


搜索型数据库白皮书

CCSA TC601 大数据技术标准推进委员会

2024 年 7 月



版权声明

本报告版权属于 CCSA TC601 大数据技术标准推进委员会，并受法律保护。转载、摘编或利用其它方式使用本报告文字或者观点的，应注明“来源：CCSA TC601 大数据技术标准推进委员会”。违反上述声明者，本推委会将追究其相关法律责任。

编写委员会

本报告的撰写得到了数据库领域多家企业与专家的支持和帮助，主要参与单位与人员如下。

❖ 主要编写单位（排名不分先后）：

大数据技术标准推进委员会、极限数据（北京）科技有限公司、拓尔思信息技术股份有限公司、中移（苏州）软件技术有限公司、星环信息科技（上海）股份有限公司、北京人大金仓信息技术股份有限公司、达梦数据技术（江苏）有限公司、腾讯云计算（北京）有限责任公司、中移（杭州）信息技术有限公司。

❖ 编写组主要成员（排名不分先后）：

刘蔚、刘思源、齐丹阳、马嘉慧、马鹏玮、闫树、姜春宇、曾勇、罗厚付、张磊、杨帆、廖石阳、李银松、肖锋、江敏、姜康、陶捷、秦华婵、刘永恒、王宝晗、刘熙、谭力鸣、李昆、张欣毅、尹强、王凯龙、刘俊、付新、张静修、鲁剑峰、高攀、任翔、黄超、徐良。

前 言

随着数据智能时代的到来，非结构化数据日益成为各类组织数据的增长主力，蕴含巨大价值，如何高效存储并利用非结构化数据，成为业内关注的焦点。以自动分词、倒排索引、相关度计算、向量检索引擎等技术为核心构建的搜索型数据库，作为高效处理非结构化数据的基础软件，逐渐成为数据智能时代的基石。

搜索型数据库结合全文检索引擎和 NoSQL 数据库的特点，相较于传统的关系型数据库，在处理大规模数据、实现快速检索和智能分析方面展现出显著的优势。近年来，搜索型数据库进一步集成了深度学习、向量搜索和 AI 大模型等尖端技术，使其不仅能够处理基础的文本搜索，还能胜任地理位置信息、图片、音视频等多模态数据的复杂检索任务。然而，搜索型数据库在性能优化、数据一致性保障以及安全性维护等方面仍面临挑战。

本白皮书全面介绍了搜索型数据库的基本概念、发展轨迹和技术架构，同时深入探讨了其在电子商务、金融、医疗健康等关键应用场景中的运用。结合当前搜索型数据库的发展现状和面临的主要问题，本白皮书提出了关于数据库的选择、应用实施的策略，并展望了搜索型数据库未来的发展方向和演进趋势，旨在为行业内外的相关利益方提供宝贵的参考和指导。

目 录

前 言	- 1 -
一、 搜索型数据库概述	- 2 -
(一) 搜索型数据库概念	- 2 -
(二) 搜索型数据库发展历程	- 3 -
(三) 搜索型数据库的挑战	- 6 -
二、 搜索型数据库技术详解	- 8 -
(一) 搜索型数据库核心技术	- 8 -
(二) 搜索型数据库关键特性	- 12 -
三、 搜索型数据库应用场景	- 14 -
(一) 业务场景	- 14 -
(二) 行业应用	- 17 -
四、 搜索型数据库能力模型	- 19 -
(一) 基础能力模型	- 19 -
(二) 能力模型的符合性验证情况	- 21 -
五、 搜索型数据库应用与实施建议	- 22 -
(一) 数据建模和索引设计	- 23 -
(二) 性能调优与监控	- 26 -
(三) 安全性考量	- 30 -
(四) 应用案例	- 30 -
六、 搜索型数据库演进趋势	- 32 -
(一) 多模态数据的融合处理和分析	- 32 -
(二) 人工智能与搜索型数据库的结合	- 33 -
(三) 云原生架构的演进	- 36 -
(四) 其它技术对搜索型数据库的影响	- 37 -
七、 总结和展望	- 38 -

图 目 录

图 1 DB-Engines 国际数据库排行榜	- 1 -
图 2 搜索型数据库的发展历程	- 3 -
图 3 DB-Engines 最新搜索引擎排名	- 6 -
图 4 搜索型数据库核心技术模块	- 8 -
图 5 搜索型数据库应用场景一览	- 14 -
图 6 搜索型数据库应用场景	- 19 -
图 7 搜索型数据库应用实施流程图	- 23 -
图 8 中国知识产权大数据与智慧服务系统	- 30 -
图 9 星环科技 Transwarp Scope	- 31 -

表 目 录

表 1 搜索型数据库与其它数据库对比	- 3 -
表 2 搜索型数据库产品	- 5 -
表 3 搜索型数据库业务场景	- 14 -
表 4 可信数据库-搜索型数据库基础能力测试通过名单	- 22 -