🤖 AI Agent - Flowable Entegrasyon Rehberi

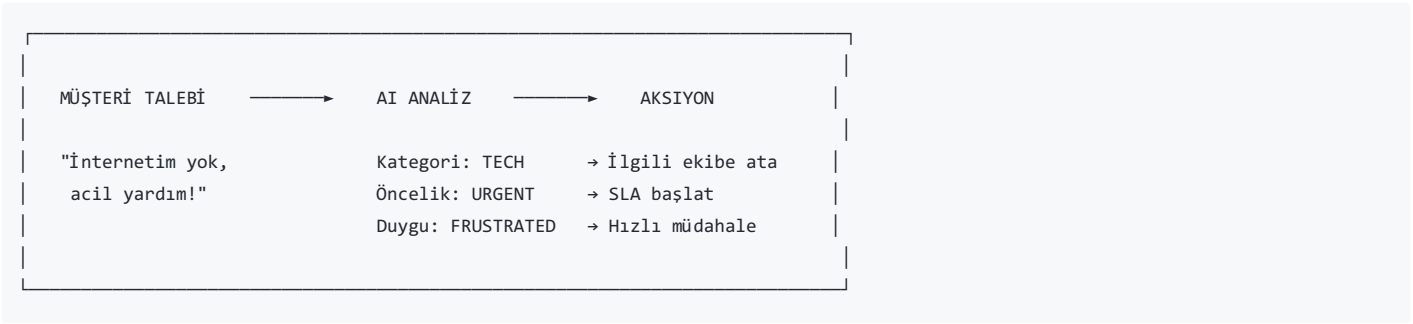
Kapsamlı Teknik Dokümantasyon

> Bu rehber, Python BPM Agent'ın nasıl çalıştığını ve şirketin mevcut Flowable sistemine nasıl entegre edilebileceğini açıklar.

📑 İçindekiler

🎯 BÜYÜK RESİM

Projenin Temel Amacı

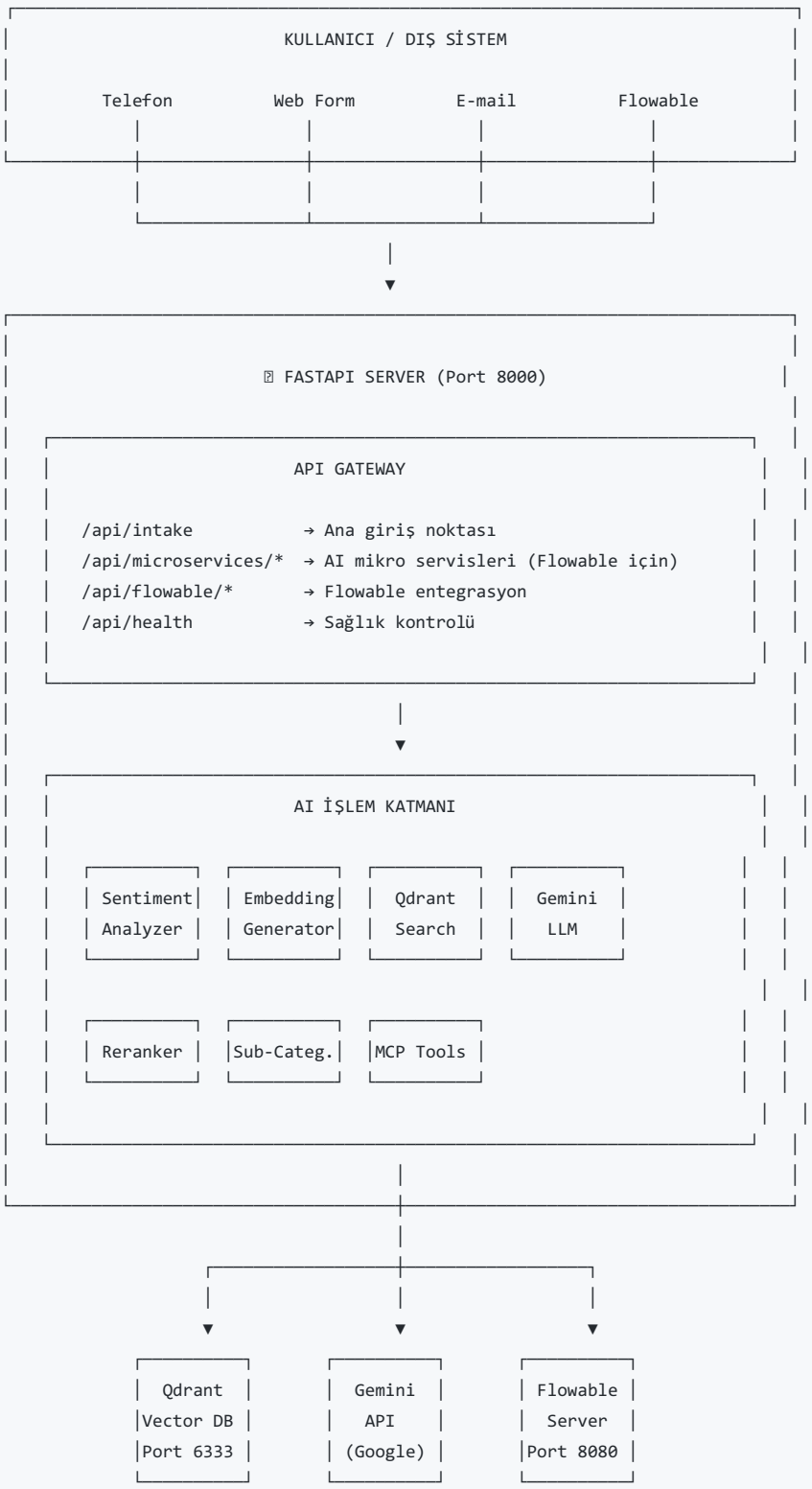


Ne yapıyor?

- Müşteri taleplerini otomatik kategorize ediyor
- Duygu analizi ile gerçek aciliyeti tespit ediyor
- Şirket politikalarına dayalı kararlar veriyor (RAG)
- Mevcut Flowable workflow'larına entegre oluyor

📦 PART 1: PROJENİN İÇ YAPISI

Katman Katman Mimari



Bir Talebin Yolculuğu (End-to-End)

ADIM 1: TALEP GELİYOR

```
┌─────────────────────────────────────────────────────────────────────────┐
│ ADIM 1: TALEP GELİYOR                                                   │
├─────────────────────────────────────────────────────────────────────────┤
│                                                                         │
│   Müşteri: "3 gündür internet yok, iş yapamıyorum, çok sinirliyim!"     │
│                                                                         │
│   JSON olarak:                                                          │
│   {                                                                     │
│     "text": "3 gündür internet yok, iş yapamıyorum, çok sinirliyim!",   │
│     "source": "phone",                                                  │
│     "customer_id": "CUST-12345"                                         │
│   }                                                                     │
│                                                                         │
└─────────────────────────────────────────────────────────────────────────┘
```

ADIM 2: DUYGU ANALİZİ (Sentiment Analysis)

```
┌─────────────────────────────────────────────────────────────────────────┐
│ ADIM 2: DUYGU ANALİZİ (Sentiment Analysis)                              │
├─────────────────────────────────────────────────────────────────────────┤
│                                                                         │
│   POST /api/microservices/sentiment                                     │
│                                                                         │
│   Girdi: "3 gündür internet yok, iş yapamıyorum, çok sinirliyim!"       │
│                                                                         │
│   AI İşlemi:                                                            │
│   - "sinirliyim" → ANGRY/FRUSTRATED                                     │
│   - "3 gündür" → Uzun süre beklemiş                                     │
│   - "iş yapamıyorum" → İş etkisi var                                    │
│                                                                         │
│   Çıktı:                                                                │
│   {                                                                     │
│     "emotion": "FRUSTRATED",                                            │
│     "intensity": 8,                                                     │
│     "justifies_urgency": true,   ← Sinirli ama haklı, acil              │
│     "reasoning": "3 gün beklemek + iş kaybı = gerçek aciliyet"          │
│   }                                                                     │
│                                                                         │
└─────────────────────────────────────────────────────────────────────────┘
```

ADIM 3: METNİ VEKTÖRE ÇEVİR (Embedding Generation)

```
┌──────────────────────────────────────────────────────────────────────┐
│ ADIM 3: METNİ VEKTÖRE ÇEVİR (Embedding Generation)                   │
├──────────────────────────────────────────────────────────────────────┤
│                                                                      │
│   POST /api/microservices/embedding                                  │
│                                                                      │
│   Girdi: "3 gündür internet yok, iş yapamıyorum"                     │
│                                                                      │
│   Gemini text-embedding-004 modeli:                                  │
│   Metin → 768 boyutlu sayı dizisi                                    │
│                                                                      │
│   Çıktı:                                                             │
│   {                                                                  │
│     "embedding": [0.023, -0.145, 0.892, ..., 0.034],  ← 768 sayı     │
│     "model": "text-embedding-004",                                   │
│     "text_length": 42                                                │
│   }                                                                  │
│                                                                      │
│   💡 Neden vektör?                                                   │
│   Çünkü "internet yok" ile "bağlantı problemi" aynı anlama geliyor,  │
│   ama kelimeler farklı. Vektörler anlamı yakalar.                    │
│                                                                      │
└──────────────────────────────────────────────────────────────────────┘
```

ADIM 4: BENZERLİK ARAMASI (Qdrant Vector Search)

```
┌──────────────────────────────────────────────────────────────────────┐
│ ADIM 4: BENZERLİK ARAMASI (Qdrant Vector Search)                      │
├──────────────────────────────────────────────────────────────────────┤
│                                                                      │
│   POST /api/microservices/qdrant-search                              │
│                                                                      │
│   Girdi: embedding vektörü [0.023, -0.145, ...]                      │
│                                                                      │
│   Qdrant'ta aramak:                                                  │
│   "Bu vektöre en yakın 10 dokümanı bul"                              │
│                                                                      │
│   Bulunan dokümanlar:                                                │
│   ┌──────────────────────────────────────────────────────────────┐   │
│   │ 1. "İnternet kesintisi durumunda 4 saat içinde müdahale..."  %92  │ │
│   │ 2. "Bağlantı sorunları için teknik ekip atanır..."           %87  │ │
│   │ 3. "Modem reset prosedürü: 1. Fişi çekin..."                 %81  │ │
│   │ 4. "SLA: Kritik bağlantı sorunları 2 saat içinde..."         %79  │ │
│   │ 5. "Network ekibi iletişim bilgileri..."                     %74  │ │
│   └──────────────────────────────────────────────────────────────┘   │
│                                                                      │
└──────────────────────────────────────────────────────────────────────┘
```

ADIM 5: SONUÇLARI SIRALA (Reranking)

```
┌─────────────────────────────────────────────────────────────────────────┐
│ ADIM 5: SONUÇLARI SIRALA (Reranking)                                    │
├─────────────────────────────────────────────────────────────────────────┤
│                                                                         │
│   POST /api/microservices/rag-reranking                                 │
│                                                                         │
│   İlk arama bazen gürültülü olabilir. Reranker daha akıllıca sıralar:   │
│                                                                         │
│   Kriterler:                                                            │
│   - Soruyla ne kadar ilgili?                                            │
│   - Doküman tipi (policy > faq > general)                               │
│   - Güncellik                                                           │
│                                                                         │
│   Sonuç: En iyi 5 doküman seçilir                                       │
│                                                                         │
└─────────────────────────────────────────────────────────────────────────┘
```

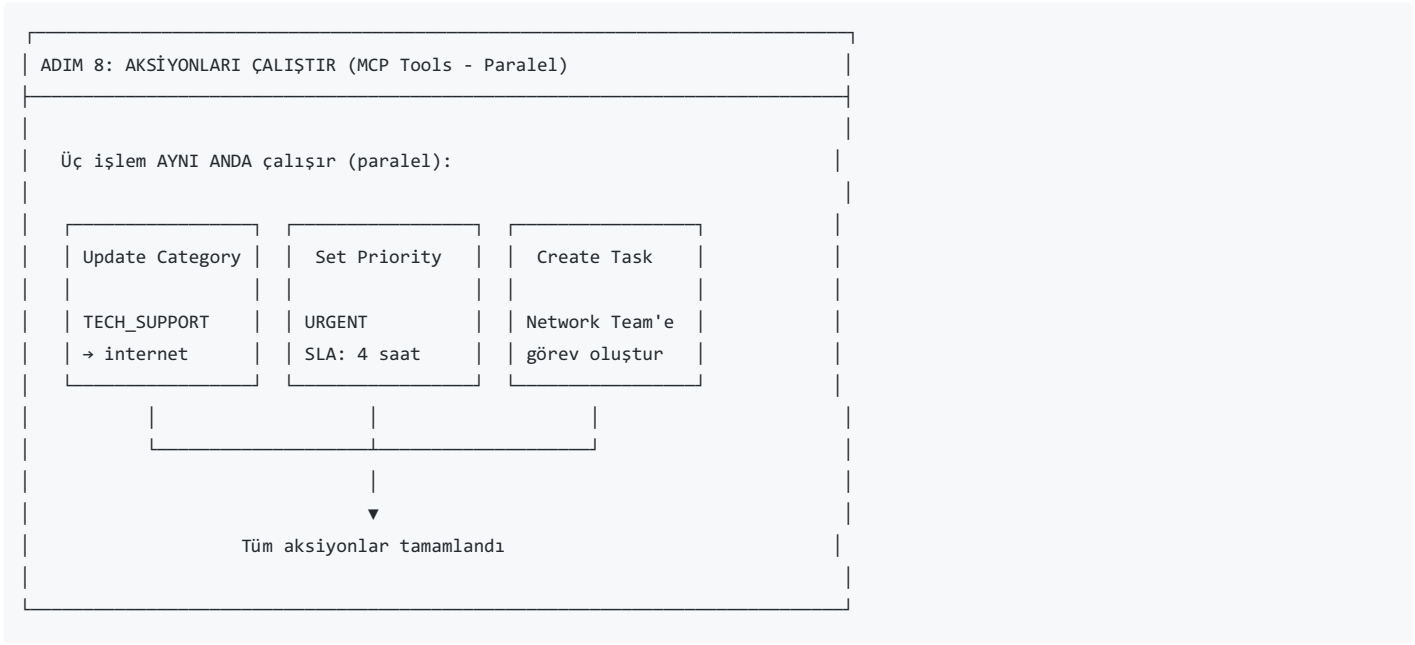
ADIM 6: LLM KARAR VERİYOR (Gemini LLM Decision)

```
┌─────────────────────────────────────────────────────────────────────────┐
│ ADIM 6: LLM KARAR VERİYOR (Gemini LLM Decision)                         │
├─────────────────────────────────────────────────────────────────────────┤
│                                                                         │
│   POST /api/microservices/llm-call                                      │
│                                                                         │
│   Gemini'ye gönderilen:                                                 │
│   ┌─────────────────────────────────────────────────────────────────┐   │
│   │ SYSTEM: Sen bir müşteri hizmetleri uzmanısın.                   │   │
│   │                                                                 │   │
│   │ MÜŞTERİ TALEBİ:                                                 │   │
│   │ "3 gündür internet yok, iş yapamıyorum, çok sinirliyim!"        │   │
│   │                                                                 │   │
│   │ ŞİRKET POLİTİKALARI (RAG Context):                              │   │
│   │ 1. İnternet kesintisi durumunda 4 saat içinde müdahale...       │   │
│   │ 2. Bağlantı sorunları için teknik ekip atanır...                │   │
│   │ 3. SLA: Kritik bağlantı sorunları 2 saat içinde...              │   │
│   │                                                                 │   │
│   │ DUYGU ANALİZİ:                                                  │   │
│   │ - Duygu: FRUSTRATED (8/10)                                      │   │
│   │ - Aciliyet haklı: EVET                                          │   │
│   │                                                                 │   │
│   │ GÖREV: Kategori, öncelik ve yapılacak aksiyonları belirle.      │   │
│   └─────────────────────────────────────────────────────────────────┘   │
│                                                                         │
│   Gemini'nin kararı:                                                    │
│   {                                                                     │
│     "intent": "internet_kesintisi_sikayet",                             │
│     "category": "TECH_SUPPORT",                                         │
│     "priority": "URGENT",         ← 3 gün + iş kaybı = acil             │
│     "auto_approve": false,        ← İnsan onayı gerekli                 │
│     "reasoning": "Müşteri 3 gündür bekliyor ve iş kaybı var. SLA aşılmış."│
│   }                                                                     │
│                                                                         │
└─────────────────────────────────────────────────────────────────────────┘
```

ADIM 7: ALT KATEGORİ BELİRLE (Sub-Categorization)

```
┌──────────────────────────────────────────────────────────────────────┐
│ ADIM 7: ALT KATEGORİ BELİRLE (Sub-Categorization)                    │
├──────────────────────────────────────────────────────────────────────┤
│                                                                      │
│   POST /api/microservices/sub-categorization                         │
│                                                                      │
│   TECH_SUPPORT kategorisi içinde nereye gitmeli?                     │
│                                                                      │
│   TECH_SUPPORT                                                       │
│   ├── internet     ← "internet yok" → BURAYA                         │
│   ├── email                                                          │
│   ├── software                                                       │
│   ├── hardware                                                       │
│   └── security                                                       │
│                                                                      │
│   Çıktı:                                                             │
│   {                                                                  │
│     "main_category": "TECH_SUPPORT",                                 │
│     "sub_category": "internet",                                      │
│     "assigned_team": "Network Team",                                 │
│     "sla_hours": 4                                                   │
│   }                                                                  │
│                                                                      │
└──────────────────────────────────────────────────────────────────────┘
```

ADIM 8: AKSİYONLARI ÇALIŞTIR (MCP Tools - Paralel)



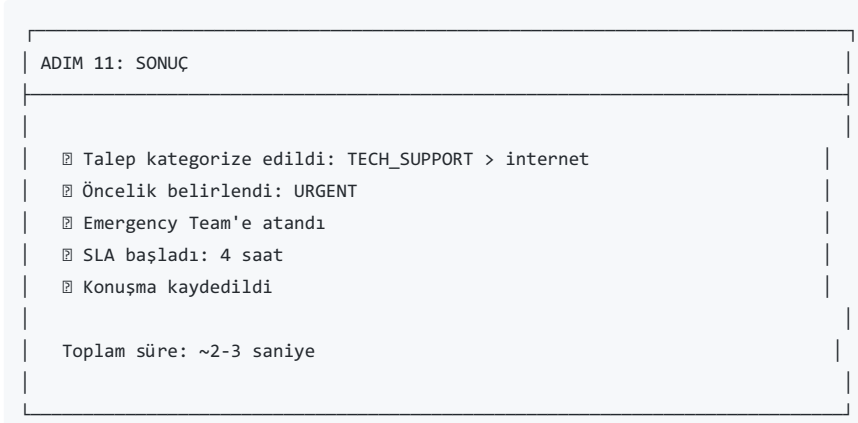
ADIM 9: PRİORİTY ROUTING (Önceliğe Göre Yönlendirme)

```
┌─────────────────────────────────────────────────────────────────────────┐
│ ADIM 9: PRİORİTY ROUTING (Önceliğe Göre Yönlendirme)                    │
├─────────────────────────────────────────────────────────────────────────┤
│                                                                         │
│   Priority = URGENT olduğu için:                                        │
│                                                                         │
│   ┌──────────┐                                                          │
│   │ URGENT   │ ──────► 🚨 Emergency Team (Acil Müdahale Ekibi)          │
│   └──────────┘                                                          │
│   ┌──────────┐                                                          │
│   │ HIGH     │ ──────► 👔 Managers (Yöneticiler)                        │
│   └──────────┘                                                          │
│   ┌──────────┐                                                          │
│   │ MEDIUM   │ ──────► 🛠️ Support Team (Destek Ekibi)                   │
│   └──────────┘                                                          │
│   ┌──────────┐                                                          │
│   │ LOW      │ ──────► 🛠️ Support Team (Destek Ekibi)                   │
│   └──────────┘                                                          │
│                                                                         │
│   Bu müşteri URGENT → Emergency Team'e atandı!                          │
│                                                                         │
└─────────────────────────────────────────────────────────────────────────┘
```

ADIM 10: KAYDET (Store to Vector DB)

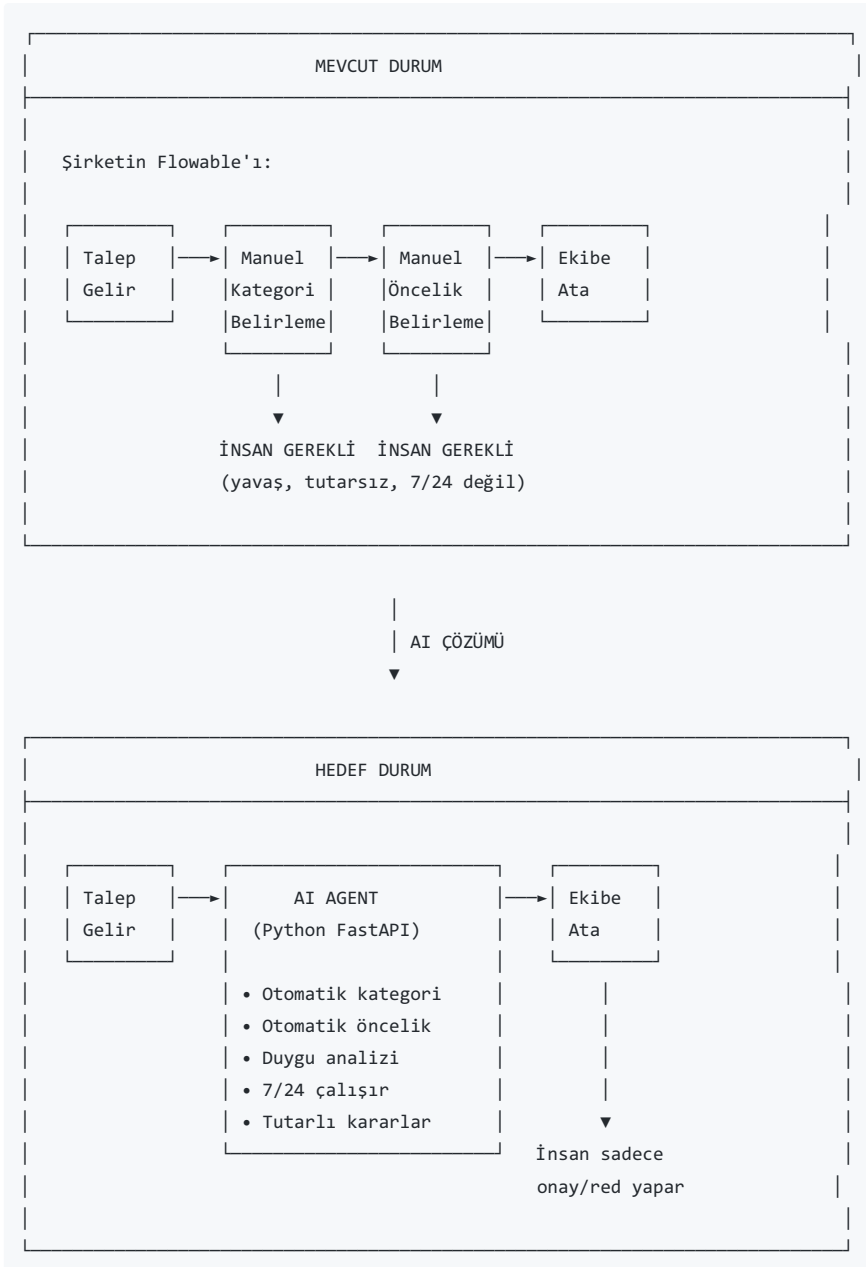
```
┌─────────────────────────────────────────────────────────────────────────┐
│ ADIM 10: KAYDET (Store to Vector DB)                                    │
├─────────────────────────────────────────────────────────────────────────┤
│                                                                         │
│   POST /api/microservices/store-vector                                  │
│                                                                         │
│   Bu konuşmayı Qdrant'a kaydet:                                         │
│   - Gelecekte benzer sorular gelirse referans olsun                     │
│   - Analiz ve raporlama için                                            │
│                                                                         │
│   Kaydedilen:                                                           │
│   {                                                                     │
│     "text": "3 gündür internet yok...",                                 │
│     "metadata": {                                                       │
│       "customer_id": "CUST-12345",                                      │
│       "category": "TECH_SUPPORT",                                       │
│       "sub_category": "internet",                                       │
│       "priority": "URGENT",                                             │
│       "resolution": "pending",                                          │
│       "timestamp": "2026-01-21T..."                                     │
│     }                                                                   │
│   }                                                                     │
│                                                                         │
└─────────────────────────────────────────────────────────────────────────┘
```

ADIM 11: SONUÇ



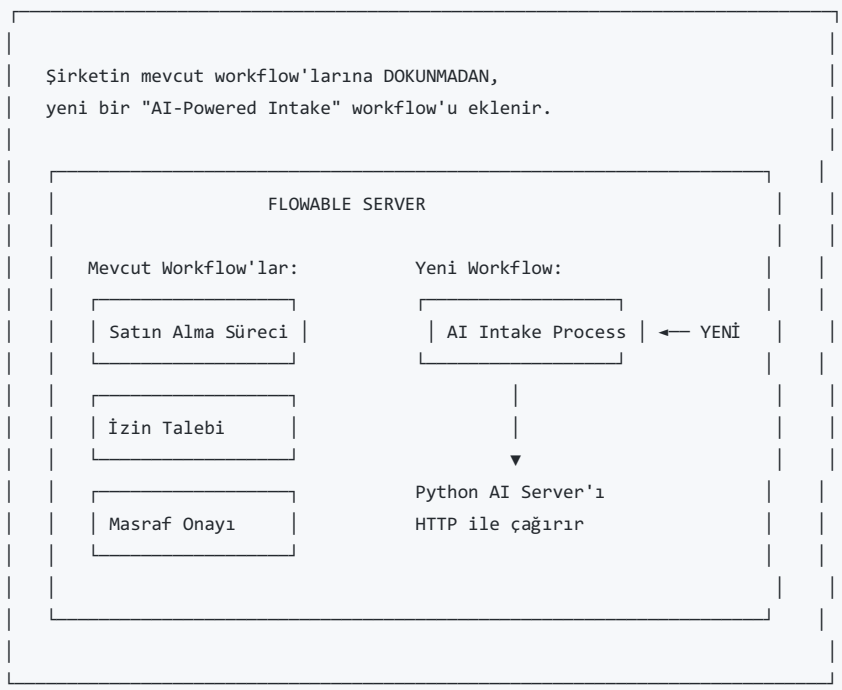
PART 2: FLOWABLE ENTEGRASYONU

Mevcut Durum vs Hedef

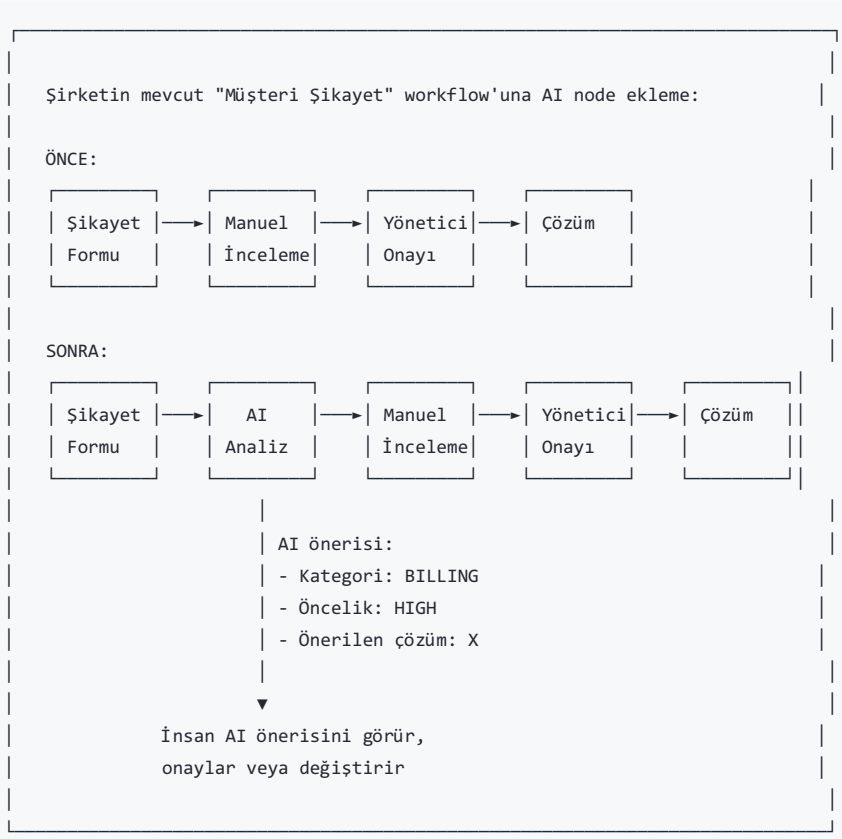


Entegrasyon Senaryoları

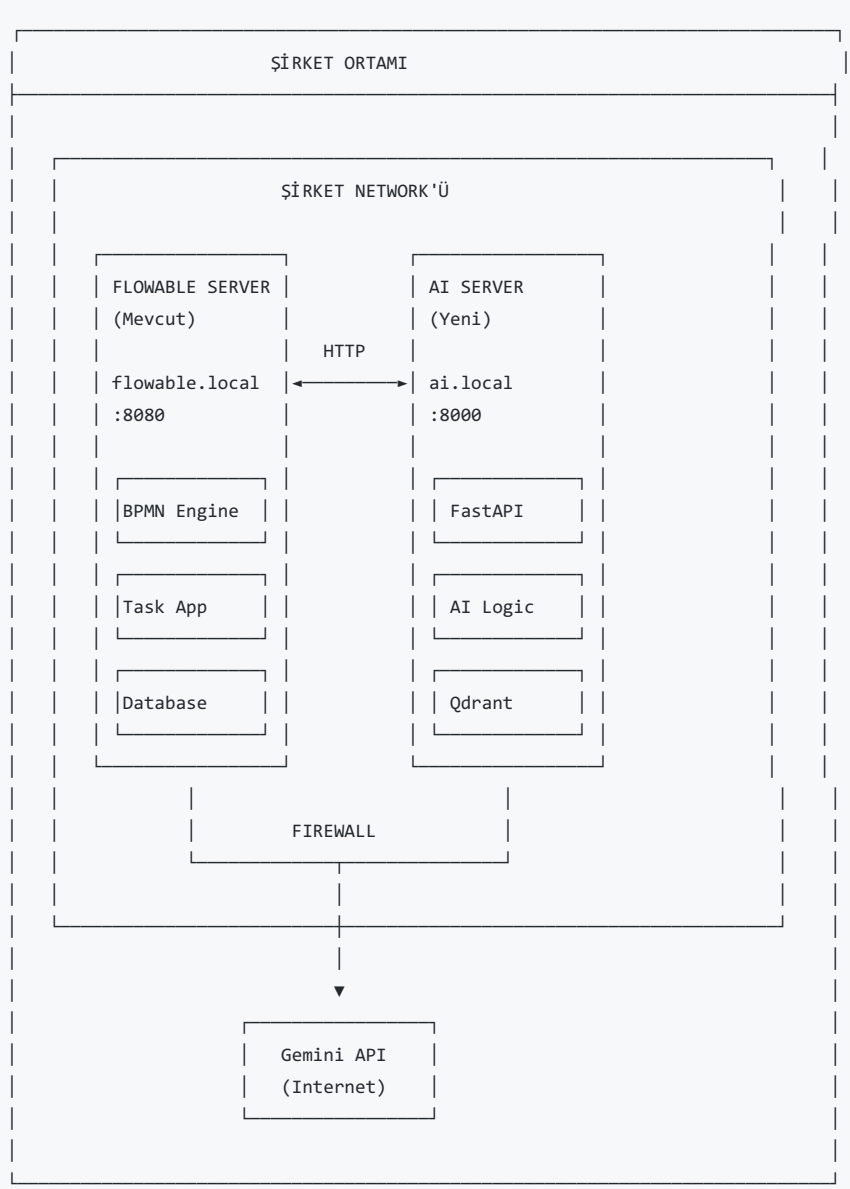
Senaryo A: Yeni Workflow Ekleme



Senaryo B: Mevcut Workflow'a AI Ekleme



Teknik Entegrasyon Detayı



BPMN'de HTTP Service Task Kullanımı

Flowable, Python API'ı şu şekilde çağırır:

```
<serviceTask id="aiAnalysis" name="AI Analizi" flowable:type="http">
  <extensionElements>
    <flowable:field name="requestMethod">
      <flowable:string>POST</flowable:string>
    </flowable:field>
    <flowable:field name="requestUrl">
      <flowable:string>http://ai-server:8000/api/microservices/sentiment</flowable:string>
    </flowable:field>
    <flowable:field name="requestHeaders">
      <flowable:string>Content-Type: application/json</flowable:string>
    </flowable:field>
    <flowable:field name="requestBody">
      <flowable:expression><![CDATA[{
        "text": "${customerRequest}",
        "source": "${source}"
      }]]></flowable:expression>
    </flowable:field>
    <flowable:field name="responseVariableName">
      <flowable:string>aiResult</flowable:string>
    </flowable:field>
  </extensionElements>
</serviceTask>
```

API Endpoint Listesi

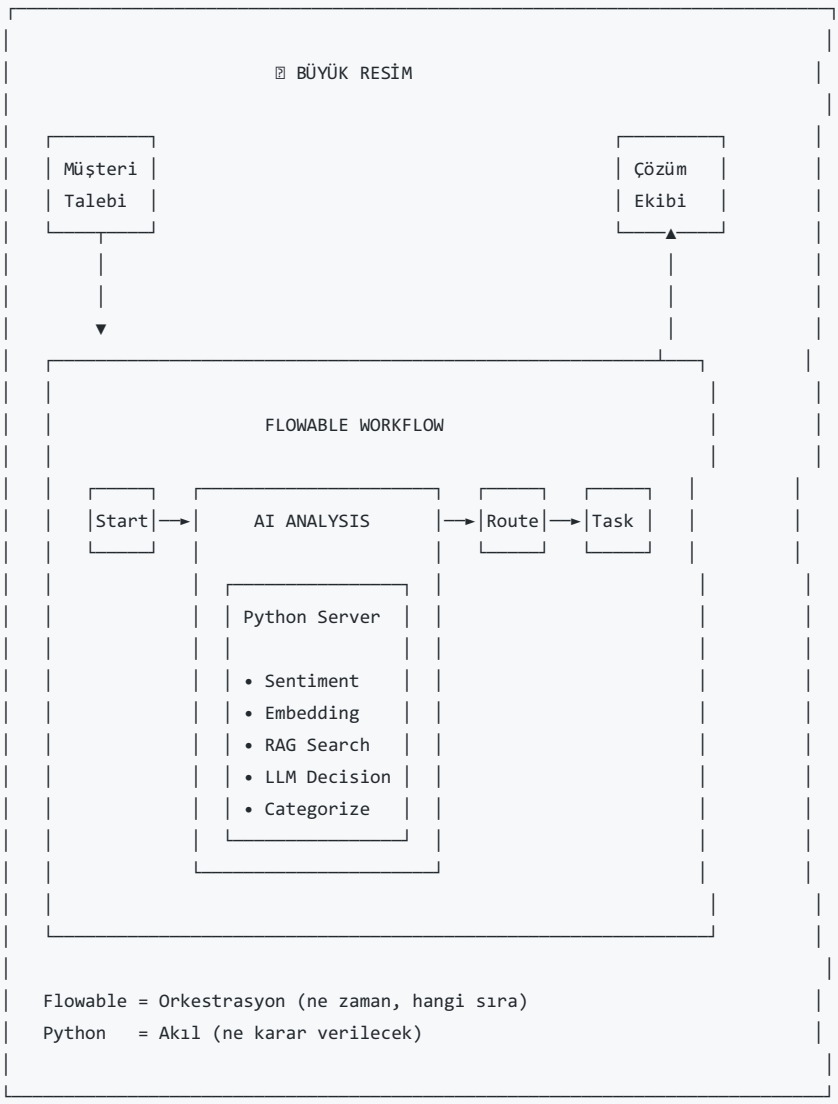
Endpoint	İşlev	Flowable Node
`/api/microservices/sentiment`	Duygu analizi	😊 Sentiment Analysis
`/api/microservices/embedding`	Metin → Vektör	□ Generate Embedding
`/api/microservices/qdrant-search`	Vektör DB arama	□ Qdrant Search
`/api/microservices/rag-reranking`	Sonuç sıralama	□ Rerank Results
`/api/microservices/llm-call`	AI karar	□ Gemini LLM
`/api/microservices/sub-categorization`	Alt kategori	□ Sub-Categorization
`/api/microservices/mcp/update-category`	Kategori güncelle	□ Update Category
`/api/microservices/mcp/set-priority`	Öncelik belirle	⚡ Set Priority
`/api/microservices/mcp/create-task`	Görev oluştur	□ Create Task
`/api/microservices/store-vector`	Kaydet	□ Store to Vector DB

□ PART 3: ŞİRKETE SUNUM

Problem - Çözüm - Değer

```
┌──────────────────────────────────────────────────────────────────────┐
│                                                                      │
│   PROBLEM:                                                           │
│   ──────────                                                         │
│   • Müşteri talepleri manuel kategorize ediliyor                     │
│   • Tutarsız öncelik belirleme                                       │
│   • Yavaş yanıt süreleri                                             │
│   • 7/24 hizmet zor                                                  │
│                                                                      │
│   ÇÖZÜM:                                                             │
│   ─────────                                                          │
│   • AI-powered otomatik sınıflandırma                                │
│   • Duygu analizi ile gerçek aciliyet tespiti                        │
│   • Şirket politikalarına dayalı kararlar (RAG)                      │
│   • Mevcut Flowable'a entegre                                        │
│                                                                      │
│   DEĞERİ:                                                            │
│   ──────────                                                         │
│   • İşlem süresi: dakikalar → saniyeler                              │
│   • Tutarlılık: %70 → %95+                                           │
│   • 7/24 çalışma                                                     │
│   • Çalışanlar rutin işlerden kurtulur                               │
│                                                                      │
│   TEKNİK:                                                            │
│   ──────────                                                         │
│   • Python FastAPI server (ayrı deploy)                              │
│   • Flowable HTTP Service Task ile bağlantı                          │
│   • Mevcut sistemlere minimum müdahale                               │
│   • Kolay rollback                                                   │
│                                                                      │
│   PLAN:                                                              │
│   ────────                                                           │
│   • Hafta 1-2: Demo ortamında POC                                    │
│   • Hafta 3-4: Test ortamında pilot                                  │
│   • Ay 2: Gerçek bir workflow ile production                         │
│                                                                      │
└──────────────────────────────────────────────────────────────────────┘
```

Özet Diagram



Hızlı Başlangıç

1. Servisleri Başlat

```
# Docker (Qdrant + Flowable)
cd python-bpm-agent
docker-compose up -d

# Python Server
source venv/bin/activate
uvicorn app.main:app --reload --host 0.0.0.0 --port 8000

# BPMN Deploy
python flowable/deploy_to_flowable.py
```

2. Test Et

```
# Health check
curl http://localhost:8000/api/health

# Sentiment analizi test
curl -X POST http://localhost:8000/api/microservices/sentiment \
  -H "Content-Type: application/json" \
  -d '{"text": "Çok sinirliyim, 3 gündür bekletiliyorum!", "source": "phone"}'
```

3. Flowable UI

- **Modeler**: http://localhost:8080/flowable-modeler
- **Task App**: http://localhost:8080/flowable-task
- **Credentials**: admin / test

Sonuç

Bu proje, mevcut Flowable BPM sistemlerine **minimum müdahale** ile AI yetenekleri ekler.

Anahtar noktalar:

- Her AI adımı bağımsız bir HTTP endpoint
- Flowable sadece bu endpoint'leri çağırır
- Mevcut workflow'lar bozulmaz
- İstediğin zaman AI'ı devre dışı bırakabilirsin
