس · أسلوب التطبيق: مختبر تطبيقي

مقدمة

تكلف سجلات العملاء المكررة ورموز المنتجات غير المتسقة والحقول غير المعرفة المؤسسات أموالاً كل يوم، في صورة شحنات متعثرة وفواتير خاطئة وتقارير غير موثوقة ومشاريع ترحيل أنظمة متوقفة. وتبدأ معظم هذه المشكلات في بيانات رئيسية لا يقيسها أحد ولا يعرفها ولا يملكها. تزود هذه الدورة من كور كونسبت المختصين بالقدرة على فحص البيانات وقياس جودتها وفق المعايير الدولية وتوثيقها ببيانات وصفية قابلة للاستخدام وبناء سجلات ذهبية موثوقة بقواعد المطابقة والبقاء، ويعمل المشاركون عملياً على بيانات من قطاعات متعددة، ويخرجون بخطة تحسين جودة البيانات والبيانات الرئيسية لمجال بيانات حرج.

أهداف الدورة

- فحص بيانات المصدر باستخدام OpenRefine SQL لقياس مشكلات الاكتمال والصلاحية والتفرد والاتساق
- تحديد قواعد وعتبات قابلة للقياس لجودة البيانات استناداً إلى نموذج ISO/IEC 25012 ومقاييس ISO/IEC 25024
- توثيق عناصر البيانات الحرجة في مسرد لمصطلحات الأعمال وقاموس بيانات يتبعان مبادئ التسمية في ISO/IEC 11179
- ضبط قواعد المطابقة الحتمية والاحتمالية وقواعد البقاء التي تنتج سجلاً ذهبياً موثقاً
- تتبع عيوب البيانات المتكررة إلى أسبابها الجذرية وتصميم مسارات عمل للأمناء البيانات تمنع تكرارها
- إعداد خطة تحسين جودة البيانات والبيانات الرئيسية لمجال بيانات حرج

الفئة المستهدفة

- أمناء البيانات المسؤولون عن سجلات العملاء أو الموردين أو المنتجات أو الأصول
- محللو البيانات والتقارير الذين يتحققون من أخطاء البيانات ويصححونها
- مسؤولو البيانات الرئيسية العاملون في بيئات ERP و CRM
- مختصو العمليات المسؤولون عن معايير إدخال البيانات وضوابطها
- منسقو فهارس البيانات والبيانات الوصفية المسؤولون عن توثيق التعريفات
- أعضاء فرق ترحيل الأنظمة الذين يجهزون البيانات للمنصات الجديدة

محاوَر الدورة

اليوم الأول: أساسيات جودة البيانات والبيانات الوصفية والبيانات الرئيسية

- تكلفة ضعف جودة البيانات: قاعدة 1-10-100 وحالات الأثر
- مصفوفة التمييز بين البيانات الرئيسية والمرجعية وبيانات المعاملات
- أبعاد جودة البيانات وتعريفاتها في DAMA-DMBOK2
- ورقة عمل تحديد عناصر البيانات الحرجة
- حصر مشكلات البيانات الحالية واستبيان أصحاب المصلحة

اليوم الثاني: معايير ونماذج الجودة والبيانات الوصفية والبيانات الرئيسية

- نموذج جودة البيانات ISO/IEC 25012: الخصائص الذاتية والخصائص المعتمدة على النظام
- مقاييس جودة البيانات في ISO/IEC 25024 والعتبات المستهدفة
- النموذج المرجعي لعمليات إدارة جودة البيانات ISO 8000-61
- مبادئ سجلات البيانات الوصفية وتسمية عناصر البيانات وفق ISO/IEC 11179
- أنماط تطبيق إدارة البيانات الرئيسية: السجل والتوحيد والتعايش والمركز الموحد

اليوم الثالث: الفحص والقواعد والبيانات الوصفية في المختبر

- مختبر فحص البيانات باستخدام OpenRefine SQL
- كتابة قواعد جودة البيانات والتحقق منها باستخدام Great Expectations
- بناء مسرد مصطلحات الأعمال وقاموس البيانات بعناصر Dublin Core
- رسم مسار البيانات من المصدر إلى التقرير
- إعداد بطاقة قياس جودة البيانات ولوحة معلومات الاتجاهات

اليوم الرابع: المطابقة والسجل الذهبي ومنع العيوب

- مطابقة السجلات الحتمية والاحتمالية بنموذج Fellegi-Sunter
- قواعد البقاء وتصميم السجل الذهبي
- متطلبات مواصفة البيانات الرئيسية وتبادلها وفق ISO 8000-110
- تحليل الأسباب الجذرية لعيوب البيانات بمنهجية DMAIC
- تصميم مسار عمل أمين البيانات وقائمة الاستثناءات

اليوم الخامس: التطبيق في المختبر وخطة التحسين

- مختبر إزالة تكرار سجلات العملاء من مصادر متعددة
- مختبر تنقية بيانات الموردين والمواد
- إعداد مدخلات فهرس البيانات الوصفية لمجال بيانات حرج
- إعداد مسودة خطة تحسين جودة البيانات والبيانات الرئيسية
- عرض الخطة ومراجعتها بين الزملاء

المهارات المكتسبة

- فحص البيانات
- قياس جودة البيانات
- تصميم قواعد جودة البيانات
- توثيق البيانات الوصفية
- مطابقة السجلات وإزالة التكرار
- إدارة السجل الذهبي
- تحليل الأسباب الجذرية لعيوب البيانات
- تصميم مسارات عمل أمناء البيانات

لماذا تحضر هذه الدورة

- تعود إلى عملك بخطة تحسين جودة البيانات والبيانات الرئيسية لمجال بيانات حرج، وقد راجعها الزملاء
- تعرض على ملاك الأعمال حجم مشكلة البيانات بأرقام يمكنهم ترتيب أولوياتها، لا روايات متفرقة
- تعالج السجلات المكررة والمتعارضة بقواعد قابلة للتكرار بدلا من حملات تنظيف يدوية لمرة واحدة
- تقارن ممارسات أمانة البيانات لديك بمختصين من قطاعات ودول مختلفة يواجهون مشكلات مماثلة

الخاتمة

تعتمد التقارير الموثوقة ومشاريع الترحيل السلسة والأتمتة المستقرة كلها على بيانات رئيسية مقيسة ومعرفة ولها مالك. تنتقل هذه الدورة من أساسيات جودة البيانات والبيانات الوصفية والبيانات الرئيسية، مروراً بالمعايير الدولية وأنماط إدارة البيانات الرئيسية التي تنظمها، ثم الفحص وكتابة القواعد والتوثيق في المختبر، وصولاً إلى المطابقة والسجل الذهبي ومنع العيوب. يطبق اليوم الأخير هذه الأساليب على بيانات واقعية لينتج خطة تحسين جودة البيانات والبيانات الرئيسية يعود بها المشارك إلى مؤسسته.