

# 《数据库发展研究报告（2025年）》 暨《中国数据库产业图谱（2025）》发布

何宝宏 | 中国信息通信研究院云计算与大数据研究所所长

# 目录 CONTENTS

01 数据库产业发展情况综述

02 数据库关键技术发展趋势

03 数据库行业应用情况综述

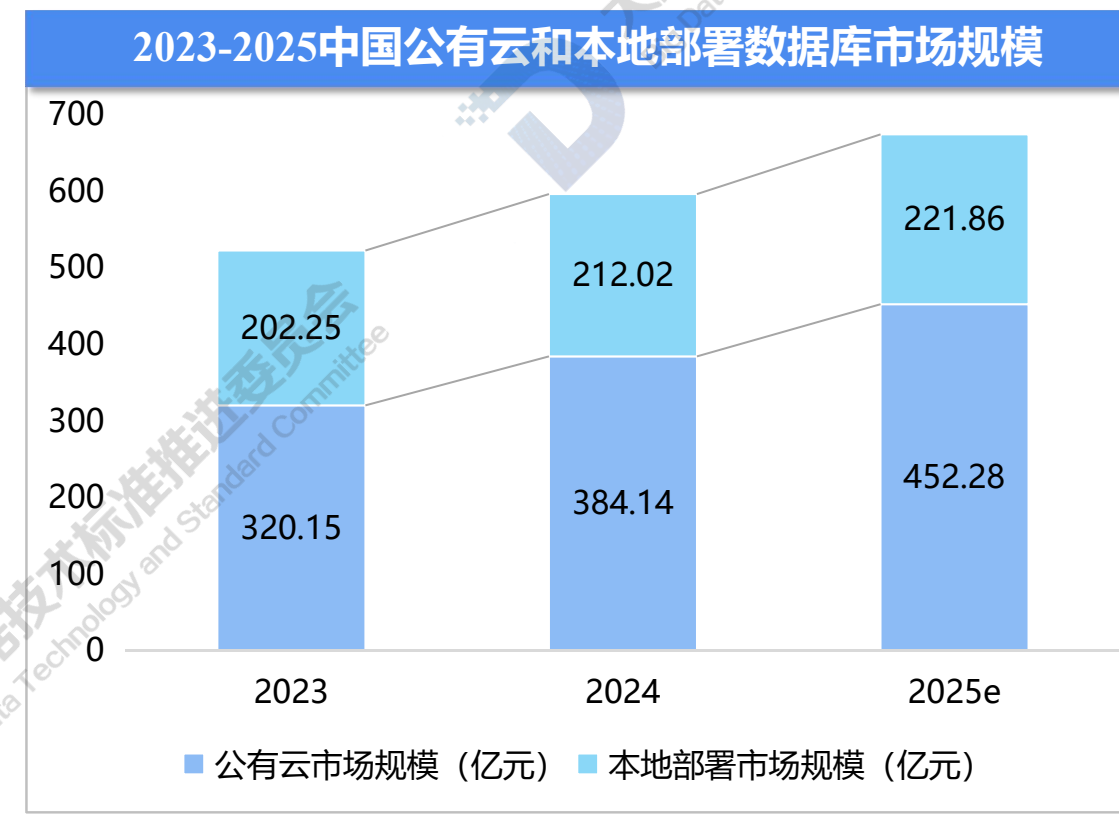
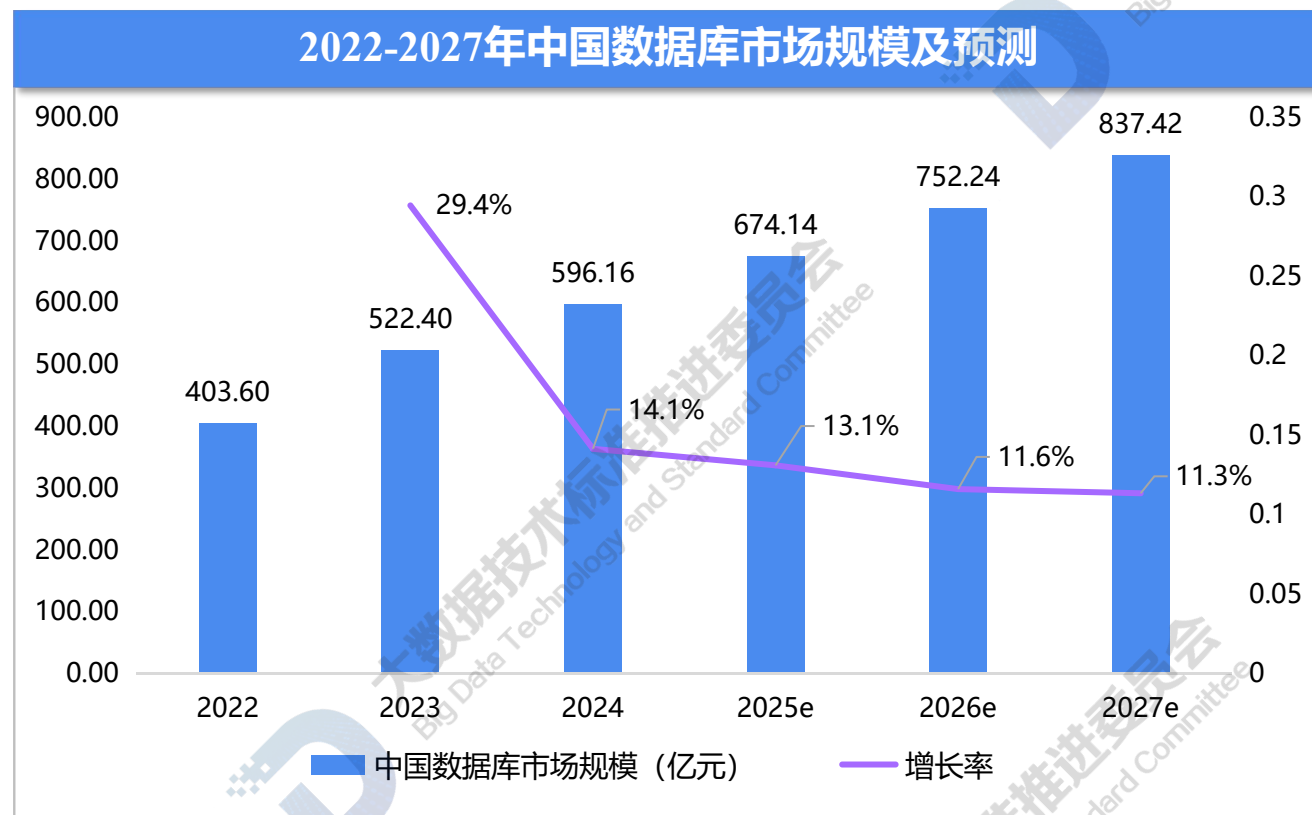
04 总结与展望

- 报告编制历时7个月
- 凝聚数据库领域百位专家智慧
- 全景呈现数据库行业现状及趋势
- 详细内容请阅览报告全文

# 洞察#1 从市场规模看，我国市场规模近600亿元，云上市场占比进一步扩大

◆ 据测算，中国数据库市场规模为83.7亿美元，约合596.16亿元人民币，占全球7.3%。预计到2027年，中国数据库市场规模将达到837.42亿元，市场年复合增长率（CAGR）为11.99%。

◆ 公有云和本地部署模式市场规模分别占总市场64.4%和35.6%，较2023年增速20.0%和4.8%；预计2025年公有云市场占比将进一步扩大达到67.1%。

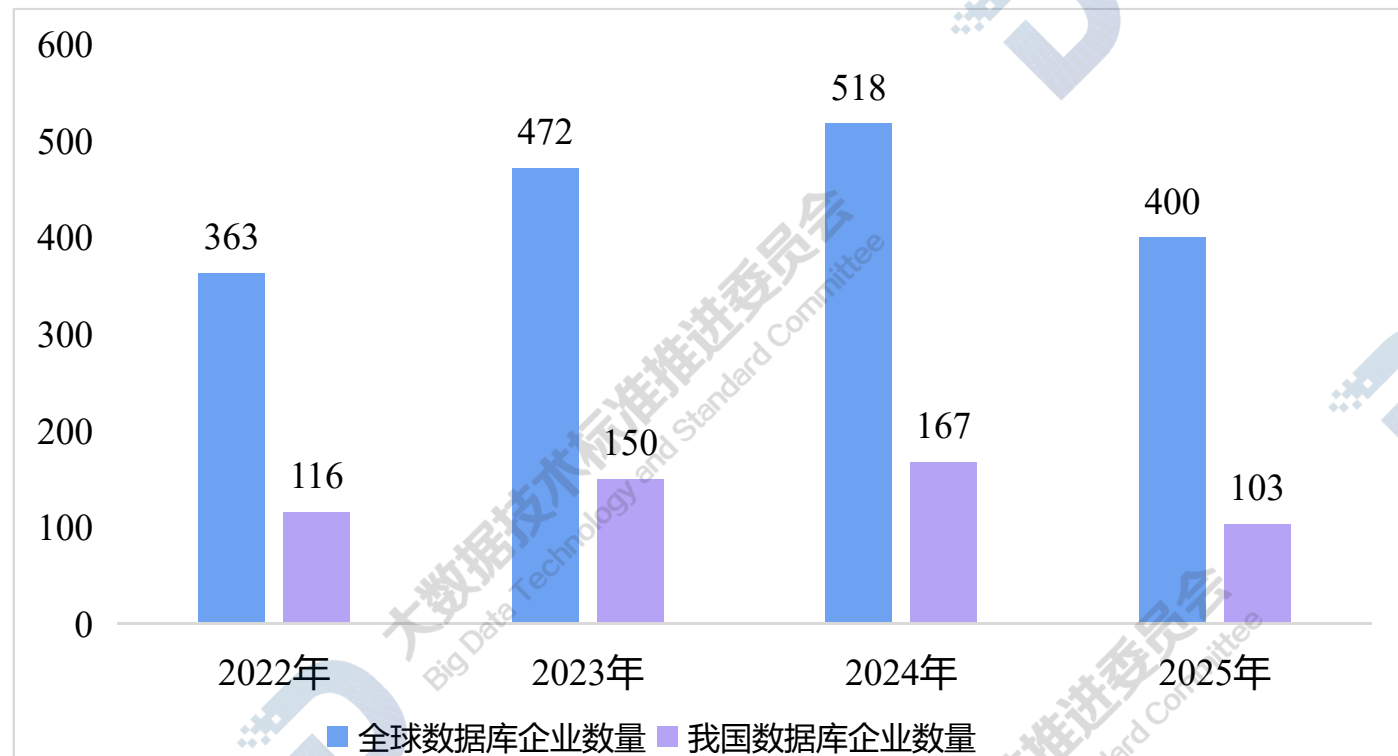


数据来源：CCSA TC601，2025年6月

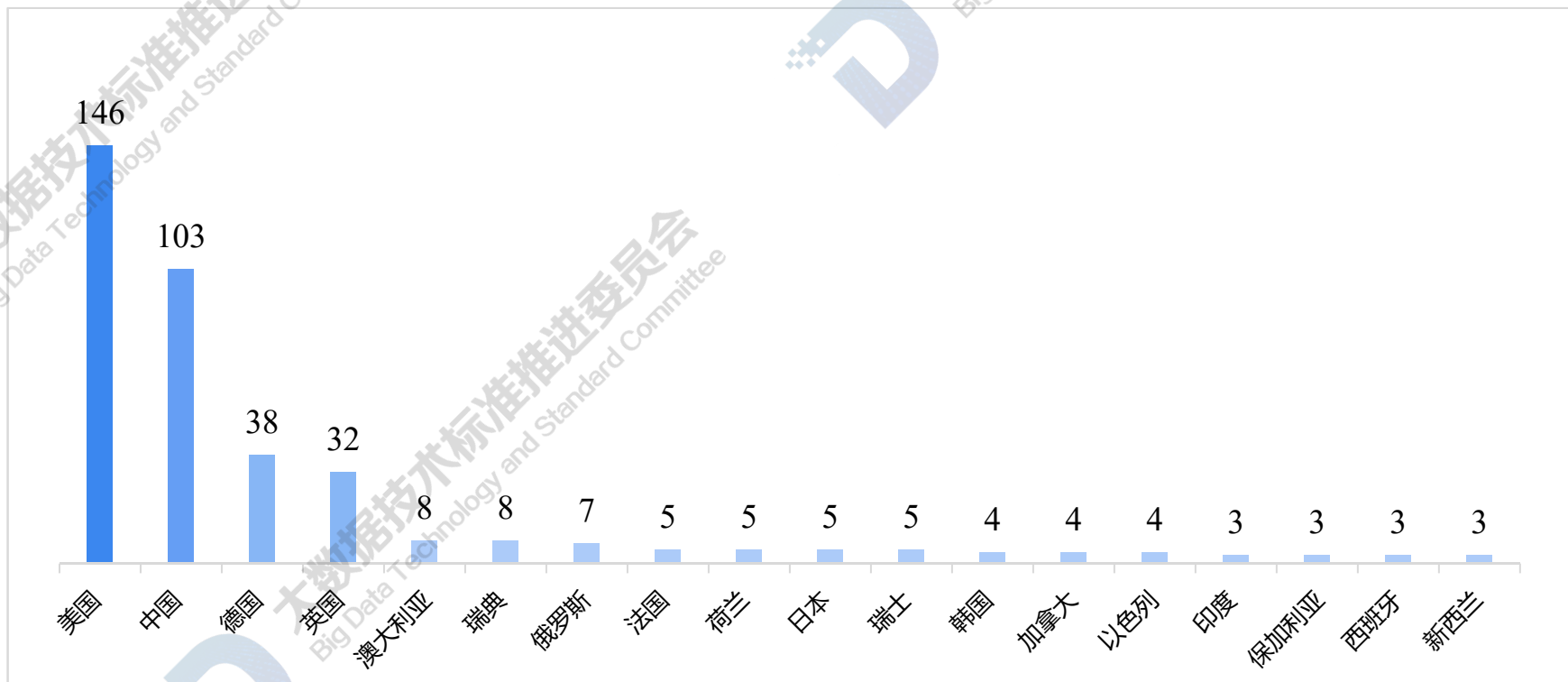
# 洞察#2 从竞争格局看，全球市场进入高质量发展阶段，我国数据库厂商数量显著收敛

◆ 截止2025年6月，全球及我国的数据库企业数量均有所回落，目前共有400家数据库产品提供商，其中美国和中国数据库厂商数量依旧处于全球领先地位，分别为146家和103家。

### 全球及我国数据库企业数量变化



### 全球数据库企业分布



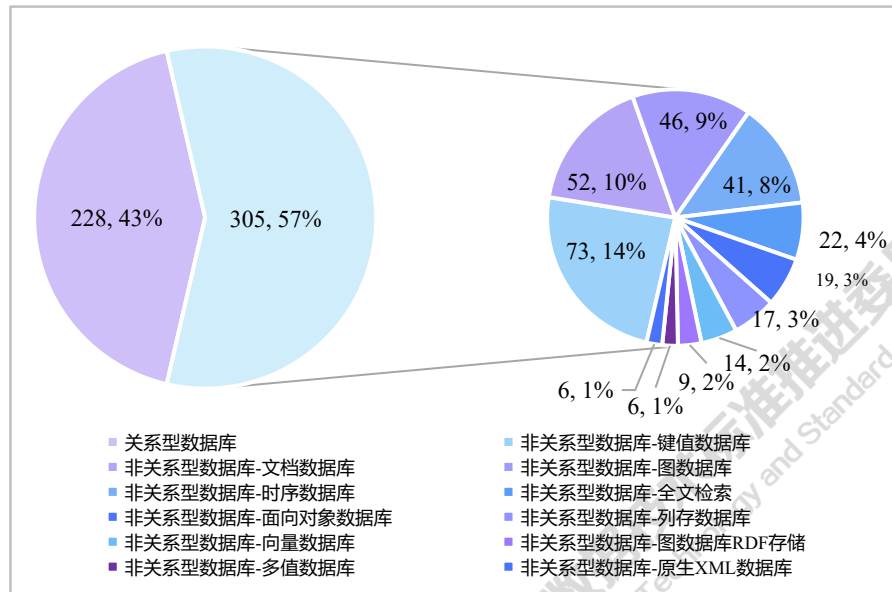
数据来源：CCSA TC601，2025年6月

# 洞察#3 从产品类型看，国内外分布各有侧重，向量数据库热度持续攀升

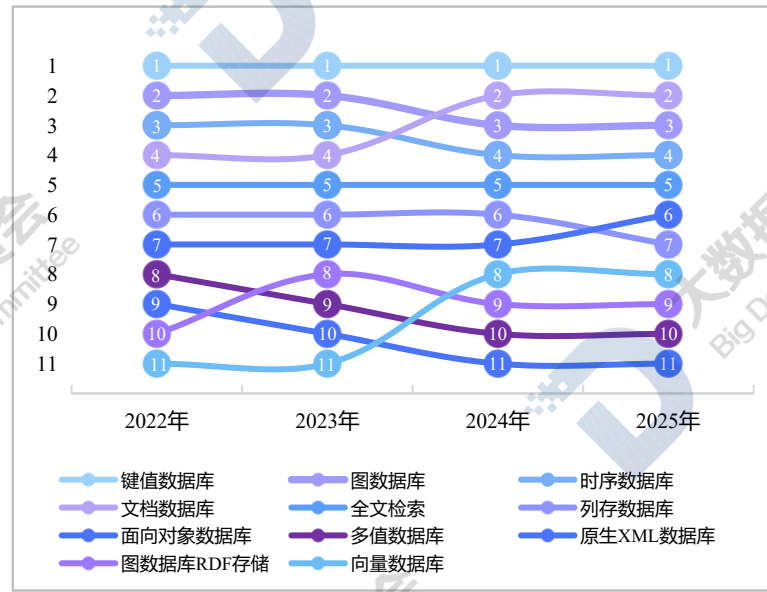
◆ 全球数据库产品数量整体分布呈现以非关系型及混合型数据库为主，关系型为辅的局面，键值数据库产品占比长期稳居第一。

◆ 与全球不同，我国数据库产品数量仍呈现以关系型为主，非关系型数据库为辅的局面，向量数据库产品热度逐年攀升。

### 全球数据库产品类型分布

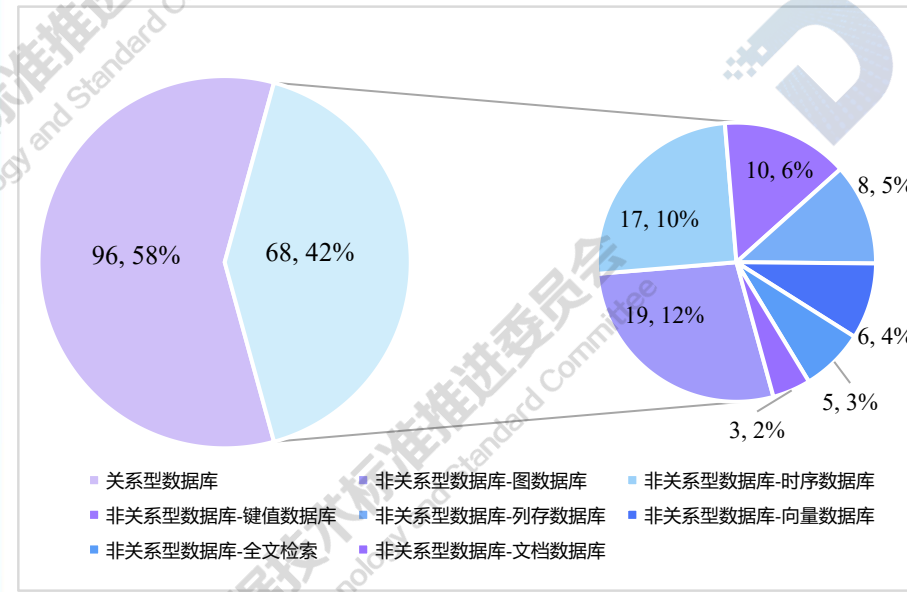


### 全球非关系型数据库数量占比排名变化

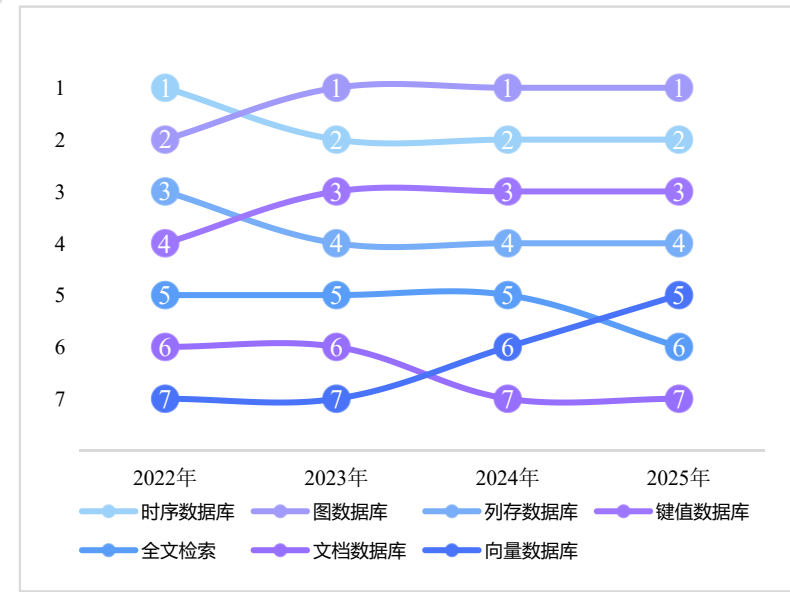


### 全球数据库产品类型

### 我国数据库产品类型分布



### 我国非关系型数据库数量占比排名变化



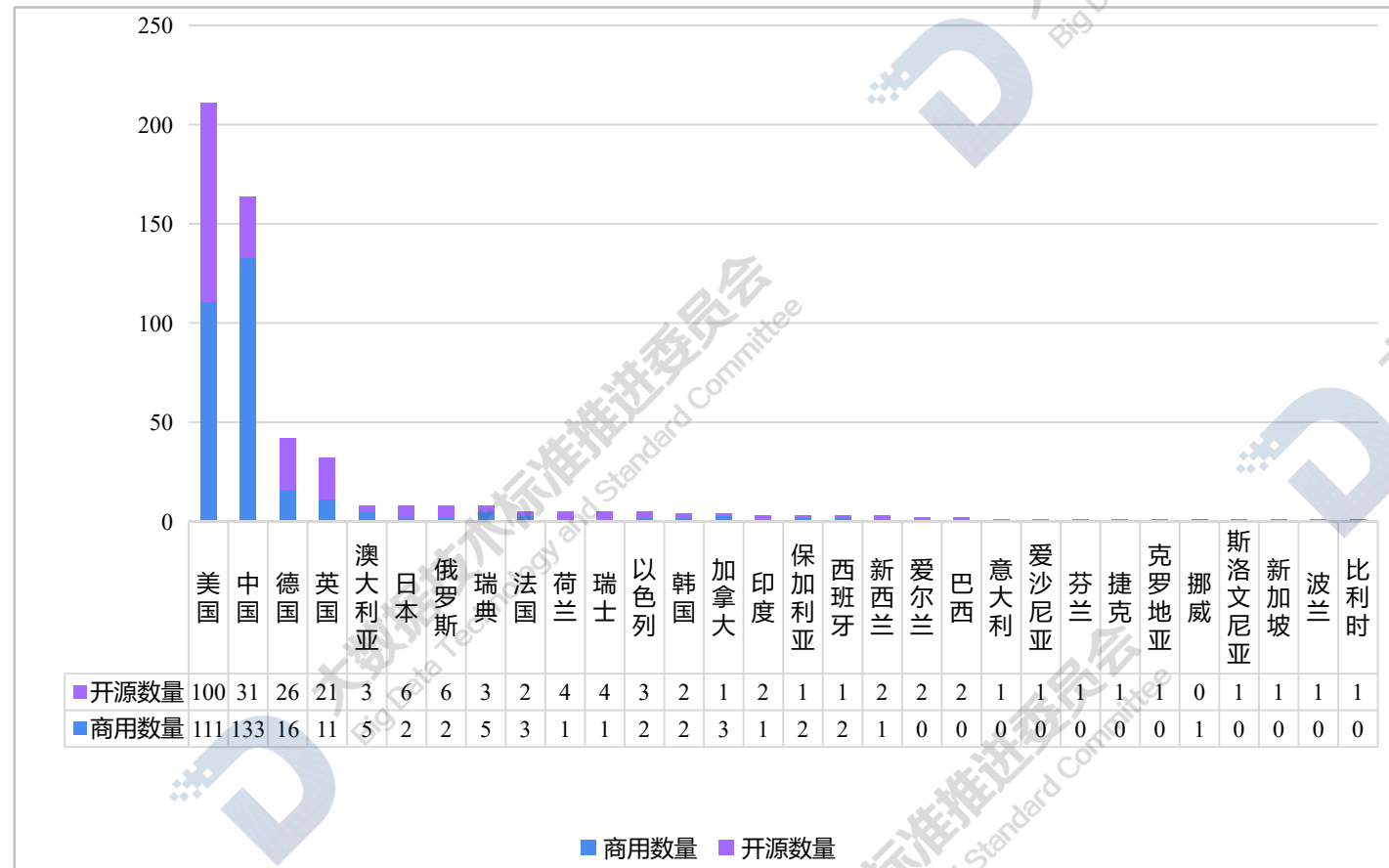
### 我国数据库产品类型

# 洞察#4 从商业模式看，商用数据库占据主导地位，PG生态获得投资青睐

◆ 欧美等国商用数据库占比略有上升，我国商用数据库占比较大，开源数据库发展前景广阔。

◆ PostgreSQL生态公司广受资本市场青睐，彰显其易用性、兼容性带来的强大吸引力，多云管理和“AI+”成为数据库投融资的焦点。

全球数据库产品商用开源对比



数据来源：CCSA TC601, 2025年6月

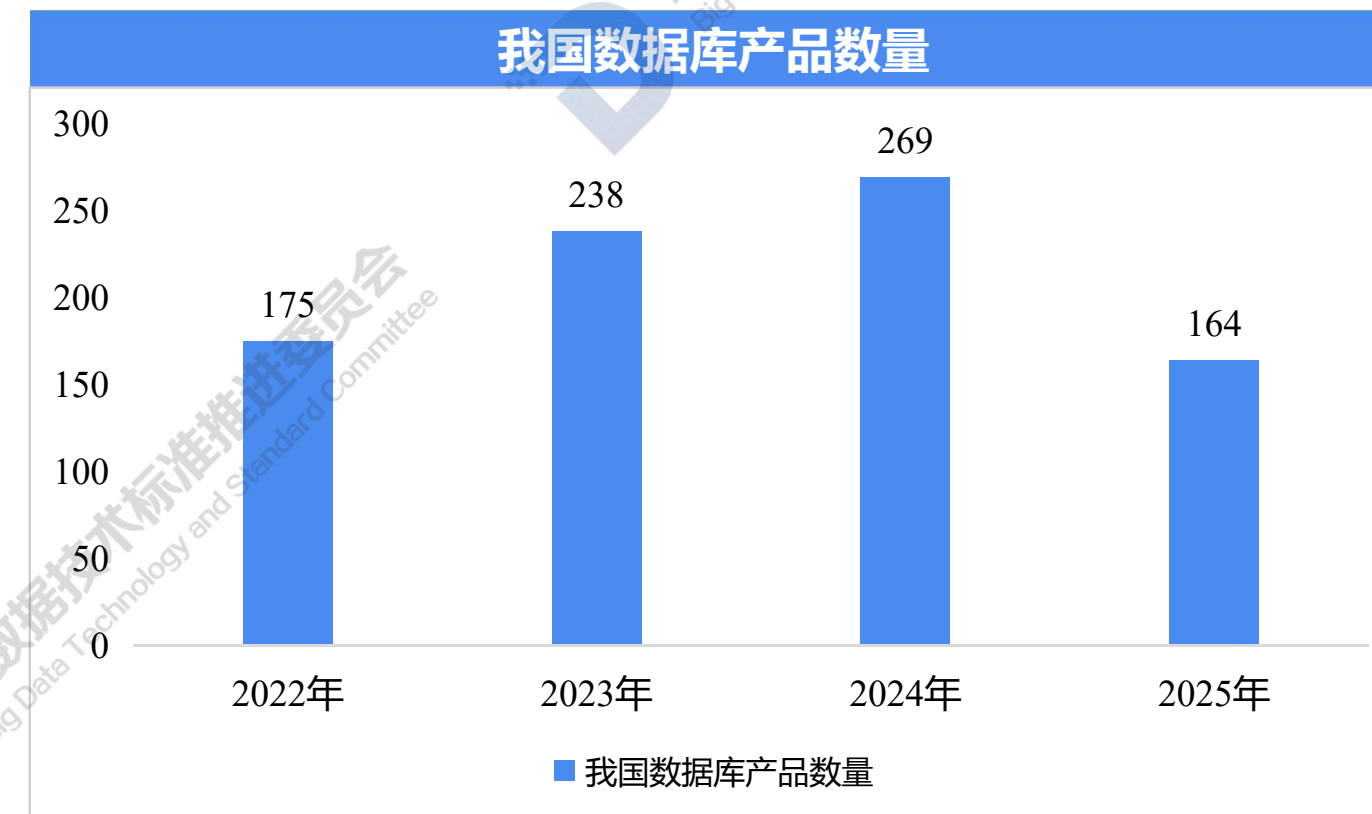
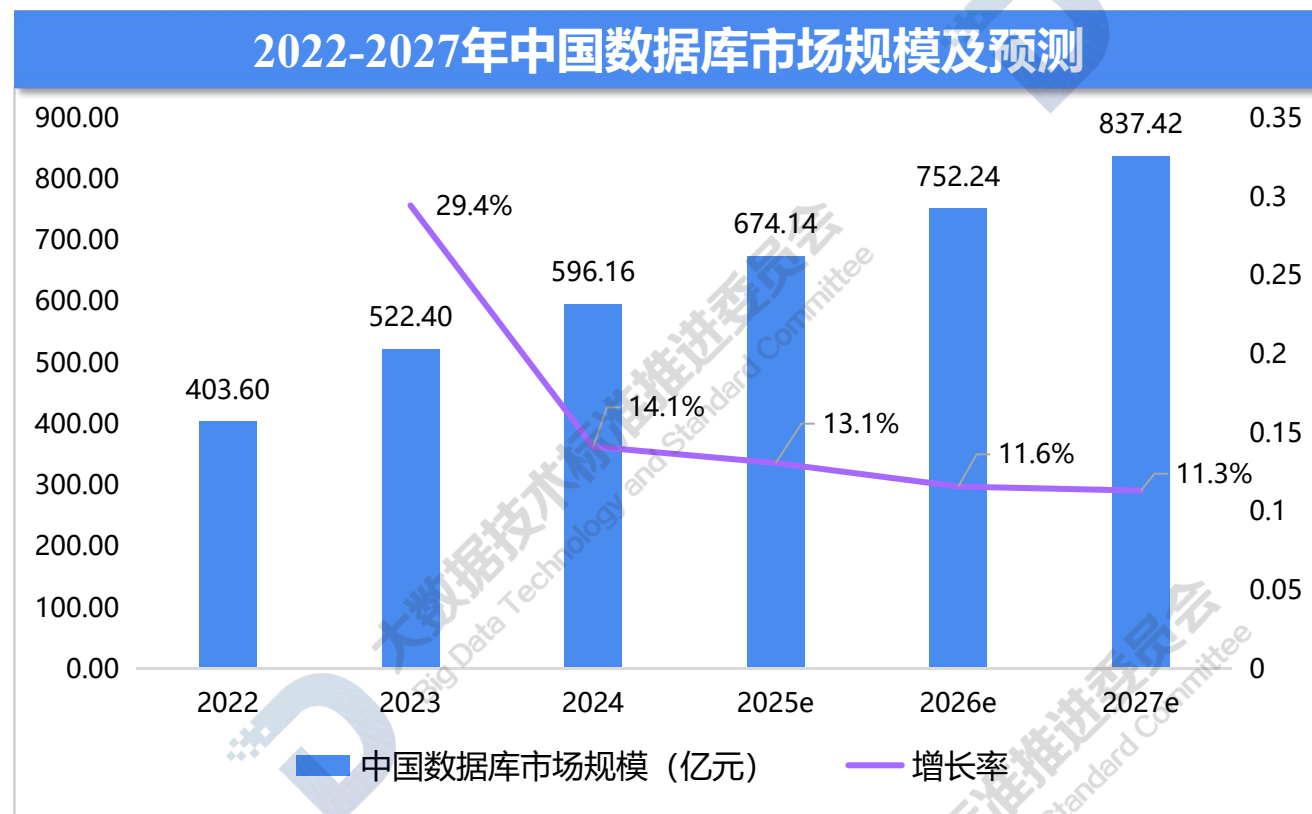
近期全球数据库企业投融资及收购情况

时间	企业名称	企业类型	涉及金额	资金用途
2025年4月	Tessell	多云数据库平台服务	6000万美金	推出基于AI驱动的对话式数据库管理服务
2025年4月	Supabase	PostgreSQL数据库云平台提供商	2亿美元	提升AI原生应用的开发体验和效率
2025年5月	Neon	Serverless Postgres托管服务商	10亿美元	Databricks收购Neon以加强其在多云环境中数据处理能力的战略布局
2025年6月	Crunchy Data	PostgreSQL的商业发行版公司	2.5亿美元	Snowflake收购Crunchy Data扩大其AI Data Cloud的产品版图

数据来源：依据公开资料整理

# 洞察#5 从我国市场看，市场规模持续扩大，市场格局逐步明晰

- ◆ 我国数据库市场头部效应明显，产业规模持续扩大，市场格局逐步显现。国产数据库从高速增长期进入高质量发展期，投资市场回归理性。
- ◆ 从我国数据库产品数量看，2022年至2024年，我国数据库产业规模逐步扩大，产品数量呈现迅速增长的趋势，由2022年的175款产品增加到2024年的269款产品，复合增长率约为24.0%；2025年，我国数据库市场规模持续扩大，但数据库产品数量收敛至164款。



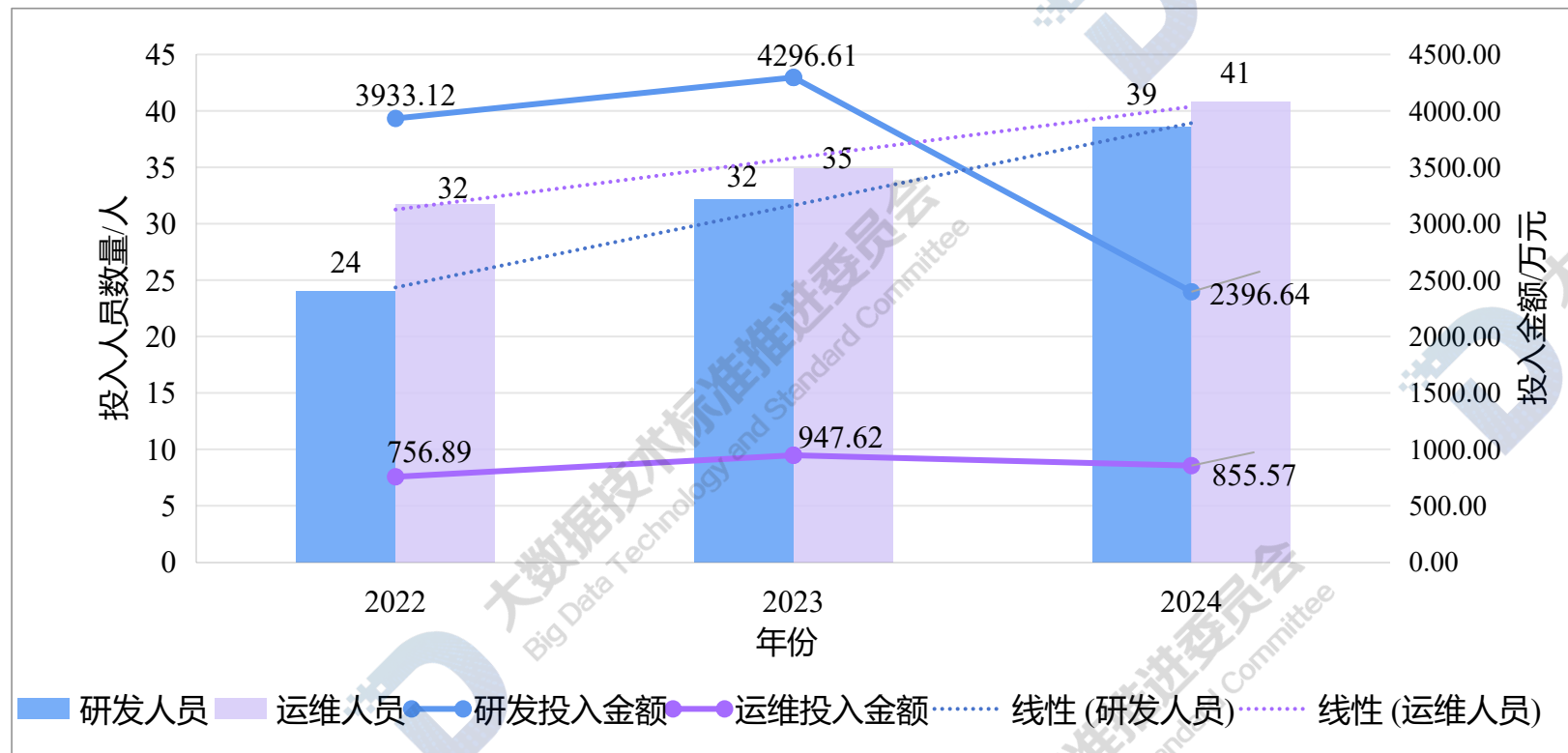
数据来源：CCSA TC601，2025年6月

# 洞察#6 从重点行业看，头部用户运维人员投入逐年增长

◆ 从重点行业看，头部用户运维人员投入逐年增长。根据“星河”数据库案例的统计，金融、电信、能源等行业的头部用户广泛应用了分布式系统，运维工作面临诸多挑战，人员投入逐年增长。

◆ 为了更好地指导业内分布式数据库系统运维管理工作，中国信通院联合多家企业研制了《分布式数据库运维管理能力成熟度模型》标准，已在多家机构落地贯标。

“星河 (Galaxy)” 案例2022-2024年研发/运维人员数量和投入资金变化



数据来源：CCSA TC601，2025年6月

分布式数据库运维管理能力成熟度模型贯标情况

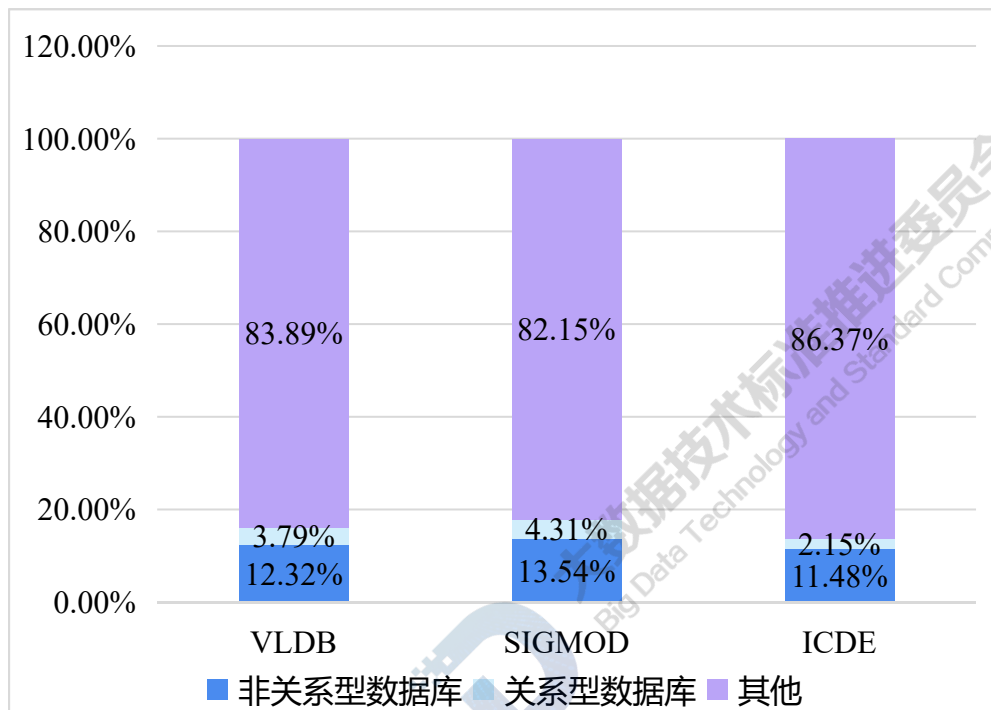
序号	企业名称	产品名称
1	中国移动通信集团浙江有限公司	浙江移动数据库运维保障体系
2	上交所信息网络有限公司	信息公司分布式数据库运维管理能力
3	中国移动通信集团北京有限公司	北京移动数据库运维管理能力
4	中国移动通信集团湖北有限公司	湖北移动信创数据库运维管理能力
5	中国移动通信集团陕西有限公司	陕西移动数据库运维管理能力
6	中国移动通信集团江苏有限公司	江苏移动数据库运维管理能力
7	中国移动通信集团上海有限公司	上海移动数据库运维管理能力

# 洞察#7 从学术创新看，多模融合成为研究重点，我国科研实力稳步增强

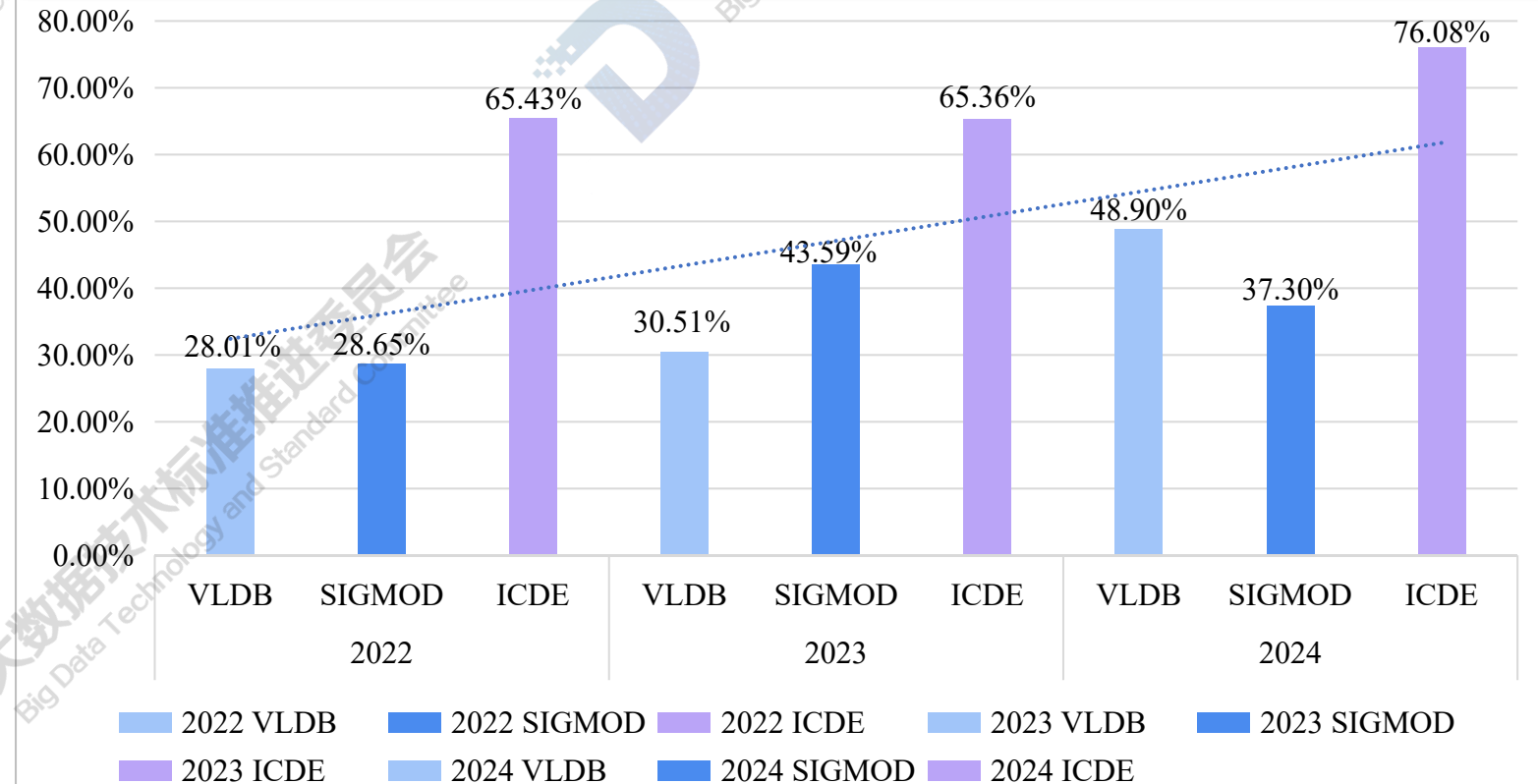
- ◆ 从三个数据库领域权威的学术会议研究方向看，当前非关系型数据库相关研究论文数量连续多年超过关系型数据库。
- ◆ “多模数据处理”成为关注热点。向量数据库、云原生数据库、图神经网络、“AI+数据库”等的热度也在持续增高。

- ◆ 我国在全球三大数据库领域学术会议的影响力进一步提升。三大会议每年贡献占比平均为40.70%、46.35%和54.09%，数量呈逐年上升趋势。企业与高校不断重视数据库理论创新与应用实践相结合，有效助力我国数据库技术发展，学术影响力不断扩大。

2024年VLDB、ICDE 和 SIGMOD 论文情况

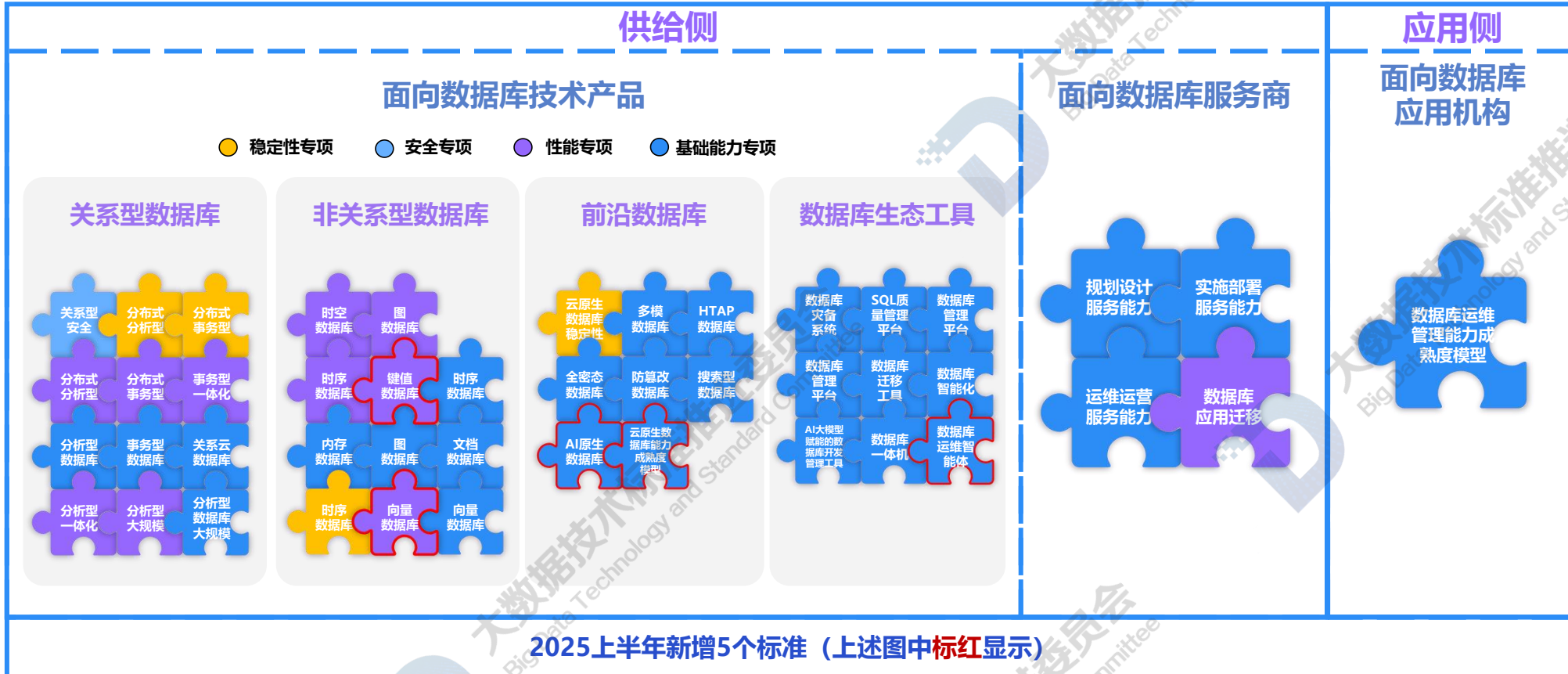


中国高校及企业贡献论文数占比



数据来源：CCSA TC601，2025年6月

# 洞察#8 从标准体系看，我国标准体系日益完善，引领产业高质量发展



**中国通信标准化协会大数据技术标准推进委员会(CCSA TC601)**紧跟国家战略，围绕数据库领域标准化工作，设立数据库与存储工作组（WG4）。自**2015**年起共推出**40**项标准，逐步构建以数据库产品、服务和应用为目标的标准体系。

**产品能力方面**，面向关系型数据库、非关系型数据库、前沿数据库和数据库生态工具，构建基础能力、性能、稳定性等技术标准。

**服务能力方面**，围绕规划设计、实施部署和运维运营，推出国内首个数据库交付标准《数据库服务能力成熟度模型》，围绕数据库应用迁移、SQL质量管理平台、数据库灾备恢复系统，推出能力分级标准。

**行业应用方面**，面向数据库应用方内部运维管理团队，推出《数据库运维管理能力成熟度模型》。CCSA TC601见证了我国数据库标准化工作有序有力进行，成为国家在数据库领域最重要的支撑单位，和数据库产业化发展的风向标。

洞察#1 从市场规模看，我国市场规模近600亿元，云上市场占比进一步扩大

洞察#2 从竞争格局看，全球市场进入高质量发展阶段，我国数据库厂商数量显著收敛

洞察#3 从产品类型看，国内外分布各有侧重，向量数据库热度持续攀升

洞察#4 从商业模式看，商用数据库占据主导地位，PG生态获得投资青睐

洞察#5 从我国市场看，市场规模持续扩大，市场格局逐步明晰

洞察#6 从重点行业看，头部用户运维人员投入逐年增长

洞察#7 从学术创新看，多模融合成为研究重点，我国科研实力稳步增强

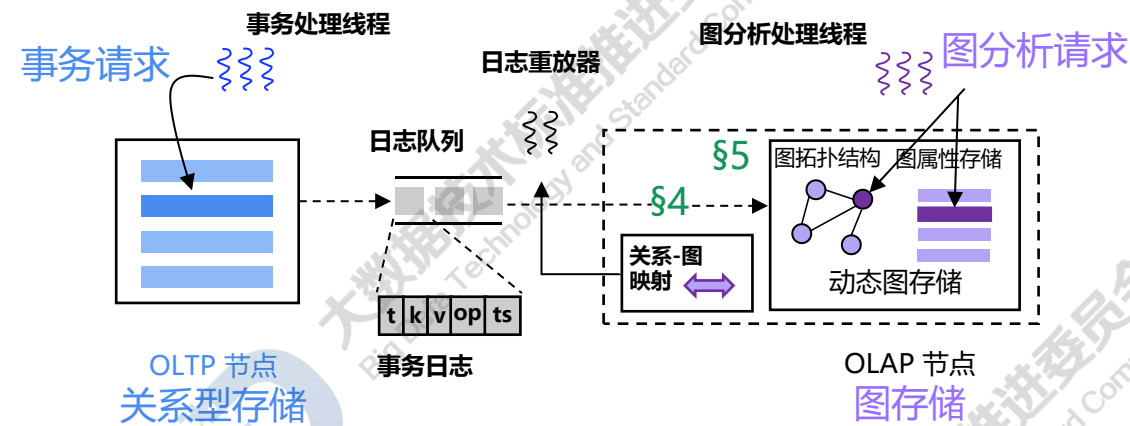
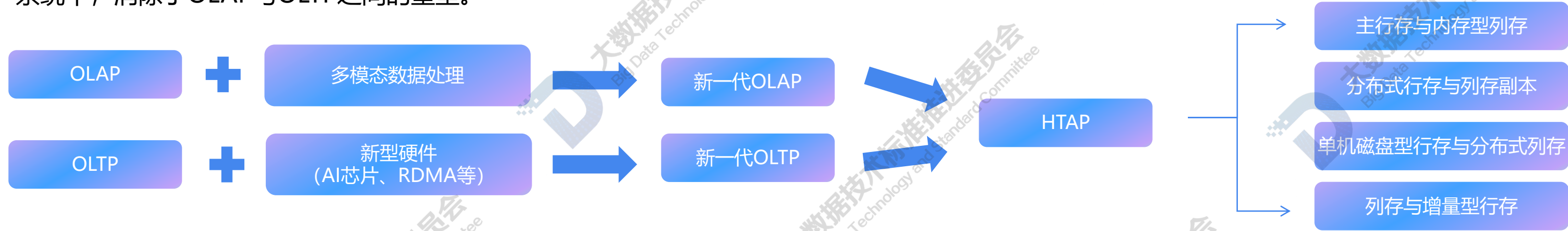
洞察#8 从标准体系看，我国标准体系日益完善，引领产业高质量发展

# 数据库技术架构持续革新迈入AI原生时代

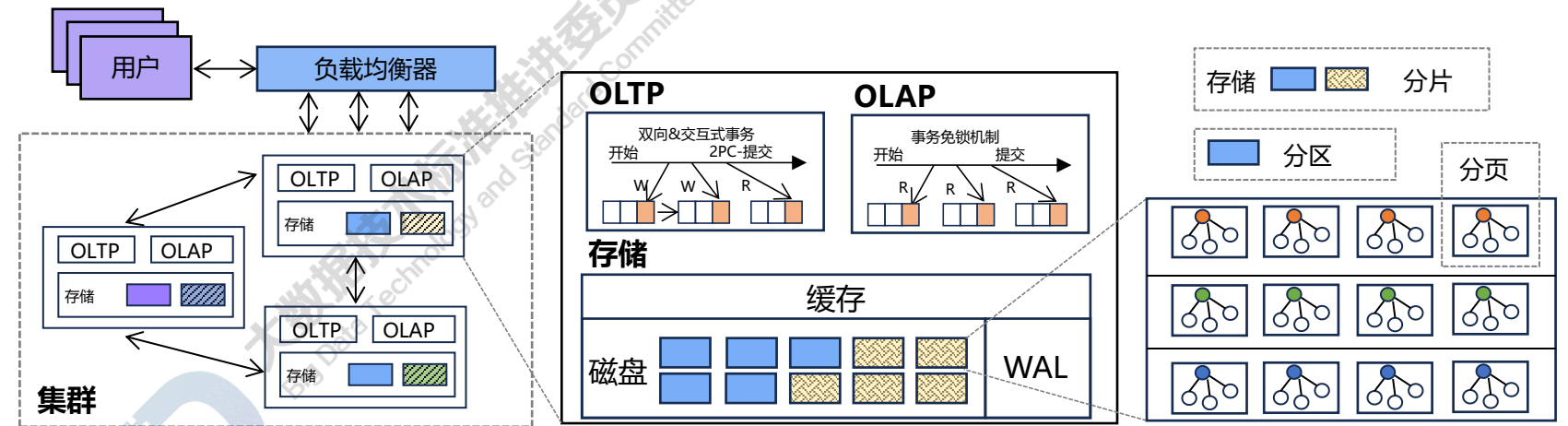


# 洞察#1 技术架构持续革新-从分而治之到交易分析一体化

◆ OLAP数据库擅长数据分析，OLTP数据库专注于事务处理。随着数智时代的到来，在事务处理的基础上，企业对于实时数据的分析需求提升，OLAP和大数据技术栈分化也给企业运维管理多套系统带来挑战。HTAP（Hybrid Transactional and Analytical Processing）架构应运而生，通过将事务处理与分析处理融合在一个系统中，消除了OLAP与OLTP之间的壁垒。



GART 系统的总体架构<sup>1</sup>



Galaxybase架构图<sup>2</sup>

## 图HTAP数据库

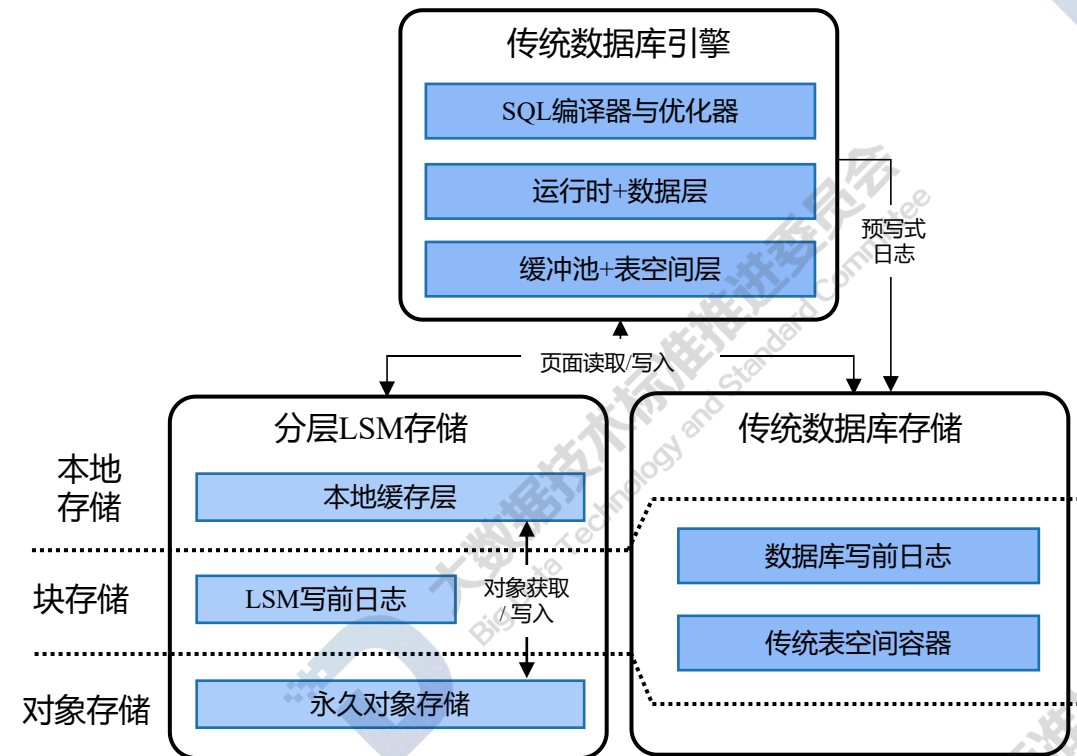
<sup>1</sup> Shen, S., Yao, Z., Shi, L., Wang, L., Lai, L., Tao, Q., Su, L., Chen, R., Yu, W., Chen, H., & Zang, B. (2023). Bridging the Gap between Relational OLTP and Graph-based OLAP. In 2023 USENIX Annual Technical Conference (USENIX ATC '23) (pp. 1-16).

<sup>2</sup> Bing Tong, Yan Zhou, Chen Zhang, Jianheng Tang, Jing Tang, Leihong Yang, Qiye Li, Manwu Lin, Zhongxin Bao, Jia Li, and Lei Chen. Galaxybase: A High Performance Native Distributed Graph Database for HTAP. PVLDB,17(12): 3893 - 3905, 2024.

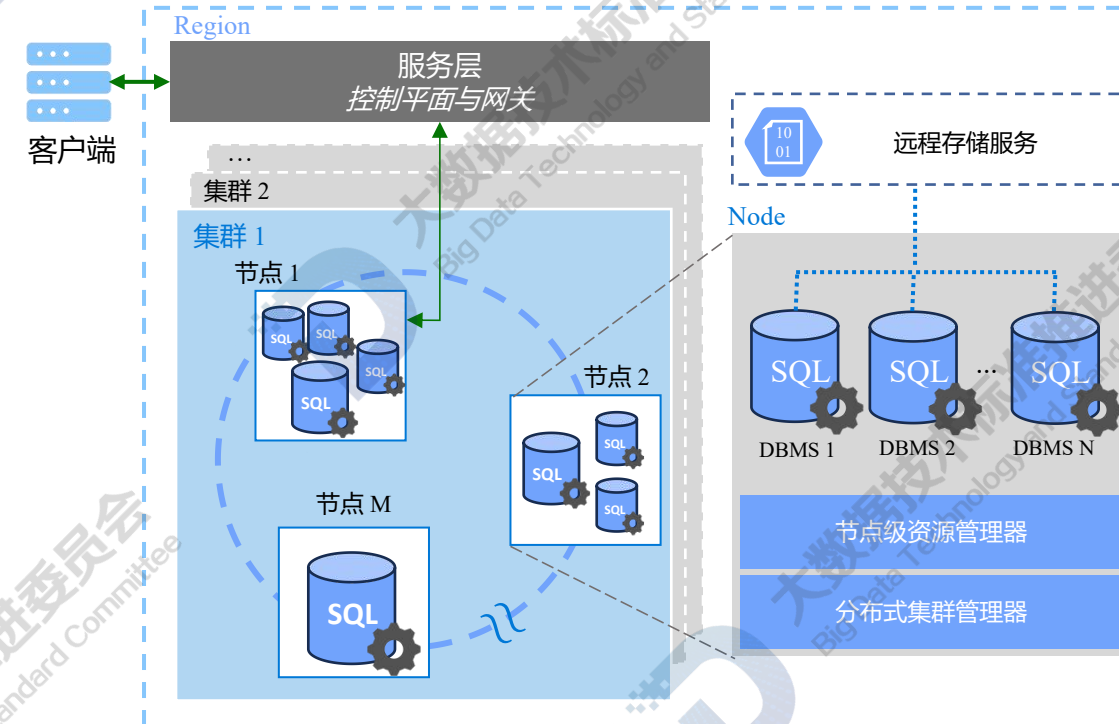
- ◆ 近年来各组织上云进程不断加速，传统数据库逐渐从私有部署转为云上部署。传统数据库为了保障系统高性能，基本采用紧耦合设计，而云计算架构的理念则是通过计算、存储的高度解耦来实现各个层次资源的高效利用和弹性扩展。为了更好地适应云环境，云原生数据库成为重要发展趋势。

## 云原生数据库实现高效资源管理

诸多企业采用的存算分离架构以及应用其他相关技术使得产品在云环境中能力得到提升。例如：采用LSM-Tree和列式存储等新型结构提升存储引擎专业化，采用资源动态调整和主动资源分配使得资源实现精细化管理，采用软件结合硬件实现软硬件协同创新加速。



Db2 LSM 架构图<sup>1</sup>



Azure SQL数据库架构概览<sup>2</sup>

## 云原生数据库能力成熟度模型标准框架

为更好指导企业对于云原生数据库研发及应用，规范云原生数据库标准化建设，信通院开展《云原生数据库能力成熟度模型》标准编制工作，将能力分为初始级、发展级、成熟级及领先级四个能力等级。



<sup>1</sup> David Kalmuk, Christian Garcia-Arellano et al. 2024. Native Cloud Object Storage in Db2 Warehouse: Implementing a Fast and Cost-Efficient Storage Architecture.

<sup>2</sup> Olga Poppe, Pankaj Arora, Sakshi Sharma, Jie Chen et al. 2024. Proactive Resume and Pause of Resources for Microsoft Azure SQL Database Serverless. In Companion of the 2024 International Conference on Management of Data.

◆ 人工智能迅猛发展影响着千行百业，数据库作为数据存储和管理的主要载体，与人工智能的融合发展正深刻地改变着数智应用的格局。数据库领域在查询优化、自然语言交互和智能资源管理等方面迎来了新的突破。

## 机器学习优化数据库应用效能

随着机器学习技术不断发展，其在数据库管理系统中的应用逐渐增多，涵盖数据库的索引、调度器、查询优化器等数据库核心组件，以及性能预测、参数调优等数据库顾问的各个方面。



ML for DB领域仍面临诸多挑战

模型效率有限

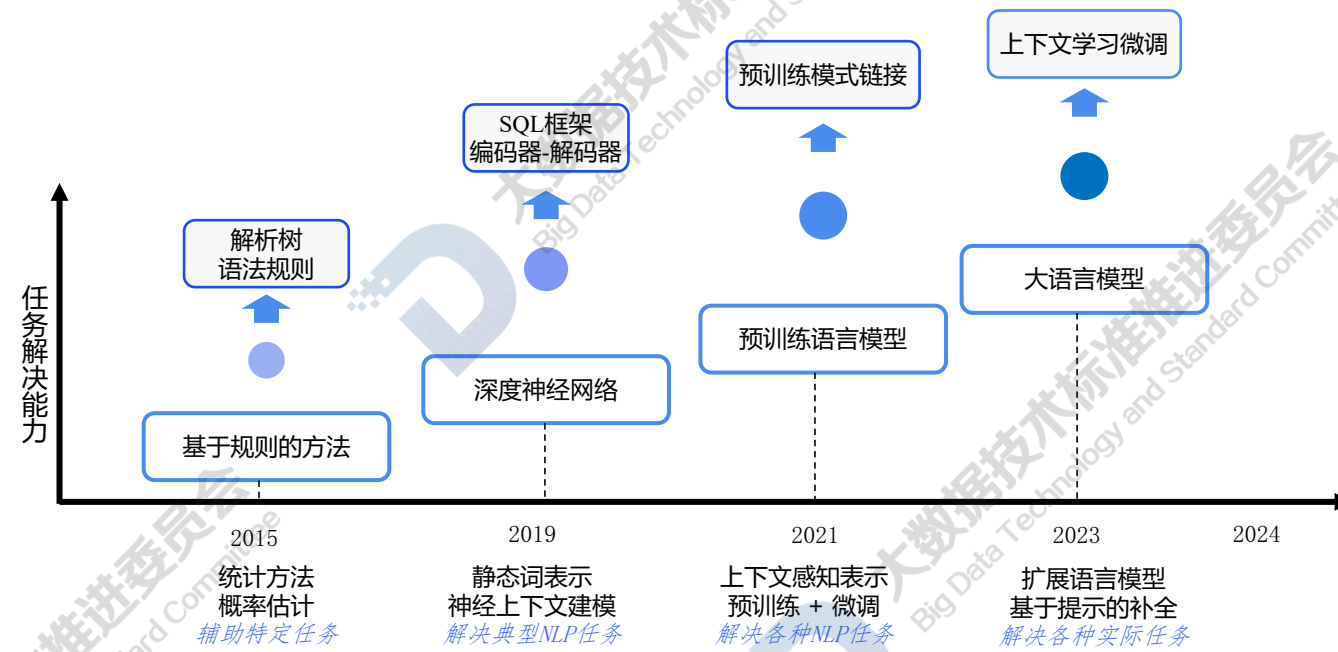
数据处理及负载动态变化

基础模型有效性依赖特定场景

训练数据集成本高且质量难以保证

## Text to SQL/Text to GQL降低数据库交互门槛

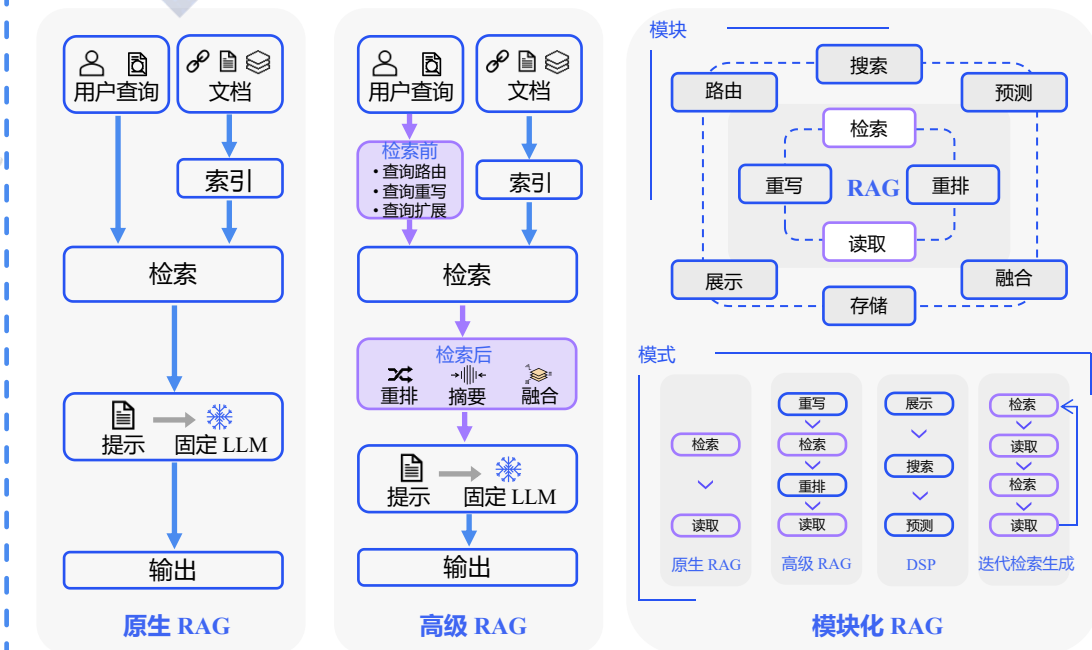
随着人工智能和自然语言处理技术的进步，Text-to-SQL技术应用而生，研究者开始利用深度学习模型对文本数据进行训练，使得系统能够更准确地理解用户意图从而降低数据库使用门槛。



Text-to-SQL发展历程

## RAG/Graph RAG提升数据库查询准确性

RAG (Retrieval-Augmented Generation) 通过整合外部数据库的知识，有效解决LLMs在特定领域或知识密集型任务中面临的幻觉、知识过时和推理过程不透明等挑战。

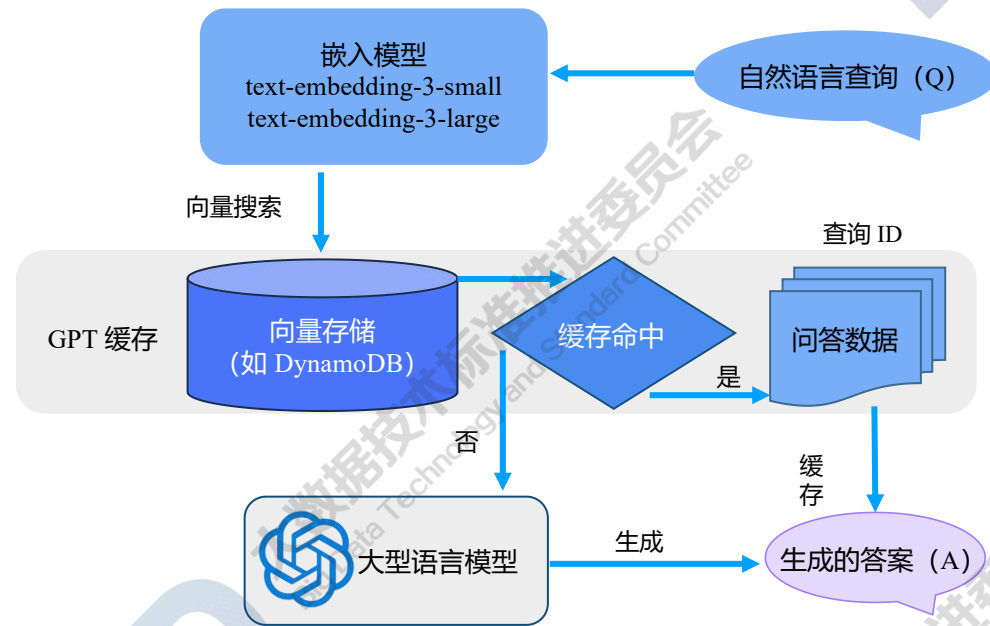


RAG 三种范式对比<sup>1</sup>

◆ 数据库技术也在不断赋能人工智能发展，正在多方位满足人工智能对数据存储、检索和处理的要求，数据库领域不断创新也为人工智能发展提供强有力支撑。

## 向量数据库提升大模型检索效率

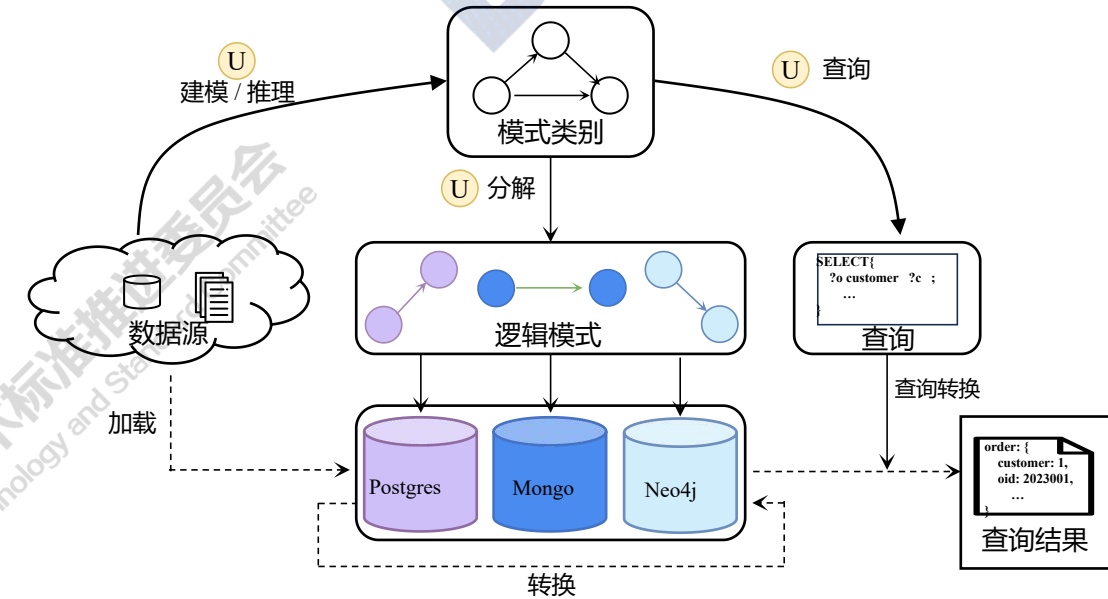
LLMs在实际应用中存在训练数据更新不及时、数据种类不足等原因导致的幻觉问题，向量数据库 (VDBMS) 作为一种专为高维向量数据设计的存储与检索系统，有效管理和检索向量数据，为解决向量检索难题、LLMs幻觉问题及知识库更新挑战提供了新路径。



使用向量数据库的GPT的语义缓存概览<sup>1</sup>

## 多模数据库赋能大模型多样化数据

非结构化和半结构化数据在人工智能应用中扮演着越来越重要的角色，如文本、图像、音频和视频等。多模数据库能够在单一系统中同时支持多种数据模型，在保证性能和一致性的同时，实现了对异构数据原生支持，提升了开发效率与数据处理能力。

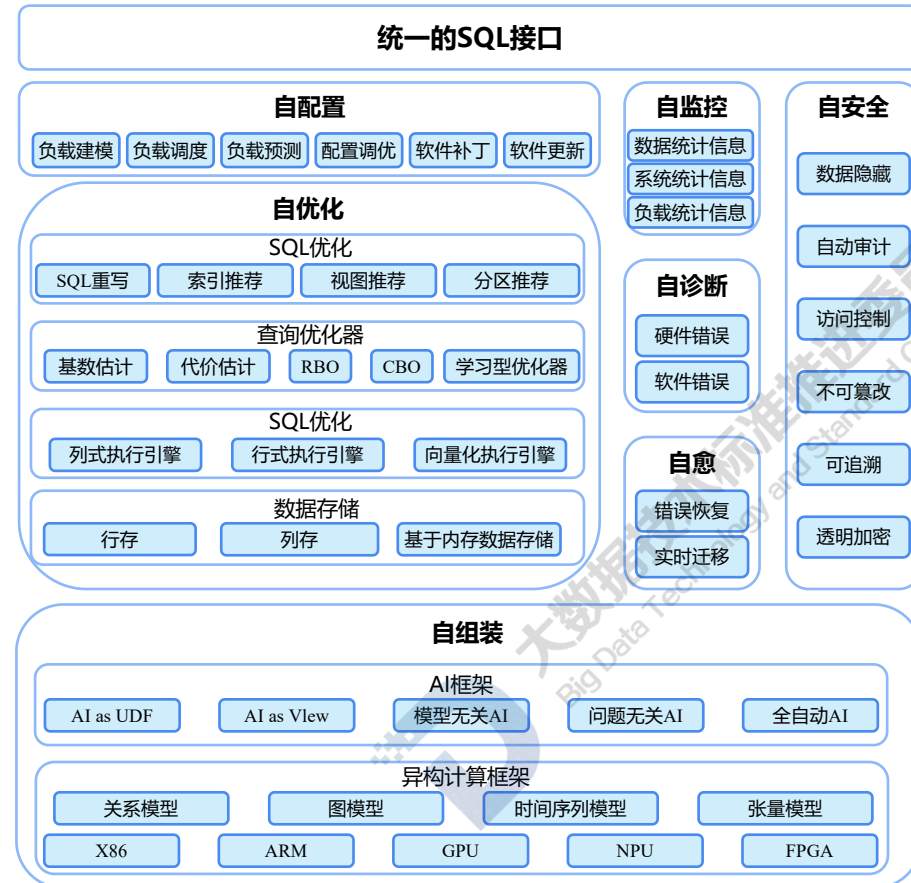


统一抽象层表达异构数据结构示意图<sup>2</sup>

◆ AI原生数据库将AI技术深度融入数据库设计与功能中，实现数据处理的智能化、自动化，显著提升数据管理与分析的效率和价值。

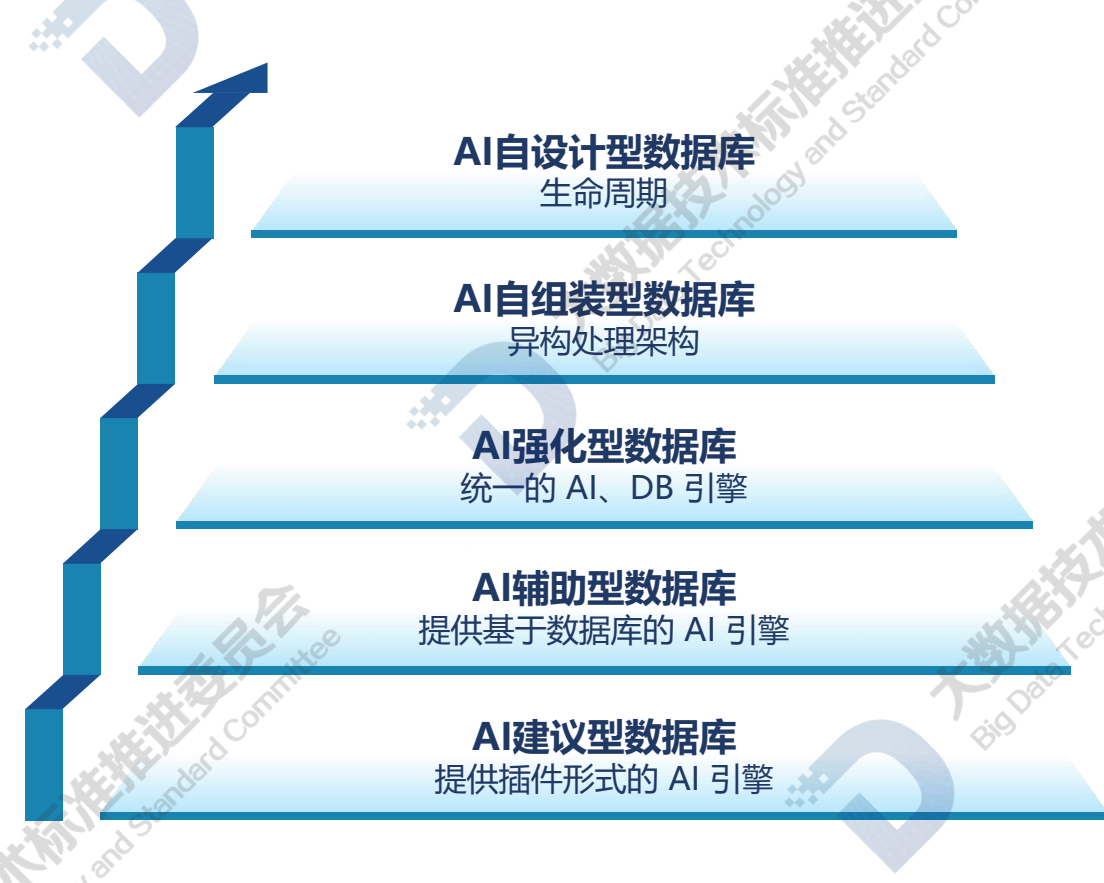
## AI原生数据库的架构<sup>1</sup>

将AI结合到数据库处理、运维和组装过程中使得数据库实现自监控、自配置、自优化、自诊断、自愈、自安全和自组装的能力，并为人工智能和数据库服务提供了统一的调用接口。



## AI原生数据库的五个阶段<sup>1</sup>

AI原生数据库依据智能程度依次分为五个阶段，AI建议型数据库、AI辅助型数据库、AI强化型数据库、AI自组装型数据库以及AI自设计型数据库。



## AI原生数据库技术要求标准框架

为更好指导企业对于AI原生数据库研发及应用，规范标准化建设，中国信通院开展《AI原生数据库技术要求》标准编制工作，标准共有9大能力域、69个能力项。

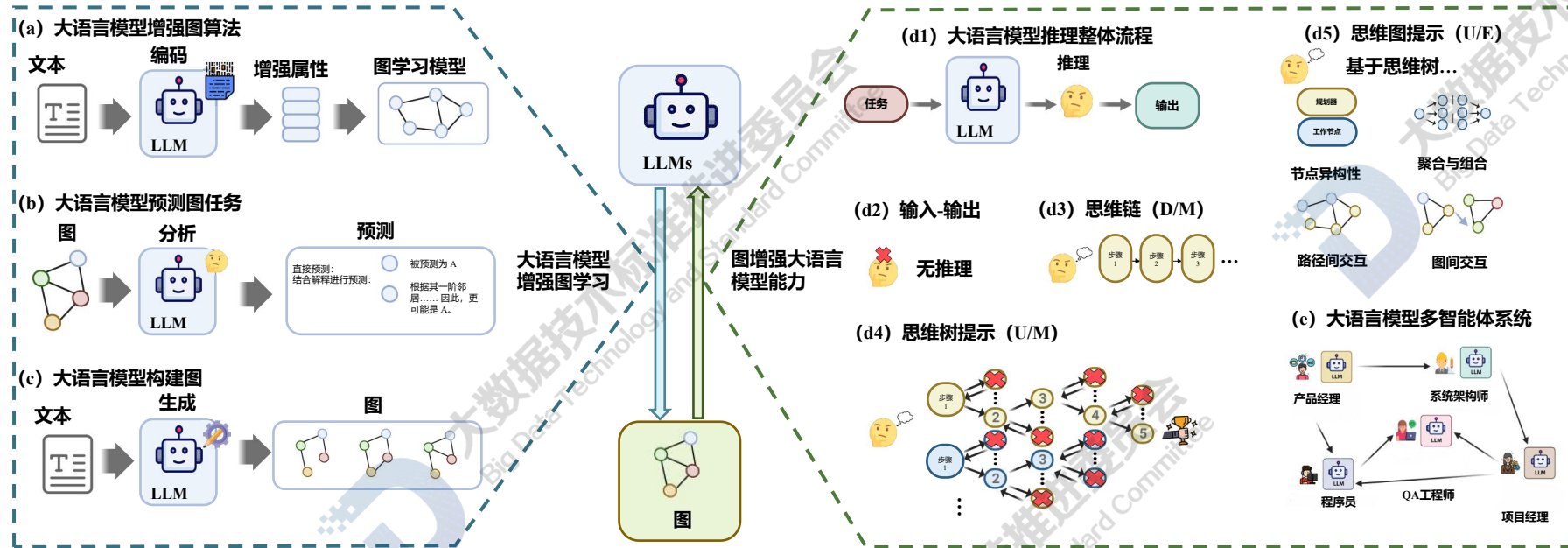


<sup>1</sup> 李国良,周焯赫,轩轶:AI原生数据库系统[J].软件学报,2020,31(3):831-844.

◆ 数据库智能体整合了自然语言处理、机器学习、知识图谱和多模态数据处理等多种技术，能够理解用户意图并高效地与数据库进行交互，为用户提供智能化、个性化的数据服务。2025年7月，腾讯云数据分析智能体TCDataAgent在“NL2SQL 全球最难榜单” BIRD-Bench中获得全球第三、国内第一的优异成绩，展现了我国企业在Agent领域的强大创新能力。

## 大语言模型 (LLMs) 与图技术相互增强的总体框架

随着分布式、多模态等数据库技术的不断发展和人工智能技术的崛起，数据库正朝着智能化方向迈进。数据库智能体技术主要基于自然语言处理与数据库交互、机器学习优化数据库应用，以及多模态数据处理技术。



## 数据库运维智能体技术要求标准框架

为更好指导企业对于数据库运维智能体研发及应用，规范数据库运维智能体标准化建设，信通院开展《数据库运维智能体技术要求》编制工作，标准包括5大能力域、74个能力项。

能力域	知识管理			容量管理				性能优化		
	知识来源	知识库维护	知识关系提取	基本信息监测	容量操作	容量变化异常	容量规划建议	对象优化	SQL优化	拓扑架构优化
日常巡检	知识分类与标注	知识检索	知识评价反馈	容量预测对象	预测可用天数	容量优化建议	自动执行	数据库整体优化	优化操作	自动执行
	基本信息监测	自定义报告	报告分发与共享	巡检报告编辑	关联分析	自动发现异常日志	故障发现	自定义报告	自动生成报告	
	日志查看检索	健康评分	慢查询识别	巡检报告对比	参数变更关联分析	系统配额监测	告警策略自定义	故障分类、分级、升级	恢复处理建议	
	服务日志分析	TOP SQL分析	健康趋势分析	巡检结果汇总展示	参数变化审计	动态感知业务	报告分发与共享	故障定位与诊断	故障自动修复	
故障管理	自定义监控指标	SQL监测	慢查询分析	参数优化建议	指标预测预警	慢查询分类	故障报告编辑			
	事务监测	阻塞监测	指标阈值优化建议	联合分析	慢查询预识别	慢查询优化				

# 03 我国数据库应用创新场景不断丰富，应用规模不断扩大



## 📶 国产数据库在电信行业得到广泛应用

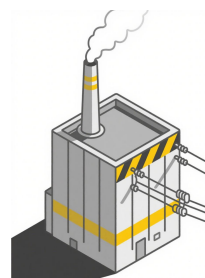
- ★ 从集中式向分布式演进，满足高并发与多中心部署需求
- ★ 从单一数据模型向多模融合扩展，应对结构多样的业务场景
- ★ 从传统静态存储向智能推理升级，支撑实时分析与自主决策

01

## 🏦 数据库技术驱动智能金融核心变革

- ★ 从“可用”到“好用”，成为金融业务高效运转的核心底座
- ★ AI与数据库融合，激活数据价值，助力智能决策与业务创新
- ★ 构建敏捷、安全的数据体系，全面支撑多元金融场景的发展

02



## ⚡ 数据库技术支撑能源行业数字化转型

- ★ 技术创新与融合方向发展
- ★ 场景不断拓展优化，贴近业务需求
- ★ 性能与安全需求提升，推动新兴技术持续应用

04

## 🚗 数据库技术助力交通运输行业平安出行

- ★ 车辆运行、驾驶行为与道路信息等数据来源多样、结构复杂
- ★ 数据库技术不断创新发展，提供强大的数据处理能力
- ★ 促进产业链上下游数据协同，推动行业高效、安全、智能发展

03



## 从产业看，全球数据库市场增速平稳，市场竞争进入深水区。

我国数据库产业规模稳步增长，市场格局头部效应明显；产业生态日益完善，数据库产品愈发成熟；人才生态逐渐丰富，开源和学术影响力逐步扩大，多模数据处理、“AI+DB”等成为学术研究的热点方向。

## 从技术看，技术融合创新发展，数据库技术迈入AI原生时代。

技术架构持续向着一体化、云原生等方向不断革新，推动数据库应用模式向灵活高效方向发展，融合型数据库成为发展方向。数据库技术迈入AI原生时代，数据库开发、运维模式变得更加精准智能，为数据库创新应用提供了更加丰富的场景和广阔的空间。

## 从应用看，应用创新场景不断丰富，应用规模不断扩大。

应用程度日益加深，高效赋能传统产业转型升级，促进数字经济和实体经济深度融合，助力新质生产力加快形成和发展。图数据库、多模数据库等新型数据库有力支持能源、交通运输业等传统行业数字化转型。

**“滴水成河，积沙成塔”**  
推动产业链上下游的协同发展，为产业的长远发展奠定坚实基础。

# 《数据库发展研究报告（2025年）》参编单位及专家

TDBC 2025  
可信数据库发展大会

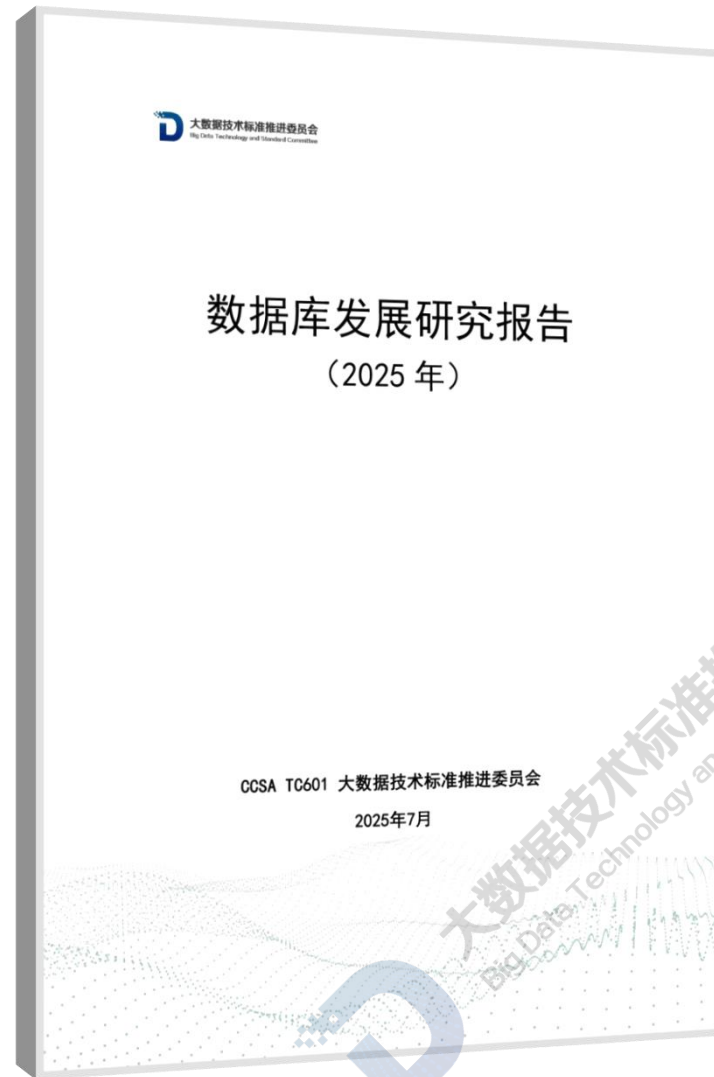
感谢《数据库发展研究报告（2025年）》编写组的努力！

## 主要编写单位（排名不分先后）：

中国信息通信研究院、中国移动集团数智化部、金篆信科有限责任公司、中电科金仓（北京）科技股份有限公司、天津南大通用数据技术股份有限公司、甲骨文(中国)软件系统有限公司、上海运熹科技有限公司、中兴通讯股份有限公司、浪潮云信息技术股份公司、北京中亦安图科技股份有限公司、腾讯云计算（北京）有限责任公司、深圳计算科学研究院、深圳九有数据库有限公司、北京奥星贝斯科技有限公司、联想（北京）有限公司、天翼云科技有限公司、西安同星恒耀信息技术有限公司、成都虚谷伟业科技有限公司、北京新数科技有限公司、北京自然原数科技有限公司、上海市数字证书认证中心有限公司、平凯星辰（北京）科技有限公司、北京海量数据技术股份有限公司、中国民航信息网络股份有限公司、中国第一汽车股份有限公司、新华三技术有限公司、北京酷克数据科技有限公司、浙江创邻科技有限公司、中移（苏州）软件技术有限公司、北京万里开源软件有限公司、杭州质变科技有限公司、上海天玑科技股份有限公司

## 编写组主要成员（排名不分先后）：

刘思源、齐丹阳、刘蔚、杨佳星、姜春宇、马鹏玮、康宸、王小玉、赵东明、吕伟初、朱业、秦延涛、左奇、董文、刘俊、胡一鸣、赵伟、冯文忠、白雪、温婧、李珈、杨鲁、杨元同、陈磊、赵衍衍、冷友方、孙路明、屠要峰、陈河堆、解海波、尚长军、邓光超、王阳、刘刚、黄远邦、郑成龙、高清华、徐晓敏、张展、郭文、欧伟杰、孟凡彬、李伟超、毛雨爽、蒋昀岂、何威、刘浔、全锦琪、苗旭成、向祥兵、王栩、王枫、李文娇、赵世凯、杨丹、钟华龙、郭立、张心怡、刘启荣、高新刚、张立、曹国顺、明玉琢、马攀、黄元霞、刘永利、朱浩、江晶、郑宇隆、刘孟占、何毓锬、刘小兵、杨晶、曹金龙、宋可欣、白玥、秦宁慰、田丰、刘中一、阎松柏、李莉、董骐、渠谨黛、郭宏伟、霍然、李永平、张战防、徐田原、唐昕、周研、张晨、李欢、吴晓峰、陶捷、刘永德、张卓、徐爽、魏闯先、陆元飞、肖煜伟、朱伟、孙洪标、俞好龙、吴国新、朱奕飞、孙成



# 《中国数据库产业图谱 (2025)》发布



# THANKS

## 感谢聆听

自主 · 创新 · 引领