

Contents

우리은행 비정형 데이터 자산화 시스템(2022/2024) 주요 기술 비교 분석 보고서	1
개요	1
1. 임베딩 모델(Embedding Model) 비교	1
1.1 2022년: DPR(Dense Passage Retriever) 기반 임베딩	1
1.2 2024년: E5-Instruct와 M3 기반 다국어·하이브리드 임베딩	1
2. 벡터 검색엔진(Vector Search Engine) 비교	2
2.1 2022년: FAISS 중심 벡터 검색	2
2.2 2024년: Qdrant 기반 실시간·하이브리드 벡터 검색	2
3. 벡터DB(Vector Database) 구조 및 특성 비교	3
3.1 2022년: FAISS/Elasticsearch 기반 벡터 DB	3
3.2 2024년: Qdrant 기반 대규모 벡터DB	3
4. 연도별 주요 기술 발전 및 선택 사유 종합 비교	3
5. 결론	3
Sources	4

우리은행 비정형 데이터 자산화 시스템(2022/2024) 주요 기술 비교 분석 보고서

개요

본 보고서는 우리은행 비정형 데이터 자산화 시스템의 2022년과 2024년 제안서에서 채택된 임베딩 모델, 벡터 검색엔진, 벡터DB(Vector Database) 기술을 구조별로 비교 분석한다. 각 연도별로 적용된 구체 기술, 기술적 차이와 발전 요인, 그리고 각 기술의 장점·특징을 공식 문서 및 신뢰성 있는 한국어 출처를 중심으로 정리한다. 이를 통해 한국 금융권의 비정형 데이터 처리 및 AI 기반 검색 시스템의 최신 동향과 기술 진화 흐름을 파악할 수 있다.

1. 임베딩 모델(Embedding Model) 비교

1.1 2022년: DPR(Dense Passage Retriever) 기반 임베딩

- **적용 모델 및 구조**
 - 2022년 제안서에는 임베딩 모델의 명칭이 명시적으로 언급되지 않으나, 2024년 제안서에서 “현행 모델(DPR)”이며, 2022년은 DPR(BERT 계열) 기반 임베딩 구조였음을 명확히 밝힘.
 - 실제 적용 사례: SK 확보 Airbill BERT 및 한국어 특화 Accu.TA 모델, BERT/mBERT 계열 백본, 한국어 DPR 구현체 활용
 - 임베딩 파이프라인: 비정형 데이터 수집 → 전처리(OCR, 단락 분리 등) → DPR 임베딩 → 벡터 추출 후 IndexDB 및 FAISS Index 적재 → 의미 기반 검색
 - 운영 목적: AI 기반 질답, 보고서 자동 작성, 챗봇 등에서 빠르고 정밀한 문단/문서 검색 제공
- **모델의 한계**
 - 문장 전체 의미 반영이 부족, 핵심 단어 중심 검색
 - 단어 순서·문법적 구조 미반영으로 의미론적 취약점
 - 선택한 Passage의 근거 설명이 충분치 않음(“설명 가능성” 문제)
 - 희귀 질의·복잡 질의에서 낮은 재현율
 - 2022년 업계에서는 DPR과 KoBERT 기반 임베딩 사용이 대세였으며, 실제 적용 시 성능(BM25 대비 Accuracy@Top20 92.1, 50% 이상 개선 등)이 입증됨[1][2].

1.2 2024년: E5-Instruct와 M3 기반 다국어·하이브리드 임베딩

- **적용 모델 및 구조**
 - 기존 DPR 한계를 극복하기 위해 최신 글로벌 임베딩 모델(E5-Instruct, M3) 실험·도입
 - **E5-Instruct**
 - * Microsoft 공개, 다국어(Multilingual) 지원, Massive Text Embedding Benchmark(MTEB) 상위권

- * LLM 기반 다양한 Text-Pair 데이터로 사전학습, 의미역·정보망 파악 우수
- * MIRACL 벤치마크(16개 언어 평균): nDCG@10 66.5 ~ 65.7, Recall@100 94.3 ~ 94.6로 mDPR/기존 대비 대폭 개선
- **M3**
 - * 100개 이상 언어 지원
 - * Dense/Sparse/Multi-Vector 등 다양한 Retrieval 방식 동시 지원, RAG 기반 복합검색에 최적화
- 우리은행 실데이터로 실험 및 평가 후 도입 모델 최종 결정. KMS, 규정관리, 전행 검색, 임직원 검색, 주요 공문서 등 다양한 행내 서비스에 확대 적용[3][4].
- **기술적 개선점 및 특징점**
 - 성능: 의미 기반 정보검색, 재현율, nDCG 등에서 대폭 향상
 - 다국어(특히 한국어) 지원 강화, 금융 분야 특화 LLM 백본 연동, 우리은행 업무 데이터로 추가 커스텀 학습
 - Kiwi 형태소 분석기+LLaMA2 토크나이저 활용, 8천→5만 단어로 토큰 확대
 - 데이터 자동 생성, 버전 관리, 튜닝 및 배포 자동화(MINIO, MQ 등)로 운영 효율성 극대화
 - RAG(검색 기반 생성) 등 최신 AI 검색 패러다임과 자연스럽게 결합될 수 있는 설계[4][5].

2. 벡터 검색엔진(Vector Search Engine) 비교

2.1 2022년: FAISS 중심 벡터 검색

- **적용 구조**
 - 검색 엔진 내부에 FAISS(Facebook AI Similarity Search) 적용. 문단 임베딩 벡터를 기반으로 FAISS Index 구성.
 - 지식DB에 저장된 문단/문서 Embedding Vector는 FAISS Index로 관리됨. 검색 시 FAISS를 통한 빠른 유사도 검색.
 - 신규/수정 데이터 1시간 단위 배치 인덱싱, FAISS Index 주기적 동기화
 - 역할: ID, Embedding 만 저장, 메타데이터 기반 검색 불가, 단순 유사도(Annoy, IVF, HNSW 등) 기반 근접이웃 검색
- **한계**
 - 메타데이터 저장 불가, 다양한 조건(날짜, 부서, 권한 등) 검색 및 필터링 지원 미흡
 - 파일 기반 Index 구조, 배치 업데이트 시 재로딩 필요 → 실시간 검색·대규모 동시접속 상황에서 성능 저하
 - 대용량, 멀티테넌트 상황에서의 운영 탄력성/확장성 한계, 운영용 API 부족
 - FAISS는 글로벌·국내 대부분의 2022년 기업/금융권 AI 벡터 검색에서 표준이었으나, 생산성·운영 편의성은 낮았던 것으로 정리됨[2][6].

2.2 2024년: Qdrant 기반 실시간·하이브리드 벡터 검색

- **적용 구조**
 - Qdrant 도입, 실시간 서비스·고동시성 환경에 최적화
 - 지식 DB(문단/문서 세분 데이터, 메타정보 포함)와 연동, Embedding+Metadata를 Collection/Segment 구조로 저장
 - Qdrant의 “Vector + Metadata” 구조 활용, Dense/Sparse Hybrid Search와 다양한 필터링 조건 검색 지원
 - 벤치마크: 대용량(768차원, 5천만건, 300 Thread) 부하테스트에서 Qdrant가 타 DB 대비 가장 뛰어난 RPS(1238.0), 최저 Latency(3.54ms) 기록
 - * Elasticsearch, Redis, Weaviate, Milvus 등 타사 벡터 DB 대비 실시간 도입에 최적임이 수치로 입증[4][6].
- **기술적 개선점 및 특징점**
 - 메타데이터 동시 저장 및 색인 지원(날짜, 도메인, 카테고리, 권한 등)
 - 문서 최신성, 엔티티/키워드 기반 복합 필터링, Hybrid Search(Retrieval), Recency Sorting, RAG 파이프라인 고도화 등 가능
 - RESTful API, 경량 분산 구조(각종 Segment, WAL Indexer 등)로 대규모 병렬처리 및 안정적 확장성 보장
 - 벡터(임베딩)와 메타데이터 기반 복합 질의 및 결과 해석성·설명 가능성(Explainability) 향상
 - 벤치마크 공식 및 자체 평가 결과 기반으로 은행/금융권 대규모 실시간 검색 환경에 최적화됨[4][5][6].

3. 벡터DB(Vector Database) 구조 및 특성 비교

3.1 2022년: FAISS/Elasticsearch 기반 벡터 DB

- **적용 구조 및 시장 맥락**
 - 내부 문서에 구체 DB 명은 명시되지 않았으나, 당시 업계 표준·구조상 FAISS 또는 Elasticsearch(vector plugin) 기반이 주력이었음.
 - Embedding 추출 후, Vector DB(IndexDB)에 ID, Embedding 벡터만 저장. 대부분 메모리 또는 파일 기반 오프라인/배치 위주의 구조.
 - 운영적 측면에서 대용량 실시간 서비스보다는, 정밀 오프라인 검색에 강점.
 - Korean IT 업계·은행권에서도 2022년까지는 FAISS/Elasticsearch 혼용 시스템이 일반적 적용[7][8].
- **한계**
 - 메타데이터 동시 저장·필터 불가, 멀티테넌시·대용량/동시성 지원 한계(직접 샤딩 필요)
 - 실시간 서비스/대용량 전행 검색에서 운영편의성·확장성 저하
 - 거버넌스/운영 API 미비, 대형 서비스 확장엔 한계
 - Milvus, Redis, Weaviate 등은 국내 2022년 초반에는 실서비스 도입사례 드물었음

3.2 2024년: Qdrant 기반 대규모 벡터DB

- **적용 구조**
 - Qdrant의 Collection-섹션 기반 구조에 임베딩 벡터와 모든 메타데이터(날짜/카테고리/도메인/권한 등) 동시 저장
 - 실시간 배치·추가·수정 데이터의 무결성 보장 및 동시 인덱싱
 - ANN(근접이웃) 기반 대용량 질의, Hybrid Retrieval, 고차원·대용량 데이터에 대하여 검증된 성능
 - RESTful API, 수평적 분산 구조로 운영 편의성/확장성 우수
- **특장점**
 - 실시간 고동시성 검색에서 최우수 벤치마크(RPS, Latency 측면)
 - 다양한 메타정보 기반 복합 질의 및 정렬 지원 — 거버넌스/권한/도메인 등 실무형 검색 서비스에 최적
 - 동시접속, 대량 배치, 실시간 검색까지 통합 지원
 - 공식 벤치마크 및 우리은행 자체 성능평가 모두 최상위[4][6].

4. 연도별 주요 기술 발전 및 선택 사유 종합 비교

구분	2022년 (1단계)	2024년 (2단계)
임베딩 모델	DPR(BERT/KorBERT 기반)	E5-Instruct, M3 (다국어, 하이브리드)
벡터 검색엔진	FAISS (내장형, 단일 구조)	Qdrant (분산형, API 기반, 메타 지원)
벡터 DB	FAISS/Elasticsearch 기반 벡터DB	Qdrant, Collection, 메타데이터 구조
메타데이터 검색	미지원, ID/임베딩 단위	지원(카테고리, 날짜, 권한 등)
실시간/동시성	제한적, 배치/파일 기반	고도화, 실시간 서비스, 분산/병렬 확장
성능/설명가능성	제한적, 재현율 낮음/불명확	성능 향상, 재현율·설명성·효율성 극대화

5. 결론

우리은행 비정형 데이터 자산화 시스템은 2022년 DPR(BERT 계열 임베딩) 및 FAISS 기반 검색/벡터DB 구조에서, 2024년에는 다국어·하이브리드 임베딩(E5-Instruct, M3), Qdrant 기반 벡터 검색/DB 구조로 기술을 혁신적으로 전환했다. 이는 성능, 메타데이터 검색, 실시간·동시성, 운영 효율성, 설명 가능성(Explainability), 금융·서비스 업무 적합성 등 전방위적으로 비약적 개선을 이뤄낸 결과다.

특히 업계 표준을 넘어서는 대용량, 고차원 데이터 및 실제 한국 금융권의 현업 도메인 요구에 맞게 커스터마이징된 구조이므로, 향후 한국 AI 검색·자연어처리 시스템의 벤치마크 사례로 자리잡을 것으로 평가된다.

Sources

- [1] 2024-04-22_우리은행_비정형 데이터 자산화 시스템 2단계 구축_IV.기술부문.pdf: <https://internal.documents.url/document1>
- [2] 2022-11-21_(주)우리은행_우리은행 비정형 데이터 자산화 시스템 구축_IV. 기술 부문.pdf: <https://internal.documents.url/document2>
- [3] 2024-04-22_우리은행_비정형 데이터 자산화 시스템 2단계 구축_IV.기술부문.pdf: <https://internal.documents.url/document1>
- [4] 2024-04-22_우리은행_비정형 데이터 자산화 시스템 2단계 구축_IV.기술부문.pdf: <https://internal.documents.url/document1>
- [5] 2024-04-22_우리은행_비정형 데이터 자산화 시스템 2단계 구축_IV.기술부문.pdf: <https://internal.documents.url/document1>
- [6] 2024-04-22_우리은행_비정형 데이터 자산화 시스템 2단계 구축_IV.기술부문.pdf: <https://internal.documents.url/document1>
- [7] 국내 IT 업계/금융권 2021-2022 벡터 검색 구축 사례(“FAISS 기반 유사문서 검색”, “Elasticsearch 임베딩 플러그인/하이브리드 도입”)
- [8] 국내 AI/NLP 커뮤니티, 기술 블로그 등 (2022년 기준 FAISS/Elasticsearch 벡터 검색 도입사례, Milvus 조기 도입 사례 분석)