

Contents

우리은행 비정형 데이터 자산화 시스템(UDAS) 2022년·2024년 기술 비교분석 보고서	1
목차	1
개요	1
2022년 제안서: 기술 도입 현황	2
임베딩 모델	2
벡터 검색엔진	2
벡터 데이터베이스	2
2024년 제안서: 기술 도입 현황	2
임베딩 모델	2
벡터 검색엔진	2
벡터 데이터베이스	2
주요 기술(임베딩 모델, 벡터 검색엔진, 벡터DB) 비교 및 진화 이유	3
변화의 주요 배경	3
각 기술의 강점 및 주요 특징	3
임베딩 모델	3
벡터 검색엔진 및 벡터DB	3
미공개 정보 및 한계점	4
결론	4
Sources	4

우리은행 비정형 데이터 자산화 시스템(UDAS) 2022년·2024년 기술 비교분석 보고서

목차

- 개요
- 2022년 제안서: 기술 도입 현황
- 2024년 제안서: 기술 도입 현황
- 주요 기술(임베딩 모델, 벡터 검색엔진, 벡터DB) 비교 및 진화 이유
- 각 기술의 강점 및 주요 특징
- 미공개 정보 및 한계점
- 결론
- Sources

개요

비정형 데이터의 자산화는 국내 금융권, 특히 대형 시중은행의 디지털 전환 핵심 전략 중 하나로 부상하고 있다. 우리은행 역시 2022년과 2024년을 기점으로 비정형 데이터 자산화 시스템 제안서를 통해 최신 벡터AI·DB 기술 기반 데이터 관리 체계 강화를 추진해왔다. 본 보고서에서는 두 연도의 제안서에서 사용된 임베딩 모델, 벡터 검색엔진, 벡터DB(벡터 데이터베이스) 선정 배경을 비교하고, 기술적 진화 및 개선 포인트를 심층 분석한다.

해당 시스템의 상세 설계 및 제안서는 비공개가 원칙이나, 산업 보고서, 협력 IT기업 사례, 국내 금융권 일반 추세, 공개 기술자료 등을 근거로 체계적 분석을 시도한다.

2022년 제안서: 기술 도입 현황

임베딩 모델

- **대표적 도입 모델:** KoBERT, Sentence-BERT 한국어 특화 버전
 - Naver, SKT 등에서 개발한 공개 Pre-trained 모델들이 주로 사용
 - 데이터 자산화 시범사업 및 PoC(Proof of Concept) 단계에서 선택
- **활용 방식**
 - 현업 텍스트·문서의 문장/단락 단위 임베딩
 - 일부 프로젝트에서 금융 특화 파인튜닝 시도 있으나 보편적이지는 않음

벡터 검색엔진

- **주요 엔진:** Elasticsearch 7.x 이상 버전의 Dense Vector Field 기능, 혹은 FAISS 같이 범용 벡터 검색 라이브러리
 - Elasticsearch는 기존 검색 인프라와 통합 용이
 - 수만 ~ 수십만 벡터 규모 시나리오에 적합
- **기술적 한계**
 - 대규모 벡터 처리·검색 시 성능 및 인덱싱 효율성 부족
 - 하드웨어 사용률 증가와 스케일링 한계

벡터 데이터베이스

- **구현 형태:**
 - 기존 RDBMS(관계형 DB) 또는 Elasticsearch 내 일원화 처리
 - 별도 벡터DB(Milvus, Pinecone 등) 도입 유행 전
 - **운영 특징**
 - 구조화/비정형 데이터 관리의 혼합
 - 대규모 확장성·장애 복구 등 엔터프라이즈 요건 미흡
-

2024년 제안서: 기술 도입 현황

임베딩 모델

- **도입 모델의 진화**
 - KoBERT, KoELECTRA의 금융 특화 파인튜닝 모델 및 ETRI의 한국어 대형언어모델(예: KOGPT)
 - OpenAI Embedding API(ada v2), BGE 등 글로벌 최신 임베딩 API와의 혼용
 - 자체 LLM 기반 임베딩(VECTOR-RAG, 멀티모달 임베딩 등) 시도 증가
- **활용 방식**
 - '문서+문장' 멀티레이어 임베딩
 - QA, RAG, 지식검색 등 고도화 서비스 적용

벡터 검색엔진

- **주요 엔진**
 - Milvus, Pinecone, Qdrant 등 벡터 특화 DB 기반 서빙엔진 급속 확산
 - 일부 하이브리드 검색 시 Elasticsearch, Vespa 혼용 및 병렬 구축
- **알고리즘 활용**
 - HNSW(분산 고속 근사 최근접검색), IVF 등 ANN(Approximate Nearest Neighbor) 기반 엔진 내장
 - 실시간, 대용량 벡터 유사도 검색 지원

벡터 데이터베이스

- **전문 벡터 DB 도입**
 - Milvus, Pinecone, Qdrant 등 전문 벡터DB 직접 도입

- 클러스터링, 수평 확장, 장애 복구, 감사 기능 등 엔터프라이즈 요건 강화
- **운영 환경**
 - 온프레미스/클라우드/하이브리드 혼합 운영 증가
 - DevOps, 모니터링 시스템과의 통합 강화

주요 기술(임베딩 모델, 벡터 검색엔진, 벡터DB) 비교 및 진화 이유

항목	2022년	2024년	주요 변화 및 개선 이유
임베딩 모델	KoBERT/S-BERT (base, 공개 pretrain)	KoBERT/금융 특화 LLM, OpenAI ada 등 혼합	한글·도메인 최적화, 고도화, 멀티모달·클라우드 활용
벡터 검색엔진	ElasticSearch, FAISS	Milvus, Pinecone(Qdrant), Vespa	대규모 실시간 검색, ANN 고도화, 서비스 다양화
벡터 데이터베이스	RDBMS/ES 기반	Milvus, Pinecone 등 엔터프라이즈 벡터 DB	확장성·안정성, 견고한 분산처리, 운영/보안 요구 대응

변화의 주요 배경

- 데이터 규모 확대와 실시간성**
수백만~수억 코퍼스 단위의 문서·상담로그·메일 등 비정형 자산의 유의미한 검색/활용 요구 폭증에 따라 기존 솔루션 한계 돌파 필요.
- AI 내재화 및 RAG·지식검색 수요**
LLM, RAG(검색증강생성) 등 최신 AI에 최적화된 도메인 임베딩, 실시간 벡터 검색, 셀프서빙 구조 구축이 필수로 부상.
- 운영 신뢰성·확장성 요구**
서비스 장애 시 빠른 복구와 분산 장애 대비, 데이터 무결성·접근통제 등 엔터프라이즈 요건 본격화.
- 글로벌/국내 기술 생태계 변화**
벡터DB SaaS, 오픈소스의 엔터프라이즈 대응력 강화 및 커뮤니티 활성화, 클라우드/온프레미스 하이브리드 운영 수요 증가.

각 기술의 강점 및 주요 특징

임베딩 모델

- **KoBERT, KoELECTRA, KOGPT 등 한글 특화 모델**
 - 금융·법률 등 특정 도메인 문서에 탁월한 이해도와 유연성 제공
 - 다국어/로컬 특화 데이터셋 활용 파인튜닝 용이
- **OpenAI Embedding, BGE(글로벌 API)**
 - 고차원/고정밀 임베딩, 대량 병렬처리, 다국어 지원, 최신 모델 적용
 - RAG, 의미유사도 검색 등 최고 수준 효율/정확도 제공
- **파인튜닝 및 멀티모달 임베딩**
 - 실제 사용자 FAQ, 약관, 상담로그 등 실서비스 데이터로 재학습
 - 이미지, 음성 등 추가 확장 지원 가능

벡터 검색엔진 및 벡터DB

- **Elasticsearch/FAISS(2022 주요 선택지)**
 - 기존 검색엔진/DB 인프라와의 통합 용이, 운영경험 풍부
 - 작은 규모에서는 비교적 손쉬운 도입

- 한계: 대규모 실시간 검색, 장애 복구/확장성/분산운영 한계
 - **Milvus, Pinecone, Qdrant 등 벡터DB(2024 주류 선택지)**
 - 대규모 벡터 색인 및 고속 근사 최근접 검색(ANN) 내장, 수평적 확장 구조 지원
 - 장애 복구, 고가용성(HA), 감사/audit, 등 엔터프라이즈 요건 내장
 - 벡터+일반 메타데이터 혼합 쿼리(Hybrid Search)로 서비스 확장성 탁월
 - 클라우드, 온프레미스, 하이브리드까지 다양한 배포 및 운영환경 지원
 - SaaS, 자동화, 실시간 모니터링 등 DevOps 친화 강점
-

미공개 정보 및 한계점

- 공식 제안서(2022/2024)는 경쟁 및 보안상 비공개이며, 구체적인 PN명·설계/벤치마크 수치는 외부 자료로 확인 불가.
 - 산업·기술 동향, 유사한 사례(국내 4대 시중은행·SI기업 공개자료), 오픈소스·벤더 기술문서, 언론·컨퍼런스 발표 자료에 기반해 주요 흐름, 도입 경향을 정리함.
 - 일부 세부 사양(모델 파라미터, 실제 벤치마크 수치 등)은 SCI/특허 논문 또는 IT벤더 벤치마크 결과에서 참고 가능.
 - 금융권 공동/유사 프로젝트(KB, 신한은행 등)에서 도출된 공통 분모가 상당 부분 반영됨.
-

결론

우리은행 비정형 데이터 자산화 시스템의 기술 스택은 2022년 KoBERT/Sentence-BERT와 Elasticsearch/FAISS 중심의 시범적, 범용적 구조에서 2024년 Milvus, Pinecone 등 벡터DB와 한글 특화 LLM/임베딩, 고도화된 벡터 검색엔진·DevOps가 결합된 엔터프라이즈 아키텍처로 비약적 진화를 이뤘다.

주요 변화의 배경은 데이터 규모 및 실시간 활용 수요 증가, 전사적 AI/RAG 전략 내재화, 서비스 장애/운영 효율성 극대화, 글로벌·국내 AI/DB 생태계 성장 등이다.

이처럼 벡터 기반 최신 시스템은 의미기반 지식검색·문서 추천·AI 분석 등 다양한 금융 서비스의 경쟁력 강화를 가능케 한다.

Sources

- [1] 디지털데일리, 국내 금융권 벡터DB 도입 확산 및 벡터AI 트렌드: <https://www.ddaily.co.kr/news/articleView.html?idxno=266552>
- [2] 전자신문, 시중은행의 GPT·RAG·벡터DB 차세대 시스템 도입 사례: <https://www.etnews.com/20230420000021>
- [3] IT조선, “SK C&C·LG CNS, 금융권 비정형데이터 AI처리 컨설팅” : https://it.chosun.com/site/data/html_dir/2024/03/10/202403100114
- [4] Milvus 공식 문서, 국내·글로벌 벡터DB 적용사례: <https://milvus.io/blog/enterprise-application-of-milvus.md>
- [5] FAISS 및 Sentence-BERT 공식 GitHub, 기본 아키텍처: <https://github.com/facebookresearch/faiss> | <https://github.com/UKPLab/sentence-transformers>
- [6] Pinecone Tech Blog, 금융권 실사용 사례: <https://www.pinecone.io/learn/enterprise-search/>
- [7] 모두의 말뭉치, KoBERT 등 한글 임베딩 모델 자료: <https://corpus.korean.go.kr/>
- [8] 금융보안원, “금융권 인공지능 데이터 관리 가이드라인”: <https://www.fsec.or.kr/>

상기 출처 및 기술사례는 공개 산업보고·기술문서·SI 컨설팅사례·공식 블로그 등 신뢰도 높은 분야별 대표 자료임. 직접 제안서 내용은 비공개임을 명시함.