**HR Q&A POC - Progress Report**

**✅ COMPLETED (75%)**

**Phase 1: Foundation Setup ✅ DONE**

* ✅ Project structure created
* ✅ Conda environment setup
* ✅ Environment configuration (.env)
* ✅ All service connections tested:
  + MongoDB Atlas ✅
  + Azure AI Search ✅
  + Azure OpenAI ✅

**Phase 2: Data Preparation ✅ DONE**

* ✅ Comprehensive HR dataset created (25 employees, 10 categories)
* ✅ Excel file with 10 sheets:
  + Personal Info, Employment, Performance, Engagement
  + Learning, Compensation, Attendance, Experience
  + Attrition, Project History
* ✅ ETL pipeline code developed

**🔄 IN PROGRESS (25%)**

**Phase 2: Data Processing 🔄 CURRENT**

* 🔄 **TODAY**: Run ETL pipeline (Excel → MongoDB)
* ⏳ Verify MongoDB collections created
* ⏳ Test data relationships

**⏳ PENDING**

**Phase 3: Search Integration ⏳ NEXT**

* ⏳ MongoDB → Azure AI Search indexing
* ⏳ Generate vector embeddings (OpenAI)
* ⏳ Create search indexes
* ⏳ Test hybrid search capabilities

**Phase 4: Query Engine ⏳ FINAL**

* ⏳ Natural language query processor
* ⏳ Azure OpenAI integration for responses
* ⏳ CLI interface for testing
* ⏳ Sample Q&A testing

**🎯 STATUS SUMMARY**

**Completed**: Foundation + Data Creation  
**Current**: Data Processing (ETL)  
**Remaining**: Search Integration + Query Engine  
**Timeline**: ~75% complete, 1-2 days remaining

**Next Immediate Task**: Complete ETL pipeline execution ✅