

PurpleFabric SDK 3.0.0 - Functional Implementation Scenarios

Table of Contents

1. [Overview](#)
2. [Authentication & Session Management Scenarios](#)
3. [Document Library \(DocLib\) Use Cases](#)
4. [Knowledge Base \(KB\) Use Cases](#)
5. [Advanced Integration Scenarios](#)
6. [Error Handling & Monitoring Scenarios](#)
7. [Real-World Application Scenarios](#)

Overview

The PurpleFabric SDK 3.0.0 is a comprehensive Python SDK that provides two main services:

- **Document Library (DocLib)**: Document storage, organization, and management
- **Knowledge Base (KB)**: Semantic search and RAG (Retrieval-Augmented Generation) capabilities

This document outlines all functional scenarios and use cases where this SDK can be implemented.

Authentication & Session Management Scenarios

1. Multi-Environment Configuration

Scenario: Managing different environments (dev, staging, production) with separate credentials.

Use Cases:

- Configure multiple credential profiles in `~/.purplefabric/credentials`
- Switch between environments using profile parameter

- Isolate development and production workloads
- Support team members with different access levels

Implementation:

```
# Development environment
kb_dev = client('kb', profile='dev')

# Production environment
kb_prod = client('kb', profile='prod')
```

2. Token-Based Authentication

Scenario: Using pre-obtained JWT tokens for authentication.

Use Cases:

- Integration with existing authentication systems
- CI/CD pipelines with service accounts
- Short-lived token usage
- Token rotation workflows

3. Credential-Based Authentication with Auto-Token Management

Scenario: Using user credentials with automatic token generation and refresh.

Use Cases:

- Long-running applications
- Applications without token management infrastructure
- Automatic token refresh before expiry (5-minute buffer)
- Fallback authentication when tokens expire

4. Hybrid Authentication (Token + Credentials)

Scenario: Using both token and credentials for enhanced reliability.

Use Cases:

- High-availability applications requiring fallback

- Applications with token expiry concerns
- Seamless token refresh without interruption
- Production systems requiring redundancy

5. Session Management for Multiple Services

Scenario: Managing multiple service clients (KB and DocLib) with shared authentication.

Use Cases:

- Applications using both document storage and knowledge base
- Shared session configuration across services
- Centralized credential management
- Service reloading after credential rotation

Document Library (DocLib) Use Cases

6. Document Repository Management

Scenario: Creating and managing document repositories (datasets).

Use Cases:

- Technical documentation repositories
- Legal document management systems
- Compliance document storage
- Research paper collections
- Product documentation archives
- Training material libraries

7. Hierarchical Folder Organization

Scenario: Organizing documents in nested folder structures.

Use Cases:

- Department-based document organization
- Project-based folder structures
- Version-controlled documentation

- Category-based classification
- Multi-level taxonomy implementation

8. Document Upload & Processing

Scenario: Uploading documents with metadata and processing options.

Use Cases:

- Batch document ingestion
- Single document uploads
- Document processing workflows
- Asynchronous large file uploads
- Multi-format document support (PDF, DOCX, TXT, etc.)

9. Document Metadata Management

Scenario: Adding, updating, and managing document metadata.

Use Cases:

- Document classification and tagging
- Status tracking (draft, reviewed, approved)
- Version management
- Department/category assignment
- Custom metadata schema implementation
- Batch metadata updates

10. Document Search & Discovery

Scenario: Searching documents with various filters and criteria.

Use Cases:

- Full-text search across documents
- Metadata-based filtering
- Tag-based document discovery
- Folder-specific searches
- Deep search in subfolders
- Paginated result retrieval

11. Document Processing Pipeline

Scenario: Drafting and processing documents for downstream use.

Use Cases:

- Document processing workflows
- Integration with knowledge base chunking
- Document transformation pipelines
- Status monitoring and tracking
- Async processing for large documents

12. Document Library Integration with Knowledge Base

Scenario: Using documents from DocLib in Knowledge Base operations.

Use Cases:

- Reusing documents across multiple knowledge bases
- Centralized document storage with multiple KB access
- Document versioning for KB content
- Access control through DocLib permissions

Knowledge Base (KB) Use Cases

13. Knowledge Base Creation & Configuration

Scenario: Creating knowledge bases with specific chunking and embedding strategies.

Use Cases:

- Technical documentation knowledge bases
- FAQ knowledge bases
- Product documentation search systems
- Internal wiki search
- Customer support knowledge bases
- Research paper search systems

Configuration Options:

- Chunk strategies (BLOCK, PARAGRAPH, etc.)
- Embedding models (Azure OpenAI, Amazon Titan)
- Table extraction settings
- Image processing settings
- Private/public asset settings

14. Document Upload to Knowledge Base

Scenario: Uploading documents for semantic search capabilities.

Use Cases:

- Technical documentation ingestion
- Product manual uploads
- Training material processing
- Research paper indexing
- API documentation ingestion
- Policy and procedure documentation

Features:

- Single and batch uploads
- Metadata and tags during upload
- Automatic chunking initiation
- Dataset and folder organization
- Async upload for large files

15. Direct Text Ingestion

Scenario: Adding text content directly as chunks without document files.

Use Cases:

- FAQ entry ingestion
- API documentation from code comments
- Dynamic content updates
- User-generated content
- Real-time knowledge updates
- Synthetic chunk creation
- Testing and experimentation

16. Semantic Search (RAG Search)

Scenario: Performing semantic search with metadata and tag filters.

Use Cases:

- Natural language query answering
- Context-aware search
- Filtered search by department/category
- Multi-faceted search with metadata
- Tag-based content discovery
- Relevance-ranked results

Search Features:

- Query-based semantic search
- Metadata filtering
- Tag filtering
- Pagination support
- Similarity threshold configuration
- Top-K retrieval with reranking

17. Chunk Management

Scenario: Managing, retrieving, and organizing chunks within knowledge bases.

Use Cases:

- Chunk retrieval and inspection
- Chunk metadata management
- Chunk tagging and categorization
- Chunk deletion and cleanup
- Grouping chunks by document
- Chunk-level search operations

18. Knowledge Base Configuration Management

Scenario: Updating KB settings for optimal search performance.

Use Cases:

- Adjusting chunk strategies for different content types

- Switching embedding models
- Configuring retrieval parameters (top_k, similarity range)
- Enabling/disabling reranker
- Optimizing search performance
- A/B testing different configurations

Configuration Parameters:

- Chunk strategy updates
- Embedding model selection
- Top-K retrieval count (1-100)
- Similarity range configuration
- Top-N final results after reranking
- Reranker model selection (Cohere, Azure AI)

19. Document Management in Knowledge Base

Scenario: Managing documents within knowledge bases.

Use Cases:

- Document listing and details retrieval
- Document metadata updates
- Document deletion
- Document status tracking
- Batch document operations
- Document-to-chunk relationship management

20. Chunking Process Management

Scenario: Initializing and monitoring document chunking processes.

Use Cases:

- Manual chunking initiation
- Chunking status monitoring
- Processing large document sets
- Retry failed chunking operations
- Chunking workflow management

Advanced Integration Scenarios

21. Event-Driven Monitoring

Scenario: Using observers to monitor operations and react to events.

Use Cases:

- Progress tracking for long-running operations
- Status change notifications
- Error monitoring and alerting
- Data caching and retrieval
- Operation statistics collection
- Custom event handling

Observer Types:

- Progress Observer: Track upload/processing progress
- Status Observer: Monitor operation status changes
- Error Observer: Error tracking with retry support
- Data Observer: Cache and track data updates
- Composite Observer: Combine multiple observers

22. Asynchronous Operations

Scenario: Using async/await for concurrent operations.

Use Cases:

- Concurrent document uploads
- Parallel knowledge base operations
- Non-blocking API calls
- High-throughput document processing
- Background task execution
- Scalable batch operations

23. Error Handling & Recovery

Scenario: Implementing robust error handling with retry logic.

Use Cases:

- Transient error recovery
- Rate limit handling with exponential backoff
- Validation error handling
- Resource not found scenarios
- Service unavailability handling
- Comprehensive error logging

24. Multi-Service Workflows

Scenario: Combining DocLib and KB services in integrated workflows.

Use Cases:

- Upload documents to DocLib, then add to KB
- Organize documents in DocLib, search in KB
- Document versioning in DocLib, semantic search in KB
- Centralized storage with distributed search
- Access control through DocLib, content search through KB

Error Handling & Monitoring Scenarios

25. Comprehensive Error Categorization

Scenario: Handling different types of errors with appropriate responses.

Use Cases:

- Client-side error handling (validation, authentication)
- Service error handling (resource not found, conflicts)
- Network error handling (timeouts, connection issues)
- Rate limit handling
- Server error handling
- Custom error recovery strategies

26. Retry Logic Implementation

Scenario: Implementing retry mechanisms for transient failures.

Use Cases:

- Network timeout recovery
- Rate limit retry with backoff
- Service unavailability handling
- Polling timeout management
- Exponential backoff strategies
- Maximum retry limit enforcement

27. Operation Monitoring & Logging

Scenario: Tracking operations with detailed logging and statistics.

Use Cases:

- Operation progress monitoring
- Success/failure rate tracking
- Performance metrics collection
- Error rate monitoring
- Usage analytics
- Audit trail creation

Real-World Application Scenarios

28. Technical Documentation Search System

Scenario: Building a searchable technical documentation system.

Use Cases:

- Developer documentation portals
- API documentation search
- Architecture documentation systems
- Technical specification search
- Code documentation indexing
- Internal wiki with semantic search

Implementation Flow:

1. Create KB with appropriate chunk strategy
2. Upload technical documents
3. Configure search parameters
4. Implement search interface
5. Add metadata for categorization

29. Customer Support Knowledge Base

Scenario: Creating a searchable knowledge base for customer support.

Use Cases:

- FAQ systems with semantic search
- Support article search
- Troubleshooting guide search
- Product documentation for support teams
- Ticket resolution assistance
- Self-service support portals

Implementation Flow:

1. Create KB for support content
2. Ingest FAQ entries as text chunks
3. Upload support articles and guides
4. Configure for natural language queries
5. Implement search with category filters

30. Research Paper & Academic Document Search

Scenario: Building a semantic search system for academic documents.

Use Cases:

- Research paper repositories
- Academic literature search
- Citation and reference discovery
- Research collaboration platforms
- Literature review tools
- Academic knowledge management

Implementation Flow:

1. Organize papers in DocLib datasets
2. Create KB with research-focused configuration
3. Upload papers with metadata (author, year, topic)
4. Implement semantic search for concepts
5. Enable metadata filtering by field, year, etc.

31. Legal Document Management & Search

Scenario: Managing and searching legal documents.

Use Cases:

- Contract management systems
- Legal case document search
- Compliance document repositories
- Legal research tools
- Document versioning and tracking
- Metadata-based legal document organization

Implementation Flow:

1. Create DocLib datasets for legal documents
2. Organize by case, client, or document type
3. Add metadata (case number, date, type, status)
4. Create KB for semantic search of legal content
5. Implement search with legal-specific filters

32. Product Documentation Portal

Scenario: Creating a comprehensive product documentation system.

Use Cases:

- Product manual search
- User guide systems
- Installation and setup documentation
- API documentation portals
- Release notes and changelogs
- Product knowledge management

Implementation Flow:

1. Organize documentation in DocLib by product/version
2. Create KBs per product or unified KB
3. Upload documentation with version metadata
4. Configure search for natural language queries
5. Implement version-aware search

33. Internal Knowledge Management System

Scenario: Building an internal knowledge management platform.

Use Cases:

- Company wiki with semantic search
- Internal process documentation
- Team knowledge sharing
- Best practices repository
- Training material search
- Organizational knowledge base

Implementation Flow:

1. Create DocLib for document storage
2. Organize by department, team, or topic
3. Create KBs for different knowledge domains
4. Upload internal documents and guides
5. Implement department/topic-filtered search

34. Content Management with Semantic Search

Scenario: Combining traditional CMS with semantic search capabilities.

Use Cases:

- Blog content search
- Article discovery systems
- Content recommendation engines
- Media library with search
- Content tagging and categorization
- Multi-format content search

Implementation Flow:

1. Store content in DocLib
2. Add metadata (author, category, publish date)
3. Create KB for semantic search
4. Ingest content as chunks
5. Implement content discovery interface

35. Compliance & Audit Document Management

Scenario: Managing compliance documents with search capabilities.

Use Cases:

- Regulatory compliance documentation
- Audit trail management
- Policy and procedure search
- Compliance reporting
- Document retention management
- Version control for compliance docs

Implementation Flow:

1. Create DocLib datasets for compliance docs
2. Organize by regulation, department, or date
3. Add compliance-specific metadata
4. Create KB for policy and procedure search
5. Implement audit-friendly search and retrieval

36. Training & Learning Management System

Scenario: Building a searchable training material system.

Use Cases:

- Training course material search
- Learning resource discovery
- Skill-based content organization
- Training progress tracking
- Course material versioning
- Multi-format training content

Implementation Flow:

1. Organize training materials in DocLib
2. Create KBs by training domain or skill
3. Upload training documents and videos
4. Add metadata (skill level, topic, duration)
5. Implement skill-based search

37. API Documentation & Developer Portal

Scenario: Creating a searchable API documentation system.

Use Cases:

- API endpoint documentation search
- Code example discovery
- Integration guide search
- SDK documentation portals
- API reference with semantic search
- Developer resource discovery

Implementation Flow:

1. Ingest API documentation as text chunks
2. Upload API guides and examples
3. Create KB with API-focused configuration
4. Add metadata (endpoint, method, version)
5. Implement developer-friendly search interface

38. Multi-Tenant Knowledge Management

Scenario: Managing knowledge bases for multiple tenants or clients.

Use Cases:

- SaaS knowledge base per customer
- Client-specific documentation portals
- Isolated knowledge bases
- Multi-organization content management
- White-label knowledge systems
- Tenant-specific search configurations

Implementation Flow:

1. Create separate KBs per tenant
2. Use tenant-specific authentication
3. Organize documents by tenant in DocLib
4. Implement tenant isolation
5. Provide tenant-specific search interfaces

39. Document Processing & Transformation Pipeline

Scenario: Building automated document processing workflows.

Use Cases:

- Automated document ingestion
- Document format conversion
- Batch document processing
- Document quality validation
- Processing status monitoring
- Error handling and retry mechanisms

Implementation Flow:

1. Set up document processing pipeline
2. Upload documents to DocLib
3. Process and draft documents
4. Add processed documents to KB
5. Monitor processing with observers
6. Handle errors and retries

40. Hybrid Search System

Scenario: Combining keyword and semantic search capabilities.

Use Cases:

- Multi-modal search interfaces
- Keyword + semantic search
- Metadata filtering with semantic search
- Tag-based + semantic discovery
- Faceted search with semantic results
- Advanced search query processing

Implementation Flow:

1. Use DocLib for keyword search
2. Use KB for semantic search
3. Combine results from both systems
4. Implement unified search interface
5. Provide search result ranking

Summary of Functional Capabilities

Core Services

1. **Document Library Service:** Document storage, organization, metadata management, search
2. **Knowledge Base Service:** Semantic search, RAG capabilities, chunk management, text ingestion

Key Features

- **Authentication:** Multiple methods (token, credentials, hybrid)
- **Session Management:** Default and custom sessions with token management
- **Document Management:** Upload, organize, search, process documents
- **Semantic Search:** RAG search with metadata and tag filtering
- **Chunk Management:** Direct text ingestion, chunk operations, metadata management
- **Configuration:** Flexible KB configuration (chunking, embeddings, reranking)
- **Event System:** Observers for progress, status, errors, data
- **Error Handling:** Comprehensive exception hierarchy with retry support
- **Async Operations:** Non-blocking operations for scalability
- **Type Safety:** Full Python type hints support

Integration Patterns

- **Standalone DocLib:** Document repository and management
- **Standalone KB:** Semantic search and knowledge management
- **Integrated Workflow:** DocLib + KB for comprehensive document and search solutions
- **Multi-Environment:** Support for dev, staging, production
- **Multi-Tenant:** Isolated knowledge bases per tenant/client

Use Case Categories

1. **Document Management:** Storage, organization, versioning, search
2. **Knowledge Management:** Semantic search, RAG, content discovery
3. **Enterprise Applications:** Internal wikis, documentation portals, knowledge bases
4. **Customer-Facing:** Support portals, FAQ systems, product documentation
5. **Specialized Domains:** Legal, research, compliance, training systems
6. **Developer Tools:** API documentation, technical documentation, code search

Conclusion

The PurpleFabric SDK 3.0.0 provides comprehensive capabilities for document management and semantic search, enabling a wide range of applications from simple document repositories to sophisticated knowledge management systems. The SDK's flexible configuration, robust error handling, and event-driven architecture make it suitable for both simple use cases and complex enterprise applications.