Comprehensive Product Requirements Document (PRD): Autonomous Hybrid AI Chatbot Builder

# 1. Product Vision

Create an autonomous AI-driven chatbot builder that drastically simplifies chatbot creation for organizations. The platform autonomously ingests internal knowledge bases, API specifications, and workflow definitions to rapidly deploy specialized chatbot agents with minimal manual intervention, effectively lowering operational costs by integrating hybrid methods (database-driven and LLM-based responses).

# 2. Target Users

- Primary: Product managers tasked with creating and managing AI chatbot agents.  
- Secondary: Developers customizing or extending the platform.  
- Tertiary: End users interacting with chatbot agents.

# 3. Key Differentiators

- Significant Reduction of Human Effort:  
 - Automated ingestion and structuring of company knowledge and workflows.  
 - Minimal manual setup, limited to providing internal resources and API documentation.  
  
- Cost Efficiency via Hybrid Methodology:  
 - Utilization of pre-stored database-driven responses for frequent queries.  
 - Restricted use of LLMs primarily for complex interactions.  
  
- Rapid Deployment with Customizable Control:  
 - Quick deployment with extensive post-deployment refinement capabilities.

# 4. Functional Requirements

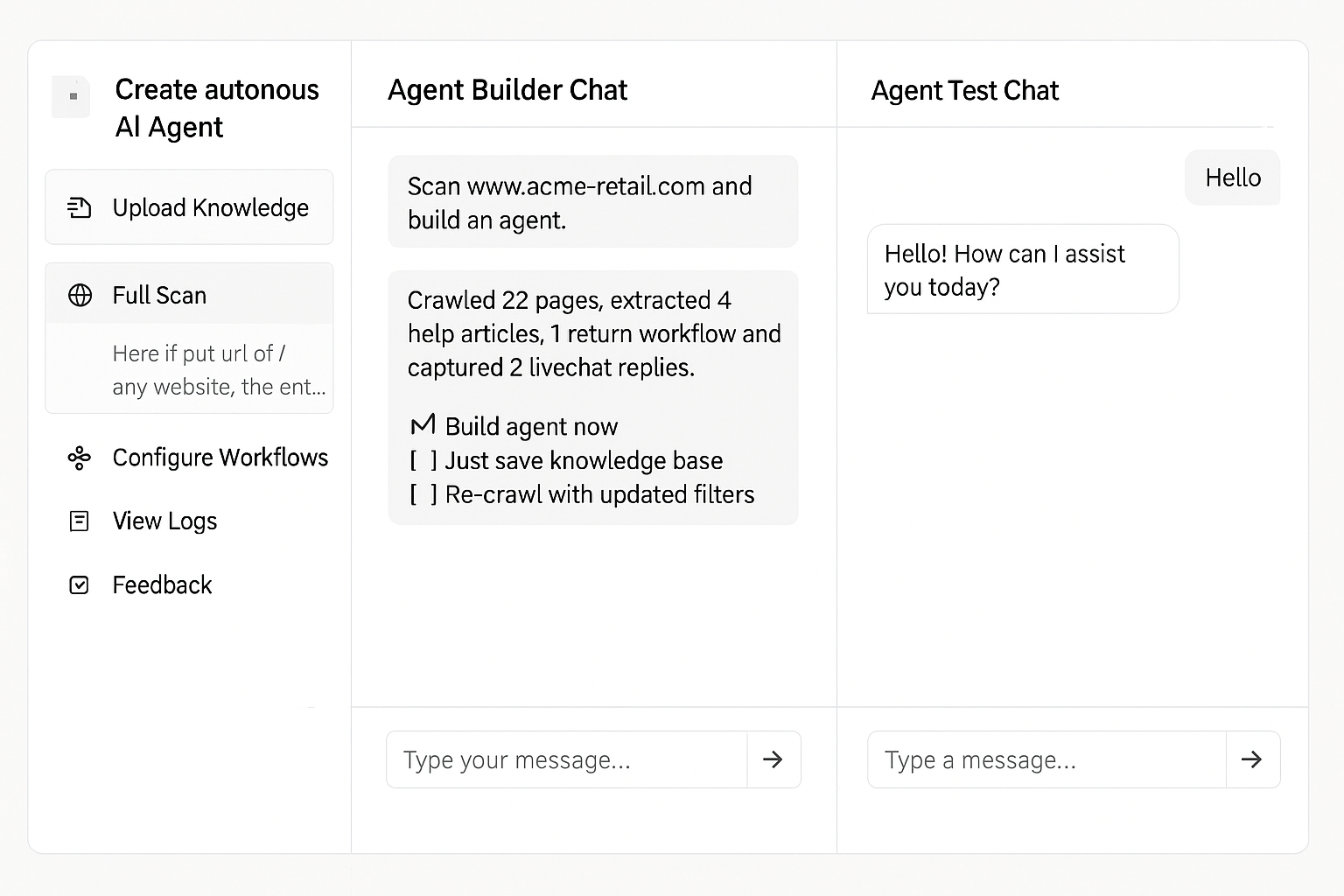
Autonomous Capabilities:  
- Automatic knowledge ingestion and structuring.  
- Automated chatbot creation and deployment.  
- Workflow automation via API-to-tool conversion.  
- Dynamic and personalized suggestion generation based on user interactions and profiles.  
  
Required Human Oversight:  
- Provisioning of internal knowledge sources, APIs, and workflows.  
- Optional manual refinements post-deployment.

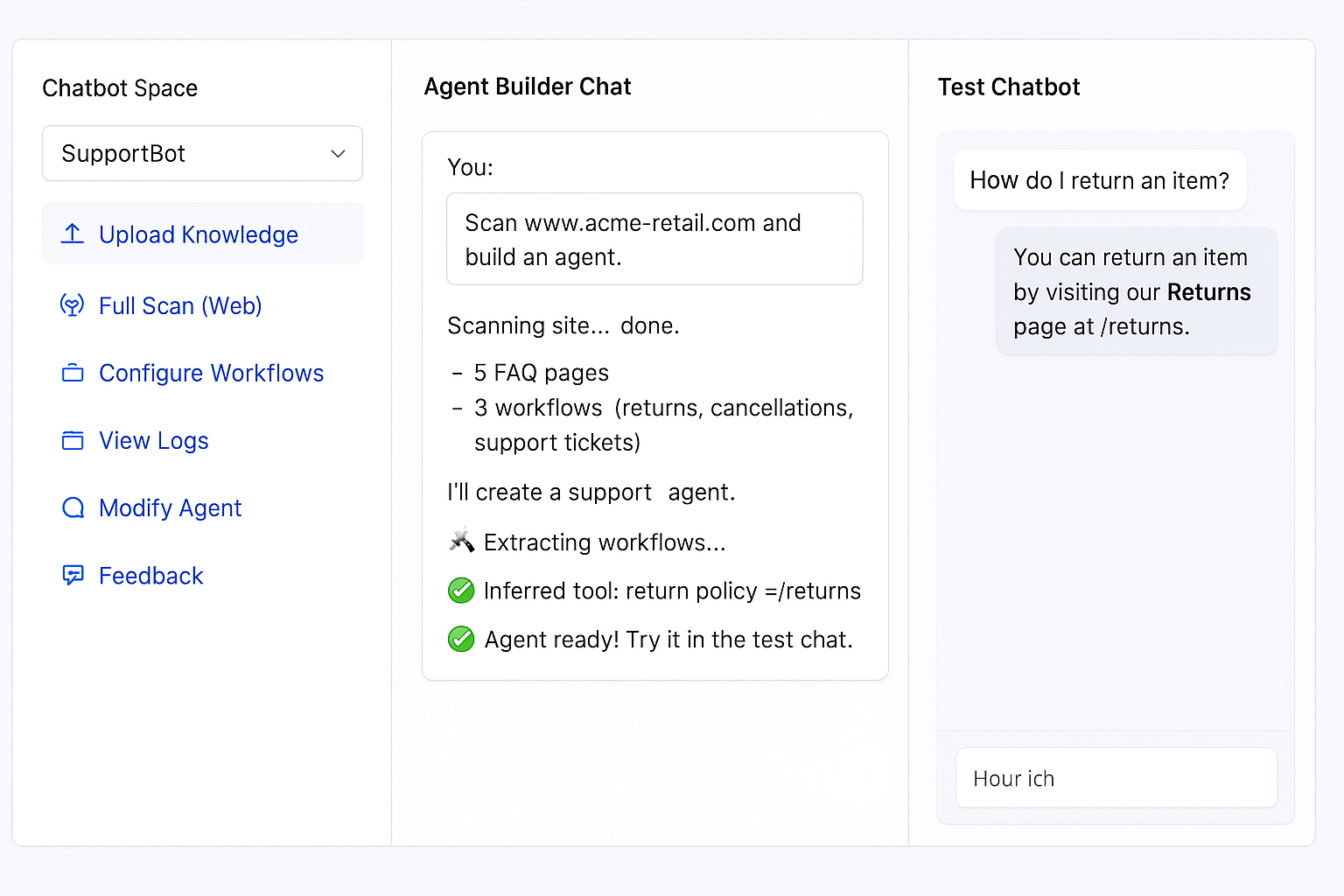
# 5. Technical Specifications

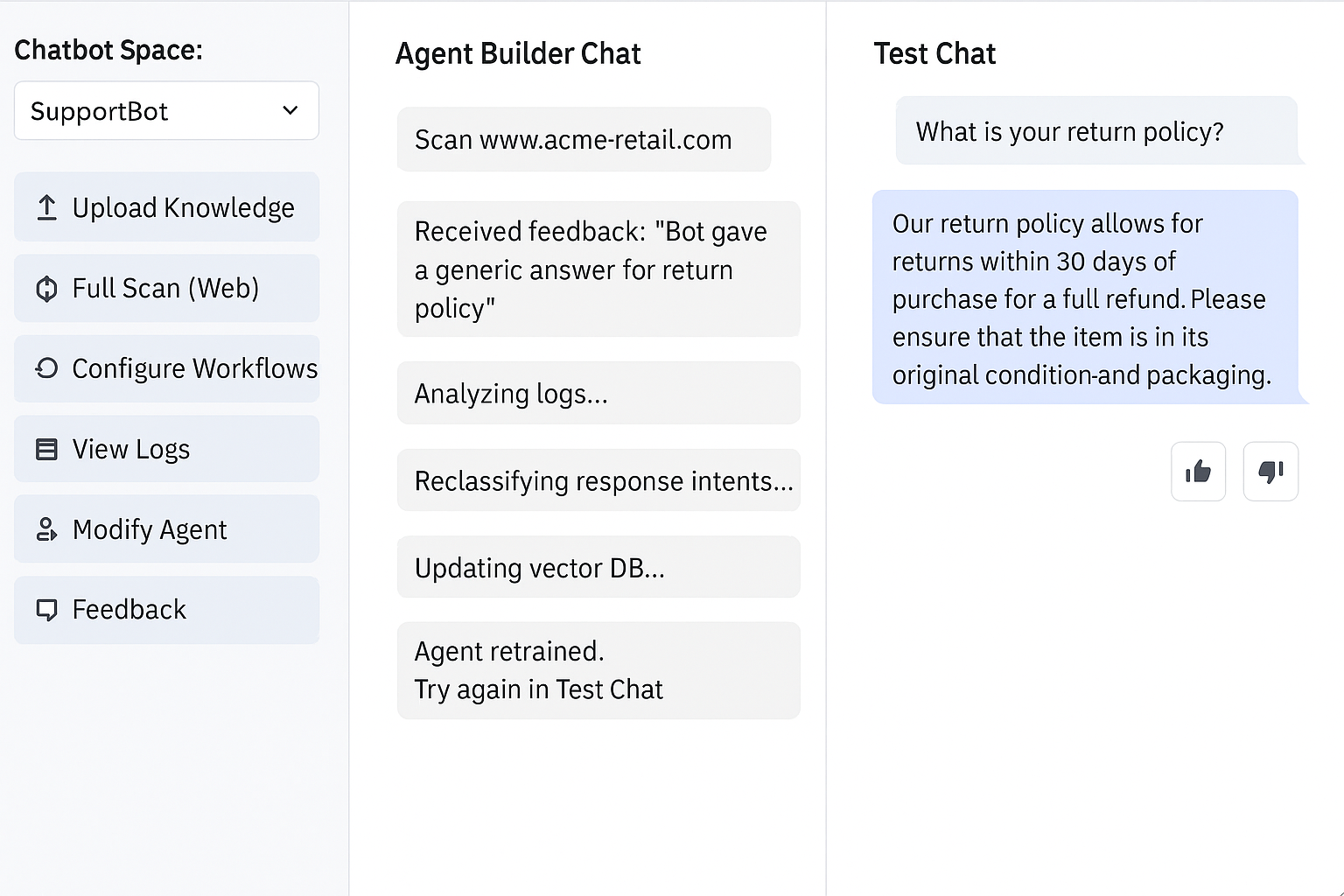
Core Technologies:  
- Python 3.9+  
- LLM Integration: OpenAI, Anthropic, LlamaCPP  
- Frameworks: LangChain, CrewAI, MCP SDK, SmolaGents  
- Document Processing: LlamaIndex, Unstructured.io  
- Vector Database: Chroma, FAISS, Pinecone  
- Knowledge Graph: Neo4j or custom  
- Frontend: Streamlit  
- API Layer: FastAPI  
  
Supporting Technologies:  
- Docker, Git  
- Testing: Pytest  
- CI/CD: GitHub Actions  
- Monitoring: Prometheus, Grafana  
- Logging: ELK Stack

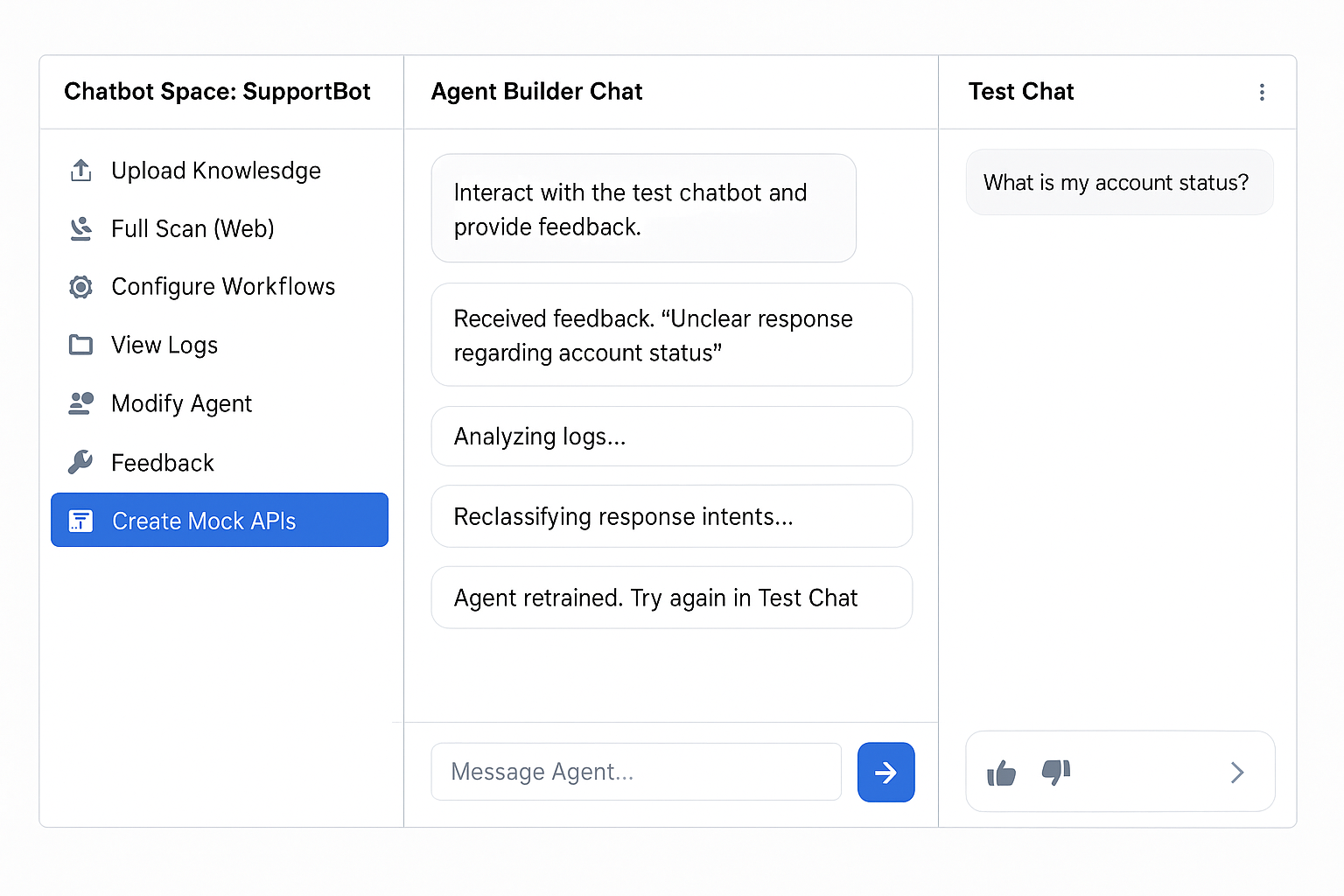
# 6. User Interface (Streamlit UI)

Panels:  
- Tools (Left Panel): Manage tools, knowledge sources, and API configurations.  
- Agent Builder Chat (Central Panel): Progress dashboard, interaction for agent configuration.  
- Test Chatbot (Right Panel): Real-time chatbot interaction and testing.  
  
Enhanced Features:  
- Chatbot spaces for scoped management.  
- Full Scan tool for website crawling and knowledge ingestion.  
- Real-time retraining via feedback from chatbot testing.  
- Mock API Generator for rapid prototyping.









# 7. Success Metrics

Primary Metrics:  
- Cost per Chat: Reduced via increased database-driven responses.  
- Human Transfer Rate: Target <5%, adjustable per industry standards.  
  
Secondary Metrics:  
- Deployment Speed: Accelerated deployment timeline.  
- Chatbot Autonomy Rate: Percentage of self-resolved interactions.

# 8. System Architecture

Core Components:  
- Autonomous Orchestrator Agent  
- Knowledge Discovery & Processing  
- Hybrid Chatbot Engine (LLM & DB-driven)  
- Suggestion Engine (Dynamic, Personalized)  
- Workflow Automation Module  
- Human Oversight Interface

# 9. Risk Management

| Risk | Impact | Likelihood | Mitigation |  
|------|--------|------------|------------|  
| Token Costs Escalation | High | Medium | Hybrid response strategy |  
| Complex Knowledge Sources | Medium | High | Robust parsing, manual overrides |  
| Security Vulnerabilities | High | Medium | Frequent audits, secure data practices |

# 10. Implementation Plan

Phase 1 (Weeks 1-4):  
- Basic knowledge ingestion.  
- Initial hybrid chatbot implementation.  
  
Phase 2 (Weeks 5-8):  
- Advanced hybrid integration.  
- Dynamic suggestion engine.  
  
Phase 3 (Weeks 9-12):  
- Workflow automation integration.  
- Enhanced personalization and analytics.  
  
Phase 4 (Weeks 13-16):  
- Final refinements and security enhancements.  
- Documentation and deployment preparations.

# 11. Future Enhancements

- Multi-language support  
- Advanced analytics dashboards  
- Enterprise system integrations  
- Mobile and voice interactions