

2023年中国数据管理行业词条报告

作者：吴金鑫

摘要 伴随着大数据时代支撑数据交换共享和数据服务应用的技术发展，不断积累的数据开始成为一种新型资产，数据管理解决方案的重要性逐渐凸显。数据管理是数据处理的核心，是指对数据的组织、分类、编码、存储、检索、维护等环节的操作，从细分领域来看，数据管理系统主要分为数据库管理、数据模型管理、元数据管理、主数据管理、数据质量管理、数据安全治理、数据价值管理、数据共享管理八大模块。数据资源已经成为数字经济时代的软件底座，几乎所有的企业数据、网络数据和边缘设备数据都需要经数据管理和分析才能够赋能上层应用或企业决策，发挥其经济时代的价值。在中国进入数字化时代后，随着存储技术和云计算的发展，数据管理解决方案对于数据在调用、处理和分析，再通过人工智能技术让数据资产化在中国企业应用和发展。

行业 头豹行业信息库、软件和信息技术服务业软件和信息技术服务业、新技术类高科技软件和信息服务

1. 数据管理行业定义

原始数据通常存在质量差、价值密度低等问题，在经过采集和存储后，通常需要经过多次加工，包括但不限于治理、建模和分析挖掘等工作，才能达到数据的规范化、资产化，实现数据的共享和流通，进而发挥数据的使用价值。数据管理是利用计算机硬件和软件技术对数据进行有效的收集、存储、处理和应用流程，旨在从原始数据中抽取、指导出有价值的信息，以支撑企业决策。数据管理解决方案支持优化各项数据管理流程的运行，助力企业用户更好的整合数据资产、提升数据质量。

2. 数据管理行业分类

数据资产管理的管理职能包括数据标准管理、数据模型管理、元数据管理、主数据管理、数据质量管理、数据安全治理、数据价值管理以及数据共享管理共八个方面。一般情况下，数据管理是自下而上的模式，从元数据管理、数据质量管理两项核心数据资产管理任务作为切入点，逐步扩展到数据模型管理、数据标准管理、数据安全治理等其他数据管理职能。

类型名称	类型说明
数据标准管理	数据标准是指保障数据的内外部使用和交换的一致性和准确性的规范性约束，通常可分为基础类数据标准、数据标准化管理标准、基础类数据标准一般包括参考数据标准、逻辑数据模型标准、逻辑数据模型标准、元数据标准、公共代码和编码标准等。指标类数据标准一部分为基础指标标准和计算指标（又称组合指标）标准，数据指标一般不单独值信息，且具有特定业务和经济含义，计算指标通常由两个以上基础指标计算得出。
数据模型管理	数据模型是现实世界数据特征的抽象，用于描述一组数据的概念和定义。数据模型从抽象层次上描述了数据的静态特征、动态行为和约束条件。数据模型所描述的内容有三部分：数据结构、数据操作（其中ER图数据模型中无数据操作）和数据约束。
元数据管理	元数据（Metadata）是描述数据的数据。元数据按用途不同分为技术元数据、业务元数据和管理元数据。元数据管理是数据资产管理的重要基础，是为获得高质量的、整合的元数据而进行的规划、实施与控制行为。
主数据管理	主数据（Master Data）是指用来描述企业核心业务实体的数据，是企业核心业务对象、交易业务的执行主体。主数据治理是一系列规则、应用和技术，用以协调和管理企业的核心业务实体相关的系统记录数据。
数据质量管理	数据质量是保证数据应用效果的基础。数据质量管理有几个典型的指标有：完整性（数据是否缺失）、规范性（数据是否按照要求的规则存储）、一致性（数据的值是否存在信息含义上的冲突）、准确性（数据是否错误）、唯一性（数据是否是重复的）、时效性（数据是否按照时间的要求进行上传）。
数据安全治理	数据安全治理指对数据设定安全等级，按照相应国家/组织相关法规及监管要求，通过评估数据安全风险、制定数据安全管理治理规范、进行数据安全分级分类、完善数据安全相关技术治理体系，保证数据被合法合规、安全地采集、传输、存储和使用。
数据价值管理	数据价值管理是对数据内在价值的度量，从数据成本和数据应用价值两方面来测算。数据成本一般包括采集、存储和计算等费用（人工费用、IT设备等直接费用和间接费用等）和运维费用（业务维护费、技术服务费等）。数据价值（收益）主要从数据资产的分流、使用频次、使用对象、使用效果和共享流通等方面计量。
数据共享管理	数据共享管理主要是指开展数据共享和交换，实现数据资产价值的一系列活动。数据共享管理包括数据内部共享（企业内部跨组织、部门的数据交换）、外部流通（企业之间的数据交换）、对外开放。

3. 数据管理行业特征

在数字管理解决方案主要向企业客户提供数字化基础设施底层、中间层的基础软件和技术服务，支持客户的技术团队及合作伙伴构建数据管理和业务应用系统，助力客户进行数字化转型。随着管理数据对象越来越多，数据治理技术越发成熟，数据应用范围越发广泛，数据管理解决方案在数据处理架构、组织职能、管理手段等方面逐渐呈现了一些新的特点和业务趋势。

数据对象持续变化	数据互联网络行业的蓬勃发展与数字化转型促进了海量数据增长和复杂的数据来源 数据作为数据管理解决方案的对象，呈现数据海量增长、来源多样化的特征。在数据量方面，单一机构的数据规模由以前的GB级上升到TB级，甚至PB级、EB级。在数据种类和来源方面，除传统的结构化数据之外，文本数据、图像数据、语音数据、视频数据等非结构化数据或半结构化数据占比越来越高，种类日益丰富，在数据来源方面，数据既包括内部数据，也包括来自第三方的外部数据，既包括传统业务处理采集的业务数据，也包括手机终端、传感器、机器设备、网站网页、日志等技术产生的数据。
数据处理架构持续迭代	数据处理的底层架构正在从ETL结构向ELT转变，分布式“计算+存储”的数据处理架构成为主流 数据的预处理流程正在从传统的ETL结构向ELT转变，传统的数集成处理流程是ETL结构，其流程是构建数据仓库的重要一环，即用户从数据源抽取所需的数据，经过数据清洗、将数据加载到数仓中。而大数据背景下的架构体系是ELT结构，其流程上层的业务需求，随时从数据源中抽取想要的原始数据进行处理分析，以Hadoop、Spark等分布式技术为核心的“计算+存储”的数据处理架构，能够支持批处理和实时的数据加载以及灵活的业务需求。
数据管理职能独立化	为了更好地数字化转型，越来越多的企业专门设立数据管理部门，带动新的岗位需求 传统的数据管理体系中，数据管理职能主要由IT部门负责，是IT部门的一项工作，业务部门配合IT部门执行数据管理，提出需求。随着数据分析与业务融合越来越深入，业务部门逐步成为大数据负责的主角，而数据管理在企业中扮演越来越重要的角色，重视数字化转型的企业设置专门的“数据管理部门”职能部门或首席数据官岗位。在这种变迁背景下，数据管理在组织架构也面临革新的需求。

4. 数据管理发展历程

数据库理论自1978年进入中国，数据库通常服务于业务，数据来自各种业务系统软件程序产生的数据，或者是由这些业务系统软件交互的用户产生的数据。1989年，数据库理论被提出，数据库通常服务于分析，数据来源于业务系统的一个或者多个数据库或者文件。数据库在形式化数据体系结构和数据管理实践方面迈出了重要的第一步，有效消除冗余数据之间的一致性。随着数据仓库的普及，数据管理开始从产品端逐渐发展成为独立化的行业。

开始时间：1978 结束时间：1996 阶段：萌芽期	行业动态：1978年，萨拜德将数据库理论引入中国。1985年，中国数据库技术首先应用于国防、军工等领域。1988年，Barry Devlin和Paul Murphy在IBM Systems Journal上发表了一篇文章，描述了一种用于数据管理的体系结构，其中的核心组件被称之为数据仓库。1989年，Oracle在《中国的册中文明称“甲骨文”。在1991年，IBM宣布数据库作为一个产品。1992年，Bill Inmon的《开发数据仓库》一书开始普及数据仓库。
行业影响/阶段特征：	数据管理伴随着数据仓库概念的发展，自1990年代开始逐渐热门。随着数据仓库落地应用与展开，企业的业务数据结构开始出现开放，数据管理的重要性凸显。数据管理的提出，在早期主要是为了解决数据对硬件、软件资源的占用问题，通过体系化的管理，降低数据冗余，同时提高数据准确性、准确性、可用性以及规范性。
开始时间：1997 结束时间：2005 阶段：启明期	行业动态：1997年，Oracle承接中国东三省电信管理局5期工程项目，中国第一代DBA（数据库管理员）诞生。DBA的核心目标是保证数据库管理系统的稳定性、安全性、完整性和高性能。1999年，国产数据库企业，北京人大金仓信息技术股份有限公司成立。2000年，中国数据库商业化企业，武汉达梦数据库有限公司成立。2003年，普元信息技术有限公司成立。
行业影响/阶段特征：	依靠手工人力电子表格数据管理模式被自动化智能的专业工具取代，越来越多的数据库管理员、业务分析师等数据管理相关岗位出现。以数据库为代表的数据库管理工具和自动化增强企业的数据库管理能力，包括梳理元数据、管理主数据，优化数据集成、提升数据质量等，自动智能的专业数据管理工具逐渐发展成系统化的数据管理解决方案。
开始时间：2006 结束时间：2025 阶段：高速发展期	行业动态：2006年，中国头部的数据可视化厂商，帆软软件有限公司成立。同年，数据智能应用软件开发商明高科技成立。2013年，互联网行业开始进入高速增长阶段，电子商务的蓬勃发展催生模式创新，互联网企业的数量开始呈指数级上升，数据管理需求释放。对服务企业客户和政府部门的“数据数据治理方案提供提前解决痛点
行业影响/阶段特征：	数字经济时代的到来催化了数据管理解决方案底层技术的迭代和终端产品的优化。数据的资产价值愈发重要，各行各业的企业对数据管理产品的需求持续增长，数字经济推动了数据的生产、分析、决策的需求释放，帮助企业数据价值变现的数据管理解决方案行业进入高速发展阶段。

5. 数据管理产业链分析

数据管理行业的上游是基础软件，一般包括大数据基础平台、数据库、数据中间件和数据科学平台。数据管理基础软件对多种场景、多种来源、多种类型的源数据进行有效的采集、存储、管理并提供相关开发平台和工具，统一管理业务中所需要的数据，让客户在数字化转型时有坚实的基础和工具支持。数据管理行业的中游是数据应用系统，构建在基础软件之上，系统架构包括数据湖、数据仓库、数据集市等，应用方向主要有数据治理、数据中台、数据决策平台、数据可视化平台、数据安全、云计算服务。数据管理行业的下游是政府部门和企业客户，数据管理解决方案为业务部门构建的数字化运营、数字化决策、智能制造、数字营销、智能风控等应用系统，助力政府部门和企业客户更加实时、准确地进行商业决策，提高业务运营效率，从而将数据价值转换为业务价值，实现数字化转型。

上游环节	上游说明	上游参与方
大数据基础平台	大数据平台是指以处理海量数据存储、计算及大数据分析数据实现并构建预测的平台。该平台由数据分发层、数据计算层、数据应用层组成。大数据平台架构包括Hadoop生态、Spark、Storm、Flume等组件集。大数据平台的应用场景可以采用开源平台，也可以采用私有云、混合云等商业解决方案，既可以部署在私有云上，也可以部署在公有云上。中国代表性的大数据平台开发商有普元信息、创略科技、亚信科技、明略科技、卓朗科技等。	普元信息科技股份有限公司、亚信科技（中国）股份有限公司、创略（上海）数据科技有限公司、北京京颐软件系统有限公司
数据科学平台	数据科学平台指利用数据、算法、机器学习AI和AI技术发现模式并构建预测的平台。该平台由数据分发层、数据计算层、数据应用层组成。大数据平台的应用场景可以采用开源平台，也可以采用私有云、混合云等商业解决方案，既可以部署在私有云上，也可以部署在公有云上。中国代表性的数据科学平台开发商有星环科技、网易、阿里巴巴、华为等。	网易（杭州）网络有限公司、星环信息科技（上海）股份有限公司、阿里巴巴（中国）网络技术有限公司、华为技术有限公司
数据库	数据库是按照数据结构来组织、存储和管理数据的“仓库”，按照所使用模型来描述数据在某一时刻的状态或内容的不问，可分为关系型数据库和非关系型数据库。其中，不同的架构模式又可分为集中式数据库和分布式数据库，以数据分片、数据集中流式及事务处理技术为核心的大数据技术体系日渐完善，数据库技术和数据服务能力持续加深，助力企业用户数据价值释放。中国代表性的数据库开发商有阿里云数据、万里数据库、达梦数据库、巨杉数据库、人大金仓等。	阿里云融创数据科技（北京）有限公司、北京万里开源软件有限公司、武汉达梦数据库股份有限公司、广州巨杉数据库股份有限公司、北京人大金仓信息技术有限公司
数据中间件	数据中间件处于数据库和用户应用系统之间，提供跨平台、跨语言、跨系统、跨架构的数据集成、数据交换、数据共享等服务。数据中间件是连接数据库和应用系统的关键，是实现数据集成和数据共享的重要桥梁。数据中间件是连接数据库和应用系统的关键，是实现数据集成和数据共享的重要桥梁。数据中间件是连接数据库和应用系统的关键，是实现数据集成和数据共享的重要桥梁。中国代表性的数据中间件开发商有星环科技、卓朗科技、网易数据、顺风科技等。	星环信息科技（上海）股份有限公司、北京顺风科技有限公司、杭州网易数据科技有限公司、天津卓朗科技发展有限公司

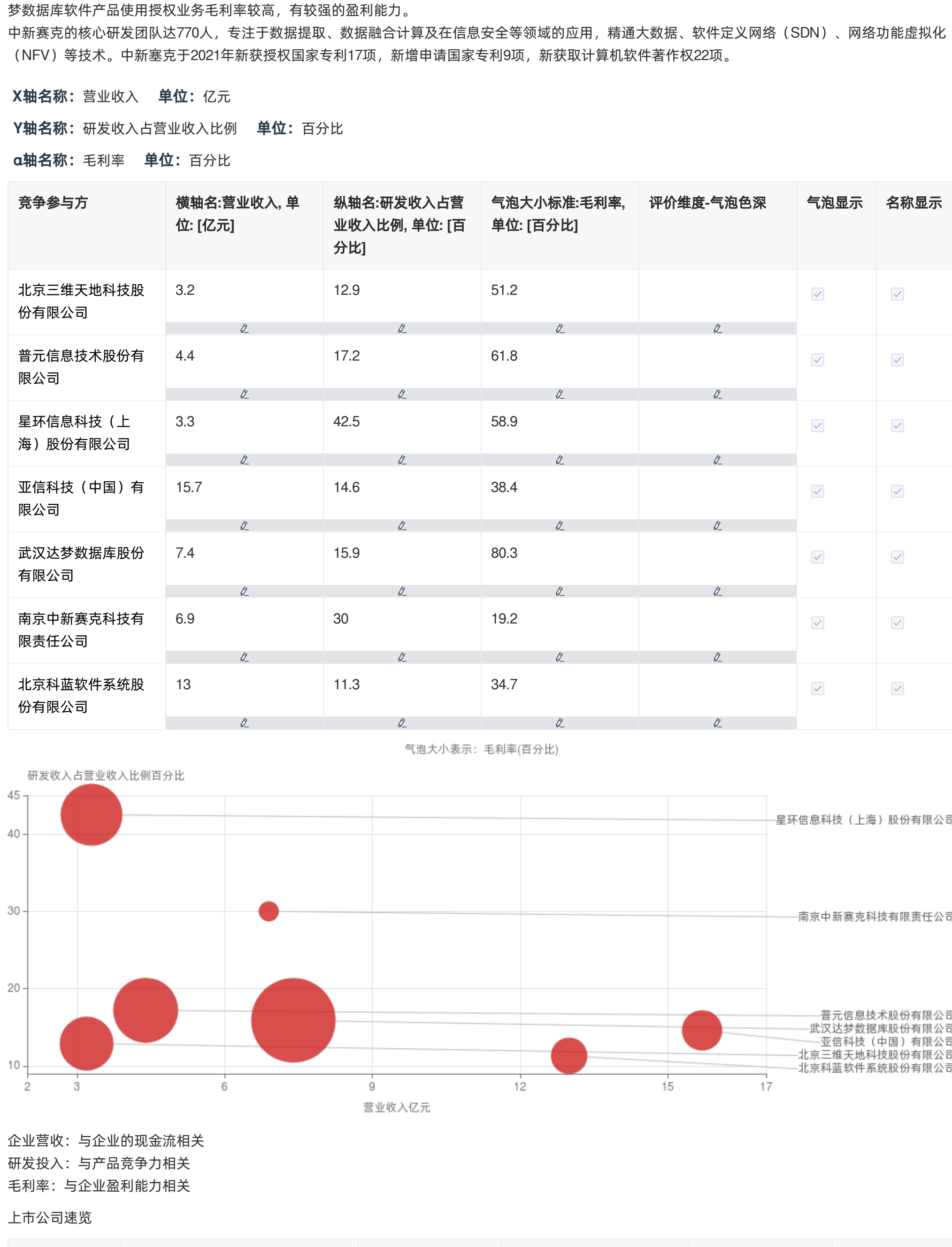
中游环节	中游说明	中游参与方
数据治理	数据治理是指以数据质量影响企业的业务决策，降低数据资产价值。数据治理的最终目标是为提高数据的质量，提高数据资产价值，是企业实现数字化转型战略的基础。数据治理是一套包括组织、制度、流程、工具的管理体系，当数据在业务系统使用中出现错误时，数据治理团队需要快速定位数据来源，修复数据错误。中国代表性的数据治理厂商有亿信华辰、易观数据、观远数据、中新赛克等。	北京亿信华辰软件有限责任公司、北京易观数据股份有限公司、杭州观远数据有限公司、南京中新赛克科技有限责任公司
数据中台	数据中台具备四个核心能力，分别为数据汇聚整合、数据治理加工、数据资产化、数据服务能力。数据中台是指通过数据技术，对海量数据进行采集、计算、存储、加工，同时统一形成标准数据，再进行存储，形成大数据资产，进而为政府部门和企业客户提供高效服务。中国代表性的数据中台开发商有普元信息、星环科技、明略科技、亚信科技等。	普元信息科技股份有限公司、星环信息科技（上海）股份有限公司、北京明略软件系统有限公司、亚信科技（中国）有限公司
数据分析平台	数据分析平台是指通过数据挖掘，帮助用户深入分析业务数据，提升业务决策能力。该平台由数据分发层、数据计算层、数据应用层组成。该平台由数据分发层、数据计算层、数据应用层组成。该平台由数据分发层、数据计算层、数据应用层组成。中国代表性的数据分析平台开发商有神策数据、网易数据、大连信元、三维天地等。	神策网络科技（北京）有限公司、杭州网易数据科技有限公司、大连信元信息技术有限公司、北京三维天地科技股份有限公司
数据可视化平台	数据可视化是指将数据以图形化的形式呈现，帮助用户更直观地了解数据。数据可视化是数据管理的重要组成部分，是实现数据价值的重要手段。中国代表性的数据可视化平台有火山引擎、思迈特软件、帆软软件、观远数据等。	北京火山引擎科技有限公司、广州思迈特软件有限公司、帆软软件有限公司、杭州观远数据有限公司
数据安全	数据安全是指为数据资产提供保护和保障的技术和管理措施，防止数据泄露、篡改、丢失等安全风险。数据安全是数据管理的重要组成部分，是实现数据价值的重要手段。中国代表性的数据安全厂商有奇安信、绿盟科技、启明星辰、天融信等。	南京中新赛克科技有限责任公司、北京奇安信科技股份有限公司、浙江绿盟科技股份有限公司、网鼎信安科技集团股份有限公司、北京科蓝软件系统股份有限公司
云计算服务	云计算是指通过互联网提供按需、弹性的计算、存储、网络等服务。云计算是数据管理的重要基础，是实现数据价值的重要手段。中国代表性的云计算服务提供商有华为、腾讯云、阿里云、百度智能云等。	华为云计算技术有限公司、腾讯云计算（北京）有限责任公司、阿里云计算有限公司、百度云计算（北京）有限公司

下游环节	下游说明	下游参与方
政府部门	云计算、大数据、区块链、人工智能等新一代信息技术赋能政府部门的政务管理系统更加智能化、便捷化，实现便民服务的同时减少繁冗的政务工作。政府部门更关注数据管理应用产品的安全性和保密性，所采购的数据管理应用产品以保障数据安全和使用安全为主。2022年，国务院发布的《关于加快数字政府建设的指导意见》进一步加强政府数字化转型，政府部门的政务管理需求将持续增长。	地方政府
企业客户	数据管理解决方案被广泛应用于金融、医疗、工程等多个行业，以金融行业为例，金融科技公司（Fintech）作为重点的数字化转型战略，在此过程中，金融科技公司对外部数据的需求持续增长，对外部数据的应用越来越深，对外部数据进行精细化管理成为当前的重要任务，部分金融机构开始着手整合现有外部数据资源，重建外部数据管理机制，梳理外部数据服务流程，在安全合规的前提下充分、高效地挖掘外部数据资产价值。中国建设银行、中国工商银行、浦发银行、华夏证券等金融机构纷纷大力发展金融科技，推动数据管理解决方案的应用。	中国建设银行股份有限公司、中国工商银行股份有限公司、上海浦东发展银行股份有限公司、华泰证券股份有限公司

6. 数据管理行业规模

在各项政策的推动和数据经济的作用下，中国企业客户对于数据应用和数据资产管理的需求持续提升，而数据管理软件和服务则是实现数据资产管理的重要基础。2021年，中国数据管理行业规模达到244亿元，2017至2021年的复合年增长率超过30%。在未来，数据管理行业将向统一数据模型的方向，随着大数据厂商技术实力的提升，逐渐出现了能够提供多种数据模型的大数据平台技术，相比多种数据库产品的集成方案，多种数据模型统一的大数据平台的优势包括提升效率、统一分析管理、降低运维成本、降低数据管理成本。多种数据模型平台将经过不断提效提升、存储引擎的处理能力、从操作响应速度、数据开发能力、数据管理能力等多个角度优化企业的数据需求，预计2026年中国数据管理行业规模将达810亿元。

数据管理行业规模参考沙利文测算的大数据软件市场规模



数据管理行业规模参考沙利文测算的大数据软件市场规模

002912	深圳市中新赛克科技股份有限公司	0	26,189.02万元	-49.69	74.00
688031	星环信息科技(上海)股份有限公司		33,086.16万元		
301159	北京三维天地科技股份有限公司	0	10,384.41万元	-20.25	44.16
01675	亚信科技控股有限公司	74.95亿	3,109,285,000.00	14.5100	33.88
688118	普元信息技术股份有限公司		4,938.85万元	26.47	31.80

数据管理代表企业分析

星环信息科技（上海）股份有限公司【688031】

企业状态：存续

注册资本：9053.1468万人民币

企业总部：市辖区

行业：批发业

8. 数据管理竞争格局

本报告共挑选了三维天地、普元信息、星环科技、亚信科技、达梦数据库、中新赛克、科蓝软件，共七家数据管理解决方案提供商进行评估。评估维度分为企业营收、研发投入、毛利率三个维度。经过评估，亚信科技、中新赛克、达梦数据库有一定的竞争优势。亚信科技的传统业务为中国三大电信运营商提供业务支持系统(BSS)、电信运营支撑系统(BSS)、计费及营销等相关的数据服务。此外，亚信科技开辟传统业务外的“三新业务”，即DSaaS（服务器资源安全保护系统即DRP）、垂直行业与企业上、OSS业务（操作支持系统）。2021年，“三新业务”收入约为15.7亿元，成为亚信科技新的业务增长点。

达梦数据库软件产品成熟度及标准化程度较高，成本主要包括软件、包装料、生产及测试相关人工投入及货物物流费用等，投入金额较低。因此达梦数据库软件产品使用费收取毛利率较高，有较好的盈利能力。

中新赛克的核心研发团队达1770人，专注于数据提取、数据融合计算及在信息安全等领域的应用，精通大数据、软件定义网络（SDN）、网络虚拟化（NFV）等技术。中新赛克于2021年新获国家专利17项，新增申请国家专利9项，新获计算机软件著作权22项。

X轴名称： 研发投入 单位： 亿元	Y轴名称： 研发投入占营业收入比例 单位： 百分比	a轴名称： 毛利率 单位： 百分比
---------------------------------	---	---------------------------------

竞争参与方	横轴:营业收入,单位:亿元	纵轴:研发投入占营业收入比例,单位:百分比	气泡大小:毛利率,单位:百分比	评价维度:气泡颜色	气泡显示	名称显示
-------	---------------	-----------------------	-----------------	-----------	------	------

北京三维天地科技股份有限公司	3.2	12.9	51.2		✓	
普元信息科技股份有限公司	4.4	17.2	61.8		✓	✓
星环信息科技(上海)股份有限公司	3.3	42.5	58.9		✓	✓
亚信科技(中国)有限公司	15.7	14.6	38.4		✓	✓
武汉达梦数据库股份有限公司	7.4	15.9	80.3		✓	✓
南京中新赛克科技有限责任公司	6.9	30	19.2		✓	✓
北京科蓝软件股份有限公司	13	11.3	34.7		✓	✓

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------

研发投入占营业收入比例百分比	气泡大小:毛利率,单位:百分比
----------------	-----------------