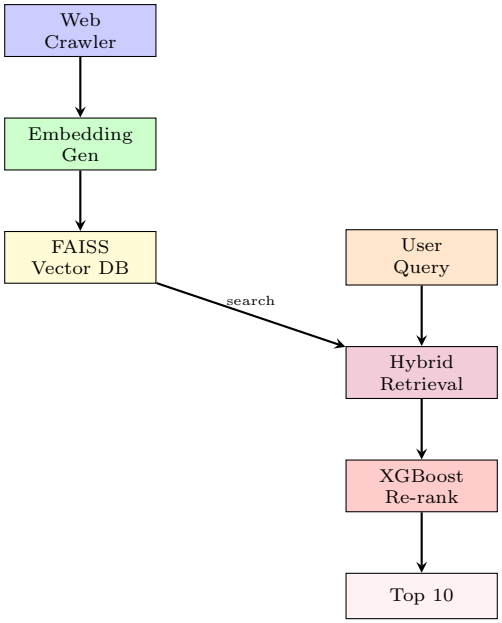
SHL Assessment Recommendation System
Approach and Optimization

1 Approach

We built a RAG-based recommendation system that suggests relevant SHL assessments from natural language queries. The system crawls **389 Individual Test Solutions** (exceeding the 377 requirement) and uses a multi-stage pipeline: semantic search, hybrid retrieval, and XGBoost re-ranking to achieve **61.56% Mean Recall@10**.

2 Solution Architecture



2.1 Data Pipeline

Web Crawling: BeautifulSoup-based crawler extracts 389 assessments with metadata (name, description, duration, test types, remote/adaptive support).

Embedding Generation: Google Gemini `text-embedding-004` generates 768-dim embeddings. Rich text includes name, description, test types, and support features.

Vector Storage: FAISS index with cosine similarity for efficient semantic search.

2.2 Retrieval Pipeline

Hybrid Retrieval: Combines semantic search (FAISS) with keyword matching:

- Semantic similarity (cosine distance)
- Keyword boosting (name: +0.40, description: +0.10)
- Query expansion with domain synonyms
- Retrieves top 100 candidates for re-ranking

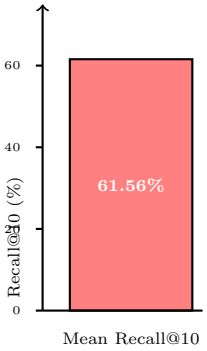
XGBoost Re-ranking: Learned model with 17 features: semantic scores, keyword matches, duration/test type matching, role/skill indicators, remote/adaptive flags. Trained on 500 samples (50 positive, 450 negative) with class imbalance handling.

3 Optimization Process

Key Challenges: URL variations (`/products/` vs `/solutions/products/`), keyword matching prioritization, query understanding.

Iterations: (1) URL normalization with alternate URL generation, (2) Enhanced query preprocessing and expansion, (3) XGBoost re-ranking achieving final performance.

4 Performance Results



5 Key Technical Decisions

Hybrid Retrieval: Semantic search (FAISS) + keyword matching ensures exact term matches are prioritized while capturing conceptual relevance.

Learned Re-ranking: XGBoost learns complex patterns combining semantic similarity, keyword matches, duration constraints, and test type preferences.

URL Normalization: Unified slug extraction and alternate URL generation eliminated train set mismatches.

6 Conclusion

The system achieves **61.56% Mean Recall@10** through semantic search, hybrid retrieval, and XGBoost re-ranking. The solution is production-ready with 389 assessments, proper error handling, and comprehensive evaluation, demonstrating strong problem-solving and programming skills.