

Sentetik/Anonim Test Dataset'i Tasarım Checklist'i — Trend (v1.0)

Asset Amaç: Bu checklist, otel projelerinde test/sandbox ortamlarında gerçek misafir verisini tamamen devre dışı bırakmak için sentetik veya anonim dataset tasarımını standardize eder. Amaç; PII leak'i sıfırlamak, gerçekçi senaryoları korumak, test verisi üretim/temizleme job'larını kanıtlanabilir loglarla raporlamak ve KVKK denetim paketine "no real PII" kanıtını eklemektir.

Kim Kullanır?: IT + yazılım ekibi, ajans, veri/BI ekibi, GM (policy onayı).

Nasıl Kullanılır?

- Data Dictionary Hazırlayın:** Sistemlerdeki tüm alanları sınıflandırarak PII barındıran kritik verileri belirleyin.
- Hibrit Sentetik Dataset Oluşturun:** Kimlik ve iletişim bilgilerini tamamen kurgusal verilerle değiştirirken operasyonel mantığı (rezervasyon tarihleri, oda tipleri) gerçekçi tutun.
- Sızıntı (Leak) Testi ve Job Loglarını Raporlayın:** Otomatik PII tarama testlerini çalıştırıp temizleme/refresh loglarını KVKK Audit Pack'e ekleyin.

A) DATA DICTIONARY CHECKLIST'İ

- [] PMS Alanları Sınıflandırıldı:** Kimlik, iletişim, rezervasyon ve finansal veriler (PII / Operasyon / Teknik) olarak ayrıştırıldı.
- [] CRM Alanları Sınıflandırıldı:** Sadakat programı, misafir tercihleri ve iletişim geçmişi etiketlendi.
- [] Web Rezervasyon Alanları Sınıflandırıldı:** Form alanları, IP/çerez logları ve ödeme adımı verileri tarandı.
- [] "Riskli Alanlar" Listesi Çıkarıldı:** Doğrudan kişiyi tanımlayan e-posta, telefon, TCKN/Pasaport ve serbest not kutuları kilit listeye alındı.

B) SENTETİK DATASET KURALLARI

- E-posta Formatı:** Tüm e-posta adresleri geçersiz ve kurgusal alan adlarıyla tanımlanır (Örn: guest_123@example.test).
- Telefon Formatı:** Telefon numaraları gerçek hatlara ulaşmayacak standart test bloklarından seçilir (Varsayım: [+90 555 000 0000](tel:+905550000000)).
- İsim/Soyisim:** Tamamen kurgusal isim veri havuzundan rastgele türetilir.
- Sezonsallık & Operasyonel Mantık:** Rezervasyon tarihleri, kalış süreleri ve doluluk oranları otelin gerçek sezonel dağılımını temsil eder.
- Segment Senaryoları:** Çocuklu aile, solo gezgin ve acente/grup rezervasyon senaryoları sentetik verilere yansıtılır.

C) PII LEAK TEST CHECKLIST'İ

- [] Gerçek Domain & Numara Taraması:** Test veritabanında gerçek e-posta uzantıları ([.com](mailto:example.com), [.net](mailto:example.net), [.com.tr](mailto:example.com.tr)) ve numara formatları regex ile taranır.

- [] **Serbest Not Alanı Taraması:** Özel istek veya resepsiyon notu kutularında unutulmuş açık isim veya TCKN/Pasaport metni analizi yapılır.
- [] **Export Örnekleri Kontrolü:** Test ortamından alınan rapor ve dış aktarım dosyaları PII varlığına karşı incelenir.
- [] **Pass / Fail Sonuç Raporu:** Tarama sonuçları audit belgesi olarak sistem tarafından **PASS / FAIL** şeklinde raporlanır.

D) TEST DATA REFRESH / CLEAN JOB

- **Job Takvimi:** Test ortamı verilerinin haftalık/aylık periyotlarla sıfırlanıp yeniden sentetik veriyle beslenme takvimi kilitlenir.
- **Job Log Alanları:** Log kaydında **job_id**, **execution_time**, **processed_records**, **result_status** zorunlu tutulur.
- **Başarısız Job Alarmı:** Otomatik veri yenileme veya temizleme işlemi başarısız olduğunda IT ekibine anlık uyarı düşer (Varsayım).

E) 3 SENTETİK VERİ SENARYOSU

1. **PMS Entegrasyon Testi:** Oda ataması, check-in/check-out ve folyo simülasyonlarının kurgusal isimlerle uçtan uca çalıştırılması.
2. **CRM Segmentasyon Testi:** Misafir profil davranışları ve segmentasyon kurallarının sahte e-posta/telefon verileriyle doğrulanması.
3. **Web Rezervasyon Funnel Testi:** Online rezervasyon adımları, ödeme ekranı geçişleri ve onay e-postalarının test domain'leri üzerinden sınanması.

F) 14 GÜNLÜK SPRINT PLAN (UYGULAMA)

- **Gün 1–3 (Dictionary & Sınıflandırma):** Tüm veri kaynakları taranarak PII haritası ve Data Dictionary tamamlanır.
- **Gün 4–7 (Hibrit Dataset Üretimi):** Operasyonel mantığı koruyan sentetik veri üretim algoritmaları ve SQL betikleri hazırlanır.
- **Gün 8–10 (Leak Test Otomasyonu):** Sandbox veritabanlarını taranıp gerçek PII kalıntıları yakalayan otomasyon kurulur (Varsayım).
- **Gün 11–14 (Refresh Job & Audit Pack Integration):** Otomatik temizleme job'ları bağlatılır, çalıştırma logları KVKK Audit Pack'e eklenir.

ÖNCESİ / SONRASI KPI TABLOSU

KPI	Önce	Sonra	Not
Test Ortamındaki Gerçek PII Kayıt Sayısı	TBD (Riskli / Gerçek Veri)	0 (Sıfır)	Sandbox ortamlarında %100 sentetik veriye geçildi
PII Leak Test Sonucu	TBD (Bilinmiyor)	%100 PASS	Düzenli otomatize taramalarla kanıt üretildi

Test Data Refresh Sıralılığı	TBD (Manuel / Düzensiz)	Otomatik (Haftalık/Aylık)	Temizleme ve yenileme job'ları loglandı
Açık Not Alanlarındaki PII İhlal Riski	TBD (Yüksek)	%0 (Regex Taraması)	Serbest metin alanları otomasyonla sterilize edildi

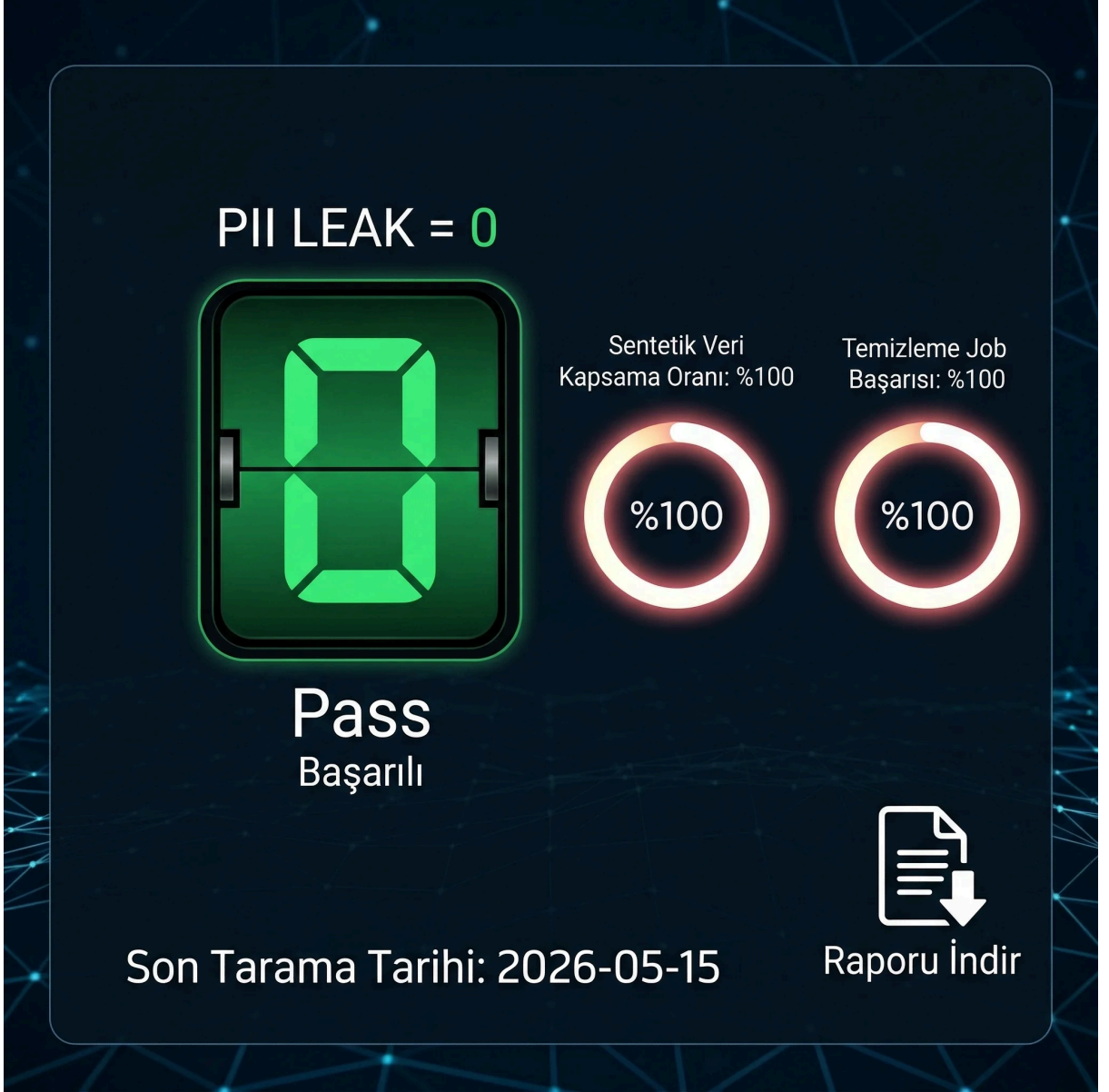
DELIVERABLES (TESLİM EDİLECEK ÇIKTILAR)

- **1x** Otel Yazılımları PII & Data Dictionary Haritası
- **1x** Sentetik Dataset Üretim Betiği & Veri Seti (v1.0)
- **1x** Otomatize PII Leak Testi & Pass/Fail Rapor Şablonu
- **1x** Test Data Refresh Job Log Mimarısı & Denetim Paketi Kanıt Eki



[Checklist Card] —“Test verisi checklist, otel projeleri bağlamı”

1200×1200px kare formatında, Otel Projeleri İçin Sentetik Veri ve Anonim Test Ortamları Checklist Kartı. Data Dictionary Sınıflandırması, Sentetik Dataset Kuralları, PII Leak Taraması ve Test Data Refresh Job Kontrol Arayüzü.



[KPI Card] — “Leak test KPI, otel projeleri bağlamı”

1200×1200px kare formatında, Test Ortamı PII Leak ve Veri Güvenliği KPI Analiz Kartı. Test Ortamındaki Gerçek PII Oranı (Sıfır), PII Leak Test Başarı Oranı (%100 PASS), Otomatik Refresh Job Durum Paneli ve Audit Pack Kanıt Göstergeleri.