

# 2020年 中国数据管理解决方案市场报告

2020 China Data Management Solutions Market Report

2020年中国ビッグデータ管理市場研究

报告标签:大数据、数据管理解决方案、数据湖、数据仓库

2021/04

报告提供的任何内容(包括但不限于数据、文字、图表、图像等)均系头豹研究院独有的高度机密性文件(在报告中另行标明出处者除外)。未经头豹研究院事先书面许可,任何人不得以任何方式擅自复制、再造、传播、出版、引用、改编、汇编本报告内容。若有违反上述约定的行为发生,头豹研究院保留采取法律措施、追究相关人员责任的权利。头豹研究院开展的所有商业活动均使用"头豹研究院"或"头豹"的商号、商标,头豹研究院无任何前述名称之外的其他分支机构,也未授权或聘用其他任何第三方代表头豹研究院开展商业活动。

# 概览说明

沙利文谨此发布中国数据管理系列报告之《2020年中国数据管理解决方案市场报告》年度报告。本报告旨在分析中国数据管理解决方案的概念定义、应用前景、技术动向及发展趋势,并识别中国数据管理解决方案市场竞争态势,反映该细分市场领袖梯队品牌的差异化竞争优势。

沙利文联合头豹研究院对数据湖、数据仓库及数据库进行了下游用户体验调查。受访者来自金融、消费、媒体、运营商、制造、物流等多个行业,所在公司规模不一,细分领域有别。

本市场报告提供的数据管理解决方案发展趋势分析亦反映出数据库行业整体的动向。报告最终对市场排名、领袖梯队的判断仅适用于本年度中国数据管理解决方案发展周期。

本报告所有图、表、文字中的数据均源自弗若斯特沙利文咨询(中国)及头豹研究院调查,数据均采用四舍五入,小数计一位。

# 概览摘要

## ■ 数据管理解决方案市场需求有望持续扩容

企业用户将逐渐加大在数据管理解决方案的投入,数据希望实现决策和运营效率的提升以建立竞争优势,这种意愿将快速拉升数据管理解决方案市场需求。数据管理人才逐渐成为数字转型企业的领导者,进一步扩大并巩固对数据管理方案的需求。

中国:数据管理系列

## ■ 国家政策逐渐完善持续增益数字经济发展

随着国家政策在个人信息保护、数据跨境流动、国家信息安全等方面的立法探索,企业与消费者对数据的重视度将进入新的阶段,激活对数据权利、数据隐私、数据安全等话题的讨论,倒逼数据管理解决方案的提供商在数据的数量、种类、速度和价值的产品维度之上,新增安全性作为新的开发方向。

## ■ 中国数据管理云部署方案将成为主流

在存算分离理念的基础上, 云化服务完美应对了弹性扩展、功能迭代、成本控制等特性需求, 在资源需求差异化的场景中实现资源的合理配置。按需所取, 按需付费的云化服务显著降低了部署成本。

## ■ 数据管理解决方案产品"走向"湖仓一体化

对数据管理解决方案的需求逐渐聚焦于数据价值损失的尽量避免和商业决策的更高支持,这种需求的不断清晰促使数据管理解决方案产品"走向"湖仓一体。赋能提升企业的数据价值创造力。

## ■ 数据管理解决方案产品与应用场景融合加深

伴隨数据管理解决方案的底层设施逐渐成熟,技术升级和服务应用场景拓展的加速及深化,数据管理解决方案行业应用逐渐向各领域的核心业务延伸。工业大数据应用深化,推进"中国制造2025";金融行业大数据应用拥有广阔前景。

**沙利**文

## 中国:数据管理系列

# 目录

<b>•</b>	名词解释		 05
•	中国数据	管理解决方案市场综述	 07
	•	概念界定	 08
	•	典型应用	 11
<b>•</b>	数据管理的	解决方案价值创造分析	 13
		行业需求分析	 14
	•	关键价值节点分析	 15
	•	商业实践分析	 16
<b>•</b>	中国数据	管理市场分析	 18
	•	市场规模分析	 19
	•	市场需求洞察	 20
	•	企业对数据管理解决方案认知分析	 21
	•	政策分析	 22
<b>•</b>	中国数据	管理解决方案发展趋势	 23
		数据管理解决方案发展里程碑	 24
	•	云部署将成为主流	 25
	•	湖仓一体化	 26
	•	数据管理解决方案与应用场景融合加深	 27
•	数据管理的	解决方案市场竞争分析	 28
	•	综合评分表现	 29
	•	中国数据管理解决方案领导者	 32
•	方法论		 35
<b>•</b>	法律声明		 36



沙利文市场研读 | 2021/4

## 中国:数据管理系列

# **Contents**

<b>•</b>	Terms	 05
<b>•</b>	Overview of China Data Management Solution Market	 07
	• Definition	 08
	Typical Applications	 11
<b>•</b>	Analysis on Data Management Solution Value Creation	 13
	Industry Demand Analysis	 14
	Value Chain Factors Analysis	 15
	Business Practice Analysis	 16
<b>•</b>	Market Size of China Data Management Solution	 18
	Market Size Analysis	 19
	User Demand Insights	 20
	Analysis on Enterprises' Perception	 21
	Policy Analysis	 22
<b>•</b>	Development Prospect of China Data Management Solution Market	 23
	Key Milestones	 24
	Cloud Deployment	 25
	Integration of Data Lake and Data Warehouse	 26
	Deepening Application Scenarios of Data Management Solution	 27
•	Competition Analysis of China Data Management Solution Market	 28
	Comprehensive Vendors Assessment	 29
	Leading Competitors	 32
<b>♦</b>	Methodology	 35
<b>•</b>	Legal Statement	 36



沙利文市场研读 | 2021/4 中国:数据管理系列

# 名词解释

- ◆ 大数据: Big Data, 无法在一定时间范围内用常规软件工具进行捕捉、管理和处理的数据集合。
- 元数据:Metadata,又称中介数据、中继数据,为描述数据的数据(data about data),主要是描述数据属性 (property)的信息,用来支持如指示存储位置、历史数据、资源查找、文件记录等功能。
- ◆ 主数据: MD Master Data, 指系统间共享数据(例如, 客户、供应商、账户和组织部门相关数据)。
- ◆ 结构化数据:又称行数据,结构化数据由明确定义的数据类型组成,由二维表结构实现逻辑表达。结构化数据严格 遵循数据格式与长度规范。主要通过关系型数据库进行存储和管理。
- ◆ 非结构化数据:数据结构不规则或不完整,没有预定义的数据模型,不方便用数据库二维逻辑表来表现的数据,非 结构化数据包括所有格式的办公文档、文本、图片、图像和音频/视频信息等等。
- ◆ 数据仓库:Data Warehouse,是为企业所有级别的决策制定过程,提供所有类型数据支持的战略集合。
- ◆ 数据湖:Data Lake,是一个集中式存储库,允许您以任意规模存储所有结构化和非结构化数据。
- ◆ 高级分析:主要通过深度学习、数据准备自动化、自然语言处理技术将数据转化为可实现的商业价值。
- Hadoop: 是一种分布式系统架构,充分利用集群的威力进行高速运算和储存。核心设计框架是为海量的数据提供存储的HDFS和为海量的数据提供计算的MapReduce。
- ◆ 联机事务处理 : OLTP,是一个在线交易处理系统。OLTP系统的主要是记录事务时的当前更新,插入和删除。
- ◆ 联机分析处理:OLAP是一个在线分析处理系统。存储输入的历史数据。它允许用户查看多维数据的不同摘要。
- ◆ 数据挖掘: Data Mining指从大量的数据中通过算法搜索隐藏于其中信息的过程。
- ◆ 决策支持系统: DSS. 是解决非结构化问题, 服务干高层决策的管理信息系统。
- ◆ 主管资讯系统: EIS. 是服务于组织的高层经理的一类特殊的信息系统。
- ◆ 商业智能: Business Intelligence指用现代数据仓库技术、线上分析处理技术、数据挖掘和数据展现技术进行数据分析以实现商业价值。
- ◆ 数据重力:强调数据在哪里收集就应该在哪里处理,这样操作的效率和成本都是最优的。换句话说,就是不必把数据运送到处理所在的位置,而是把处理推送到数据所在的位置。
- ◆ 数据智能: Data Intelligence, 是以数据作为生产资料,通过结合大规模数据处理、数据挖掘、机器学习、人机交互、可视化等多种技术,从大量的数据中提炼、发掘、获取知识,为人们在制定决策时提供有效的数据智能支持,减少或者消除不确定性。
- ◆ 数据资产:Data Asset 指由企业拥有或控制的,能为企业带来未来经济利益的,以物理或电子方式记录的数据资源。
- ◆ 数据治理: Data Governance, 作为数据管理的一个核心职能, 对数据资产管理行使权力和控制的活动集合。





沙利文市场研读 | 2021/4 中国:数据管理系列

# 名词解释

- ◆ 公有云: Public Clouds 指第三方提供商用户能够使用的云,通过互联网使用,可能是免费的或成本低廉的。
- ◆ 私有云: Private Clouds 指为一个客户单独使用而搭建的,公司拥有基础设施,或部署在一个安全的主机托管场所。
- ◆ 混合云:Hybrid Cloud 是公有云和私有云两种服务方式的结合,提供弹性需求,同时保障安全和控制。
- ◆ 数据沙箱: Data Sandbox, 通过实现数据所有权和数据使用权分离,以确保数据流通过程安全可控的产品。
- ◆ 流数据(或数据流):指在时间分布和数量上趋于无限的一系列动态数据集合体,数据的价值随着时间的流逝而降低、因此必须实时计算给出秒级响应。
- ★ 流计算:对流数据执行实时加载实时计算,对包含单条记录或几条记录的微批量数据进行相对简单的模型计算。
- ◆ 并行计算:一种一次可执行多个指令的算法,目的是提高计算速度,及通过扩大问题求解规模,解决大型而复杂的 计算问题。
- ◆ 分布式计算:Distributed computing,是一种把需要进行大量计算的工程数据分割成小块,由多台计算机分别计算, 在上传运算结果后,将结果统一合并得出数据结论的科学。
- ◆ 内存内计算:把计算单元嵌入到内存里面,这样的话内存既是一个存储器,也是一个计算机,它并不需要从内存中读取数据,数据是直接进出CPU的。不但不受内存的性能限制,而且还提高了效能比。







中国数据管理解决方案市场综述

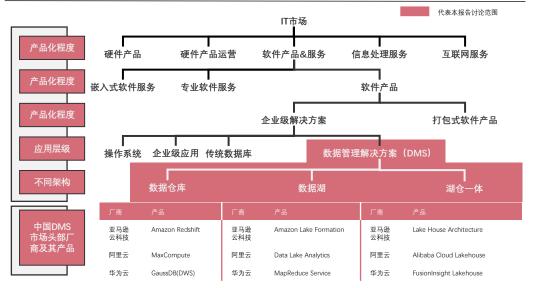
- ◆概念界定
- ◆典型应用

沙利文市场研读 | 2021/4 中国:数据管理系列

# 数据管理解决方案概念界定

沙利文将数据管理解决方案 (DMS, Database Management Solutions) 定义为由服务提供商向公司与组织提供的一站式数据管理系统

## 数据管理解决方案



#### ■ 数据仓库、数据湖构成核心模块

数据湖和数据仓库是实现正式数据管理解决方案的两种主流架构:

数据仓库:专注处理结构化数据,重视大数据处理的效率,偏重于企业数据管理的成长性。

数据湖:储存海量的实时的原始数据、兼容非结构化数据、偏重于增加企业数据管理的灵活性。

#### 数据管理解决方案的行业价值

数据湖和数据仓为各行业提供数据资产化的基础,数据的资产化将从营销、研发、供应链等关键价值节点重构企业价值链。

数据管理解决方案利用计算机硬件和软件技术对海量数据进行有效的收集、存储、计算、分析和应用的过程,旨在从原始数据中抽取、推导出有价值的信息,以支撑企业决策。

数据管理解决方案的目的并非简单的整理和存储数据,而是实现数据的高级分析,高级分析将 直接为企业提供更及时的决策和观察。数据湖和数据仓库通过连结数据管理解决方案要素和提 供数据高级分析的基础,实时支持企业决策,加速了企业数据的价值创造。



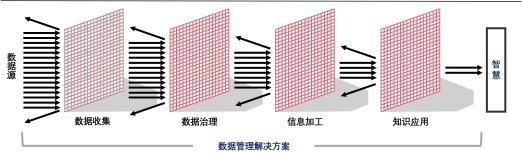


沙利文市场研读 | 2021/4 中国:数据管理系列

# 数据管理解决方案概念界定

数据管理解决方案利用计算机硬件和软件技术对海量数据进行有效的收集、存储、计算、 分析和应用的过程,旨在从原始数据中抽取、推导出有价值的信息,以支撑企业决策

## 数据管理解决方案



## 数据 现象采集后变为数据

- 结构化数据
- csv./ison. 非结构化数据
- 文本/图片/影片
- 半结构化数据 xml./html.

信息

- Flink/ Storm 并行计算
- HDFS/HBase
- 分布式计算
- Yarn/Spark 内存计算
  - Spark/SAP HANA

# 数据清洗后变为信息

## 信息关联后变为知识 关系型信息整合

- 二维表格模型
- 非关系型信息整合
  - 键值型数据库
  - 列储存数据库
  - 文档型数据库
    - 图形数据库

## 智慧

## 知识应用后变为智慧

- 实时决策
  - RTDSS决策支持 EIS主管咨讯
- BI商业智能
- 机器学习

- 数据智能
- 数据沙箱

## 数据管理解决方案让毫无意义的"数据"转化为让企业价值增长的"智慧"

#### 打包数据转化流程,一键式导出

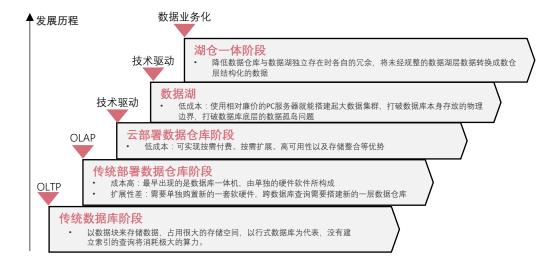
数据源的种类多,数据的类型繁杂,数据量大,并且产生的速度快。企业在处理数据过程中, 一方面需要保证数据转化的可靠性和高效性,另一方面需要控制操作成本。数据管理解决方 案通过专业的软件硬件技术与实施计划,能够提供高性价比、快速、准确的数据转化效果。

#### 拓宽现象与智慧的通道,加速形成竞争优势

在未来的行业竞争中,需求洞察、产品制造、审发营销、用户追踪等关键职能都离不开企业 系统、机器系统、互联网系统、社交系统所产生的海量数据。通过数据湖、数据仓库的应用、 打破系统之间的物理壁垒,优先整理出与现实最匹配的行业数据、业务数据、内容数据、线 上行为数据、线下行为数据,掌握第一手知识。帮助决策者先发占领行业新兴市场高地,并 为行业竞争动态提供最新视角, 提高企业的决策弹性。

# 数据管理解决方案概念界定

数据湖和数据仓库的边界正在慢慢模糊,数据湖自身的治理能力、数据仓库延伸到外部存储的能力都在加强,湖仓一体的出现让数据管理的灵活性与成长性得到了统一



#### ■ 成本驱动

数据管理面对的根本问题是对数据的存储与查询。

在传统数据库阶段,存储与查询都面临着巨大的成本与难度;传统部署的数据仓库阶段,提高了对数据的治理能力降低查询的成本与难度,但可扩展性的局限决定了其降低查询成本的能力下确界;云部署数据仓库阶段,大大拓宽了可扩展性的边界,且免去托管、运维、软件投入等成本,按需收费实现高资源利用率,但无法解决数据仓库本身对非结构化数据不兼容的属性;数据湖的实践,一举实现了存储性能的飞跃,兼容实时的、海量的、类型广泛的数据,且真正打破了数据库之间的物理壁垒;湖仓一体的出现,吸收了数据湖对存储的优势和数据仓库对查询的优势,将大数据的使用门槛进一步降低。

#### ■ 企业用户的需求进程

对调用数据的灵活性和成长性需求、会随着企业用户的发展不断的演进。

初创阶段的企业,数据从产生到消费的周期还很长,往往只需要联机事务处理(OLTP)系统记录业务事件,这是传统数据库的应用;而对不同业务的数据集中分析,才需要对数据清洗后储存在数据仓库中,提供联机分析处理(OLAP)分析,为企业提供成长性空间;当业务成长到一定的规模后,企业需要对海量且多样的数据进行储存与分析,强调对灵活性的需求,数据湖的优势凸显。分析方法则因为数据量的增加可以拓展到数据挖掘(Data Mining),进而支持决策支持系统(DSS)、主管资讯系统(EIS)分析出更具价值的信息与知识,帮助构建商业智能(BI)。

而对于主流互联网企业的营销和运营等场景、电信行业的经营分析与金融行业的风控管理这些应用场景中,数据湖对海量数据 存储能力与数据仓库处理高度结构化数据的提取能力都变得十分重要。在数据重力的作用下,巨量的数据传输成本已将实际的 业务抵在重压之下,湖仓一体正是在数据业务化的需求中诞生。

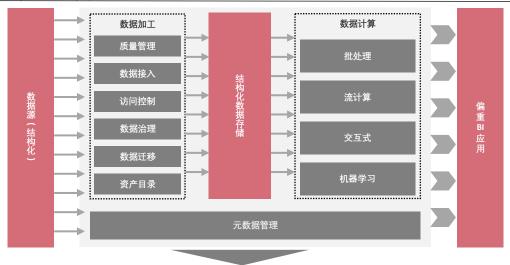


中国:数据管理系列

# 数据管理解决方案典型应用(1/2)

数据湖和数据仓库是实现正式数据管理解决方案的两种主流架构,数据仓库更重视大数据处理的效率,偏重于企业成长性

数据管理解决方案的两种典型应用——数据仓库架构



位于多个数据上的大容量存储区,以实现业务型、交易型数据记录和查询功能

#### 本质

- ▶ 内置的存储系统,数据通过抽象的方式提供(例如采用Table或者View),不暴露文件系统
- 数据需要清洗和转化,通常采用ETL/ELT方式
- ▶ 强调建模和数据管理,供商业智能决策

#### 描述

▶ 数据

来自事物系统、运营数据库和业务线应用程序的关系型数据

> 数据质量

可作为重要事实依据的高度监管数据

Schema

设计在数据仓库实施之前 (写入型Schema)

分析

批处理报告、BI和可视化

▶ 用户

业务分析师

▶ 性价比

更快的查询结果只需要较低的存储成本

## 优势

- > 引擎深度理解数据,存储和计算可做深度优化
- > 数据全生命周期管理,完善的血缘体系

- 细粒度的数据管理和治理
- 完善的元数据管理能力,易于构建企业级数据中台

中国:数据管理系列

数据来源:大数据行业典型企业官网、产品手册、年报、沙利文整理

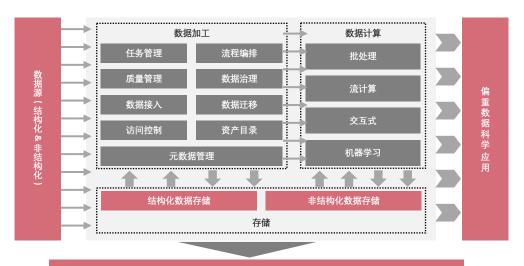


沙利文

# 数据管理解决方案典型应用(2/2)

数据湖和数据仓库是实现正式数据管理解决方案的两种主流架构,数据湖可处理非 结构化数据。偏重干增加企业数据管理灵活性

数据管理解决方案的两种典型应用——数据湖架构



一个集中式存储区,可存储所有类型数据,能实现对非结构化数据的深入分析

## 本质

- 统一的存储系统
- > 存储原始数据

- ▶ 丰富的计算模型/范式
- 数据湖与上云无关

## 描述

数据

来自设备、网站、应用和媒体等的关系型数据及非关系型数据

数据质量

任何无法进行监管的数据 (例如原始数据)

Schema

写入在分析时(读取型Schema)

#### 分析

机器学习、预测分析、数据发现和数据分析

中国:数据管理系列

数据科学家、数据开发人员和业务分析师

性价比.

更快的查询结果只需要较低的存储成本

## 优势

- 收集和摄入所有数据源,获得整个无孤岛的数据集合
- 支持对实时和高速数据流执行ETL(抽取-转置-加载)功能 > 具有人工智能的高级分析
- > 可扩展性和灵活性

数据来源:大数据行业典型企业官网、产品手册、年报、沙利文整理



利 文

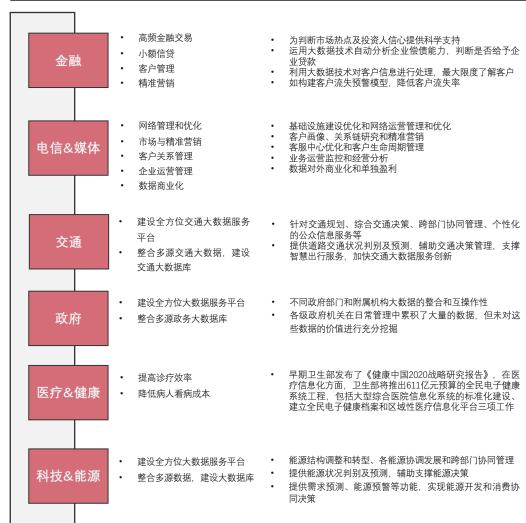


# 数据管理解决方案价值创造分析

- ◆行业需求分析
- ◆价值节点分析
- ◆商业实践分析

# 数据管理解决方案价值创造分析——行业需求分析

## 各行业大数据管理需求



数据来源:阿里云、华为云、腾讯云、亚马逊云科技、IBM等官网,沙利文整理

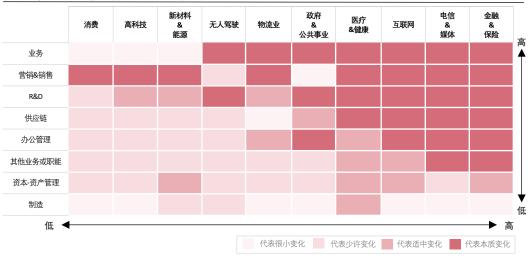




# 数据管理解决方案价值创造分析——关键价值节点分析

数据湖和数据仓为各行业提供数据资产化的基础,数据的资产化将从营销、研发、供应链等关键价值节点重构企业价值链,构建数据生态圈良性循环

## 大数据管理对行业核心价值节点带来的变化



数据来源:沙利文

## ■ 价值节点变化

大数据管理本质改变了各行业对业务、营销和研发的价值节点实践。数据价值体现在展示现象的发展过程。描述现象的发展本质并对现象发展的趋势作出预测。

转化数据推导出相关方向上的预测、趋势分析、交叉销售策略推荐、客群画像匹配,进而掌握 客户真正在意的产品和服务的要素。这种数据智能的应用将以数据为生产资料,驱动商业重塑。

## ■ 价值链变化

数据云化的核心思想是"一切皆服务",解放生产力,让企业可以专注在其最核心的领域内。

数据的资产化不仅仅对营销与研发人员产生变化,下游对数据的监测、洞察能够为在价值链条的其他功能构成比如仓储、生产计划、人员管理等提供优化建议。同时,数据的远程同步与实时可达性,能够让价值链的任何一个部分都能随时随地调整功能决策。

分布式创新潜能得到释放,在改善了客户体验的同时,在价值链上的每一个环节都开辟出新的价值创造空间,驱动各类参与者创造出新效率。

#### ■ 数据生态圈良性循环

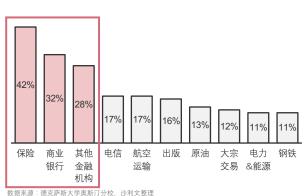
数据的连通在横向维度上,一方面在外部能够集结不同行业的一系列客户,另一方面在内部能够促进价值链条上不同环节的紧密协作;在纵向维度上,让生态圈的参与者在其角色中增强体验,甚至在其中发挥主导作用。

通过网络效应, 生态圈能够提供单个企业凭借自身无法提供的产品和服务, 从而吸引新的客户, 生成更多的数据。

# 数据管理解决方案价值创造分析——商业实践分析

数据管理对各行业收益均有明显的放大作用,其中在保险业和金融业的效果最为明显; 更及时的决策/观察是数据管理的主要受益来源

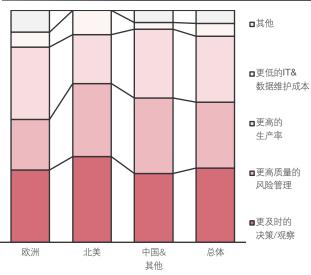
各行业每10%数据管理的投入对资产收益率的提升 [百分比]



## 保险、商业银行、金融机构的行业共性:

中国:数据管理系列

## 正式数据管理解决方案的主要收益来源占比



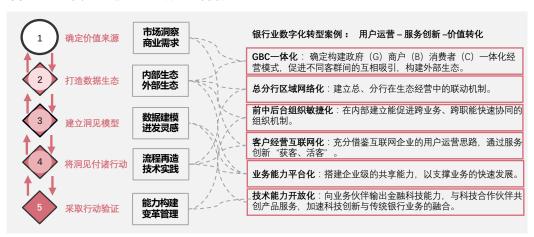
数据来源:沙利文

## 数据管理解决方案产生收益的三种方式:

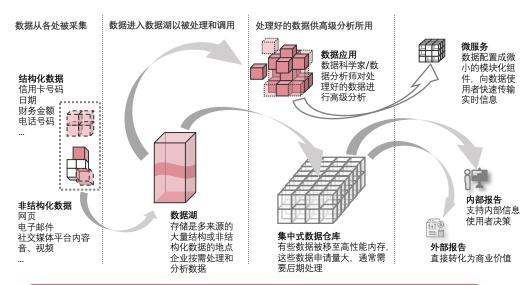
## 无论何种增长方式都离不开大数据计算能力:

# 数据管理解决方案价值创造分析——商业实践分析

数据湖和数据仓库通过连结五大数据管理要素和提供数据高级分析的基础,实时支 持企业决策。加速了企业数据的价值创造



## 数据湖和数据仓库作为加速器参与大数据管理



数据湖和数据仓直接加速了数据的采集、存储、计算、分析和调用





中国:数据管理系列



# 中国数据管理市场规模

- ◆市场规模分析
- ◆市场需求洞察
- ◆企业对数据管理解决方案认知分析
- ◆政策分析

# 中国数据管理解决方案市场规模——市场规模分析

利好政策的持续颁布以及大数据技术的创新融合推动中国的大数据产业快速发展, 未来伴随更多的数据应用场景逐渐落地,数据管理解决方案市场有望持续扩容

中国数据管理解决方案市场规模。2015-2024年预测

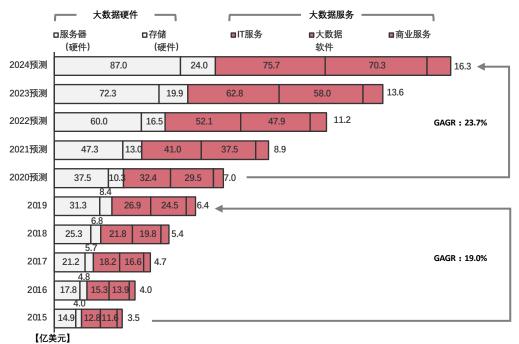
#### □ 大数据硬件

#### □ 大数据服务

#### □ 数据管理解决方案市场

- 预计2024年中国大数据硬件 市场规模为111.0亿美元
- 预计2024年中国大数据服务市 场规模为**162.3**亿美元
- 预计2024年中国大数据整体市场规模为 273.3亿美元

中国:数据管理系列



#### ■ 市场需求扩大

企业用户将逐渐加大在数据管理解决方案的投入,数据希望实现决策和运营效率的提升以建立竞争优势,这种意愿将快速拉升 数据管理解决方案市场需求。数据管理人才逐渐成为数字转型企业的领导者,进一步扩大并巩固对数据管理方案的需求。

#### ■ 数据管理解决方案趋于成熟,性价比高

数据仓库与数据湖已经成熟,市场竞品丰富,产品安全稳定功能齐全,能为各行业各企业的需求提供差异化定制。湖仓一体的发展结合了两个成熟产品的优势,满足灵活性、成本、性能、安全、治理等企业级特性,同时进一步降低整体成本。

#### ■ 国家安全政策利好

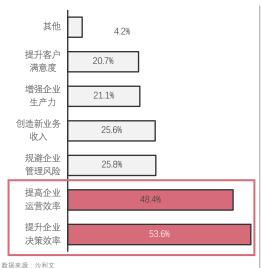
随着国家政策在个人信息保护、数据跨境流动、国家信息安全等方面的立法探索,企业与消费者对数据的重视度将进入新的阶段,激活对数据权利、数据隐私、数据安全等话题的讨论,倒逼数据管理解决方案的提供商在数据的数量、种类、速度和价值的产品维度之上,新增安全性作为新的开发方向。同时,对国家数据安全的考虑,中国厂商将得到远高于外资厂商将会得到的市场与政策支持。

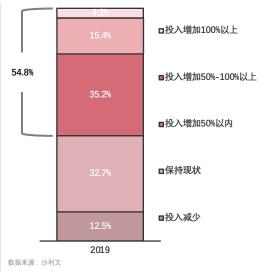
# 中国数据管理解决方案市场规模——市场需求洞察

企业用户将逐渐加大在数据管理解决方案的投入,希望实现决策和运营效率的提升以建立竞争优势,这种意愿将快速拉升数据管理解决方案市场需求

2020年中国大数据服务企业用户需求动因

2019年中国企业用户在大数据服务的投入意愿





#### ■ 提升企业决策效率和运营效率是企业用户的首要需求

企业在运营、决策、风险管理、业务创新、生产及客户服务等效率提升意愿明显。

#### ■ 企业用户在数据管理解决方案的投入意愿较高

近55%的企业用户计划增加在数据管理解决方案的投入,将为大数据服务商提供较大发展空间。

## 数据管理解决方案驱动市场需求的逻辑

1. 有效提高财务表现:员工生产效率、净资产收益率、投资回报率、资产利润

2. 有效增进客户关系:增益创新能力获得新产品线的营收增长、拓展客户基础

3. 有效增益经营管理:资源利用水平、预测与生产计划、交付周期、服务条款

#### ■ 边际投入兑换数据红利

不同程度对数据管理解决方案的投入,都会首先获得基础的数据存取与数据质量的明显增益。 随着投入程度的增加,数据的不同属性包括质量、易用性、智能度、可及性和灵活性都会得到 相应比例的增益。企业通过调整不同比例的投入与高级分析的策略,建立在行业中的竞争优势。



沙利文市场研读 | 2021/4 中国:数据管理系列

# 企业对数据管理解决方案认知分析

企业对数据管理人才的认知从专家型转变为领导者,随着企业对数据管理的认知发生变化,数据管理人才在企业的地位正在提升

企业对数据管理人才认知的变革

Γ 怪才型 专 1995年 家 型 认 专家型 知 2000年 布道者型 2005年 领 助力者型 导 2010年 者 型 认 管家型 知 2015年 领导型 2020年

将数学与计算机科学转化为战略实践、获取竞争优势。

帮助企业建立专业的数据团队,并将数据科学作为一种专业技能。

协助企业加深对数据与分析价值的了解,并共享数据,企业纷纷开始设置 "首席分析官"一职,其中,金融服务等行业拔得头筹。

帮助企业确立明晰的愿景和路线图,以打造数据驱动型企业,但这类首席分析官通常为自己的"地盘"而争斗,较少站在高管的角度考虑问题。

这类首席分析官常常在公司内部长期担任领导职务,不一定具有技术背景, 但为人处世有方法,致力于在组织层面帮助公司处理阻碍发展的问题。

以领导者的角色,致力于在企业、IT部门及其自身之间构建"三方合作"以确保业务取得成效,并采用切实可行的方法,达成目标。

#### ■ IT时代讲入DT时代

随着云服务和移动互联的普及,数据制造的能力向中小企业和消费者转移,各行各业竞争需要精细化运营,通过大数据分析洞察需求,生成情景化的预测性知识以建立差异化竞争力。

#### ■ 大数据与物联网、人工智能将进一步融合

在新时代的竞争中,领导者应基于混合云、5G、物联网和边缘计算能力重新设计工作流程,提高企业适应力。开发基于数据的AI策略,将数据置于每一项业务决策的核心依据,制定清晰的商业计划,打造认知型企业。

#### ■ 数据决策需要领导者认知

数据决策:一个完整的预测支持决策的回路包括历史数据输入,模型训练,数据预测,决策,执行,结果收集和数据反馈。数据分析支撑管理决策,其前提是细化目的数据需要支撑哪部分的管理决策。比如:新功能上线后的用户活跃度,还需要进行哪些修改;新品的销售占比和区域表现,后期该如何倾斜;会员的购买力分析,如何针对性推送优惠信息等。只有明确了系列目的,数据才有可能针对性地被收集和使用,要求数据领导者做好全局规划,保证数据回路的有效推进。





# 中国数据管理解决方案市场规模——政策分析

伴随数字经济建设速度的加快,政府对大数据行业发展重视程度逐渐提升,大数据 产业多层次政策体系逐渐完善,数据管理解决方案等细分产业持续受益

颁布主体	政策名称	颁布日期	政策要点
深圳市人大常委会	《深圳经济特区数据 条例(草案)》	2020-12	公共数据属于新型国有资产,其数据权归国家所有。深 圳市政府代为行使区域内公共数据的数据权,授权市数 据统筹部门制定公共数据资产管理办法并组织实施。
全国人大常委会	《中国人民共和国个 人信息保护法(草 案)》	2020-10	确立了个人信息处理应遵循的原则,即强调处理个人信息应当采用合法、正当的方式,具有明确、合理的目的, 处理信息应遵循公开、透明的原则。
全国人大常委会	- — — — — — — — 《数据安全法(草 案)》	2020-07	第十条 国家积极开展数据领域国际交流与合作,参与数据安全相关国家规则和标准的制定,促进数据跨境安全、自由流动。
工信部	《关于工业大数据发 展的指导意见》	2020-04	针对工业大数据现阶段,在加快数据汇聚、推动数据共享、深化数据应用、完善数据治理、强化数据安全、促进产业发展六个方面全盘布局,系统推进。
国家互联网信息办公室	《数据安全管理办法(征求意见稿)》	2019-05	维护国家安全、社会公共利益,保护公民、法人和其他 组织在网络空间的合法权益,保障个人信息和重要数据 安全。
国家标准化管理委员会	《信息安全技术大数据交易服务安全要求》	2017-12	有助于理清数据交易安全界限,促进数据交易行为合法 合规,推动中国数据交易机构的安全建设,促进数据交 易行为合法合规,促进全国数据要素有序流通,充分释 放数据红利,助力"数字中国"建设。
工信部	《大数据产业发展规 划(2016-2020年)》	2017-01	以大数据产业发展中的关键问题为出发点和落脚点,以强化大数据产业创新发展能力为核心,推动促进数据开放与共享、加强技术产品研发、深化应用创新,打造数据、技术、应用与安全协同发展的自主产业生态体系。
国家发改委	《关于组织实施促进 大数据发展重大工程 的通知》	2016-01	政府将重点支持大数据示范应用、共享开放、基础设施统筹发展,以及数据要素流通。国家发改委将择优推荐项目进入国家重大建设项目库审核区,并根据资金总体情况予以支持。

数据来源:人大常委会、国务院、发改委、工信部、国家标准委、国家互联网办公室、沙利文整理

## ■ 强调个人信息保护,明确权利属性,数据权立法探索初现

数据作为一种新型生产要素育次被写入其中,与土地、劳动力、资本、技术等其他生产要素并驾齐驱。自然人对个人数据依法享有 数据权;公共数据属于新型国有资产,其数据权归国家所有;要素市场主体也有数据权、任何组织和个人不得侵犯。奠定了发展数 字经济的法理基础,直接要求数据流通市场规范化。未来,可预见面向数据监管、隐私安全、权利归属的解决方案提供商将建立起 一条新的数字经济赛道。

#### ■ 加强数据跨境流动的立法,保障国家信息安全

中国在建立数据跨境流动管理体系,平衡国家安全、个人隐私保护、产业竞争等多方利益、既要满足全球化经济所要求的数据流动 也要满足安全所需对数据的监察控制。国际关系或政治敏感度将成为所有相关企业要重点追踪的维度,直接影响了外资数据管理方 案提供商在中国市场上所需面临的数据跨境流动监管力度的加大。另外,消费者或企业对于数据认知的提升,也将直接影响其选择 同类解决方案的偏好,国产数据管理方案提供商将通过占领部分外企撤出中国市场的份额实现增长。

■ 大数据逐渐向各细分应用领域延伸,强调大数据技术与具体应用场景的结合

"大数据"作为热词早已深入消费者、企业和各地政府的心,回归理性,找到适合企业当前时期且行之有效的数据管理解决方案才 是数字经济发展的大厦根基。政企合作,服务民生问题,建立跨部门的数据互通,加速审批流程,实现惠民便民的大数据应用将成 为服务提供商与开发商的热点,在各个方向重现抗疫期间大数据迸发的应用潜能。





中国:数据管理系列



# 中国数据管理解决方案发展趋势

- ◆数据管理解决方案发展里程碑
- ◆云部署方案将成为主流
- ◆数据管理解决方案产品的湖仓一体化
- ◆数据管理解决方案与应用场景融合加深

# 数据管理解决方案发展里程碑

此前数据管理解决方案的重点分别为数据的数量、种类和速度三个主要阶段,现阶 段数据管理解决方案的重点在于数据的价值、驱动供给侧开发新一代数据库产品

数据管理解决方案的演进

15MB 一家全球性银行的人力 资源数据库

500MB 一家零售企业一个品类 单品年周销售数据

2TB 一家全球性咖啡公司 计划的会员数据

2.5PB 一家全球性超市 的交易数据

中国:数据管理系列

PB/ZB 数 其他新的数据来源 据 量 级 TB 网络 GB 客户关系管理 MB 企业资源规划 客户细分 采购具体信息 产品具体信息 采购记录 客户联络 支付记录 支持联络方式 数量、种类

数据来源:Teradata. 沙利文分析

## 三类软件产品细分的关键差异

代表数据库产品对大数据主流应对方式 NoSQL 海量数据批处理 Hadoop M-R 数据 海量数据管理 NewSQL 价 MonGoDB/Hbase 值 密 数据仓库 度 OldSQL 内存数据分析 DB2 Blu/HANA 传统事务处理 Oracle/DB2 流计算/内存计算 内存关系型 Stream/\$park 数据库

- 用于支持企业价值创造的数据管理在处理数据的 种类、数量和速度上均呈现爆发式增长。
- 数据量级和价值密度的增长将催生对新一代数据 库的需求。

- 开源数据库生态不断完善:如M/SQL、PostgreSQL MongoDB等开源数据库,正在以价格低廉、性能相 等、功能丰富的特点在中小型企业中占领商业数据 库的市场。



ТВ

EΒ

PB

# 云部署方案将成为主流

在存算分离理念的基础上, 云化服务完美应对了弹性扩展、功能迭代、成本控制等特性需求, 在资源需求差异化的场景中实现资源的合理配置

## ■ 具备数据储存计算处理分析能力的完整平台

以往,为了应对网络速度的不足、数据在各节点之间交换时间较长的问题,大数据分布式框架采用存储与计算耦合的形式,使数据在自身的储存点上完成计算,以降低交互。这是传统部署DMS。

#### 存储与计算耦合造成了额外成本

实际业务中,对于数据存储空间与计算能力的需求是各自变化的,使得两类资源的需求配比不可预见,当其一出现资源瓶颈时,资源的横向扩展必然导致储存或者是计算能力的冗余,并且对数据的迁移也造成了额外的成本。

#### ■ 存算分离有效控制成本

将储存和计算两个环节剥离形成两个独立的资源集合, 互不干涉又通力合作, 每个资源内部充分体现资源的规模聚集效应, 使得单位资源的成本尽量减少, 同时兼具充分的弹性以供横向扩展。当出现资源紧缺或富裕的时候, 只需要对资源进行获取或回收, 使用具备特定资源配比的专用节点进行, 减少冗余, 在资源需求差异化的场景中实现资源的合理配置。

#### ■ 按需所取的云化服务优势显著

在存算分离的基础上,Serverless、云原生的概念的提出,使数据的处理分析能力摆脱了对于完整平台和工具的需求,大大降低了开发周期、节省开发成本,同时服务应用由提供方运维,实行按需付费,消除了复杂的运维过程和成本。

#### 数据管理系统部署方案

# 战略 ・ 商业实质 ・ 产品生命周期 ・ 主要目标 营销方法 ・ 商销 ・ 销售方法 人力资源 ・ 开发者应有的技能组合 产品开发 ・ 产品特征组合

#### 传统部署DMS

- 推向市场的速度是关键
- 中长期
- 好声誉和良好的客户关系
- 销售优先,产品其次
- 关系型营销
- 仅直接销售
- 沟通和需求满足是关键
- 聚焦于满足个性化需求用户的定制型产品

#### 云部署DMS

推向市场的速度是关键

中国:数据管理系列

- 短期
- 高市场占有率
- 产品优先、销售其次
- 规模营销
- 直接销售和间接销售
- 创造性和编程技巧是关键
- 聚焦于可满足大部分用户需求的通用型产品

机构可以选择云端、本地或混合部署 在公有云、私有云或混合云 之间进行选择 平衡风险控制与成本效率

数据来源:中国信通院,沙利文整理



沙利文

25

# 数据湖与数据仓库的深度融合

对数据管理解决方案的需求逐渐聚焦于数据价值损失的尽量避免和数据对商业决策的更高支持,这两种需求的不断清晰促使数据管理解决方案产品"走向"湖仓一体

#### 数据价值难以有效发挥的原因:

■ 数据孤岛普遍存在,缺乏统一的数据的视图

由于技术或管理制度的原因,企业的数据存在数据孤岛,散落在各个业务系统中。

■ 数据质量低下

数据质量也决定了业务决策的分析,数据资产的价值建立在不断提升的数据质量的前提上。

■ 缺乏安全的数据环境

数据安全风险包括了数据泄露与数据滥用。一旦发生数据安全事件,会对企业经营造成损 失和对用户的隐私造成侵犯,束缚数据价值的释放。

■ 缺乏数据价值管理体系

企业沒有建立起一个有效管理和应用数据的模式,包括数据价值评估、数据成本管理等,对数据的服务和应用缺乏合规性的指导。

## 数据管理解决方案服务架构的变化



■ 提升企业的数据价值创造能力

统一的数据管理解决方案为企业提供数据资产化基础

■ 提升企业的的数据管理解决方案效率

云原生的数据分析核计算架构提高了数据处理的流畅性

数据来源:阿里云、华为云、腾讯云、亚马逊云科技、IBM等官网,中国信通院,沙利文整理

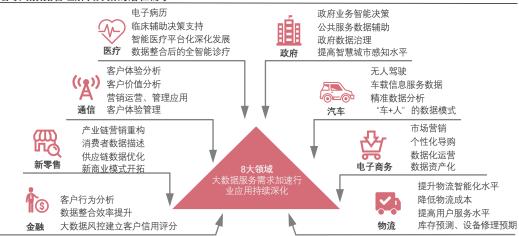


沙利文

# 数据管理解决方案将更加专业化

伴随数据管理解决方案的底层设施逐渐成熟,技术升级和服务应用场景拓展的加速 及深化、数据管理解决方案行业应用逐渐向各领域的核心业务延伸

## 各领域数据管理解决方案的潜在需求



#### ■ 大数据产业

大数据产业是以数据及数据所蕴含的信息价值为核心生产要素,通过数据技术、数据产品、数据服务等形式,使数据与信息价值在各行业经济活动中得到充分释放的赋能型产业。

#### ■ 衍生业态为各行业提供大数据融合应用

大数据融合应用在各行业、各领域中,提供与行业紧密相关的应用软件和整体解决方案。需 求逐渐由感知型应用转向预测型和决策型应用。

## ■ 工业大数据应用深化,推进"中国制造2025"

5G、物联网、边缘计算、区块链等技术逐步融入工业领域,支持企业敏捷管理和精细化运营的大数据管理解决方案推动制造业转型升级提速,更高质量地支撑工业大数据采集、汇聚、流通、分析、应用的价值闭环、助力更高效的产品工艺研发,实现数据价值变现。

#### 金融大数据应用拥有广阔前景

金融行业同时储存了大量的关于客户、账户、产品、交易等结构化数据和反映客户偏好、社会关系、消费习惯等信息的包括语音、图像、视频等非结构化数据。数据管理解决方案在金融业的深度应用,在交易欺诈识别、精准营销、黑产防范、信贷风险评估、供应链金融、投资行情预测等多领域的具体业务中不断释放数据价值。打破金融机构之间数据壁垒,实现跨机构的数据融合将成为趋势。





中国:数据管理系列



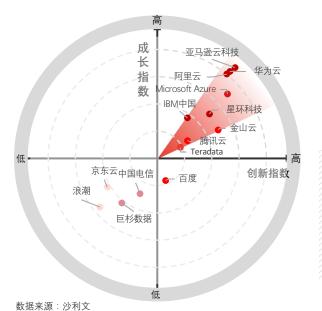
数据管理解决方案市场竞争分析

- ◆综合评分表现
- ◆中国数据管理解决方案领导者

# 综合竞争表现

中国数据管理解决方案市场处于平稳增长阶段,竞争主体在产品创新能力、成长能力和基础能力三个维度各具竞争优势

中国数据管理解决方案市场综合竞争表现——Frost Radar (弗若斯特雷达)TM



注:圆环按由内向外递增的逻辑对应由低至高的综合评分,竞争力由"创新指数"、"成长指数"以及"基础指数"综合得出(圆环仅适用于第一象限竞争者)。

中国:数据管理系列

中国数据管理解决方案市场发展处于平稳增长期 本报告对竞争主体数据管理解决方案产品和服务 综合竞争力的分析结论仅适用于该阶段数据管理解决方案市场发展情况。

沙利文将持续关注数据管理解决方案市场, 捕捉竞争动向。

## 代表领导者区间

位于该区间的厂商为中国数据管理方案市场领导者

- 横坐标代表"创新指数":
- 衡量竞争主体在数据管理解决方案的创新能力,位置越靠右侧,数据管理解决方案服务职能丰富度和产品调优能力越强
- 纵坐标代表"成长指数":
- 衡量竞争主体在数据管理解决方案产品在产品功能、性能成长维度的竞争力、位置越靠上方数据管理解决方案产品成长能力越强
- 色深代表"基础指数":
- 衡量竞争主体的数据管理解决方案产品在基础能力维度的竞争力,色深越深,能力越强

# 评分维度说明(1/2)

本报告设定成长指数(功能成长)、创新指数(能力创新)、基本指数(基本数据分析能力)、对竞争主体旗下数据管理解决方案竞争力进行评估

分析能力),对竞争王体旗下数据官理解决万案竞争刀进行评估				
评价维度	一级指标	要点		
	与AI结合程度	机器学习算法能力 对语言文本的理解能力 多语种的语义理解 存储资源利用 数据资源利用 多温数据管理 易用性(易用易管理) 数据模型灵活度(多维分析系统) 数据查询自由度		
	大数据的支持能力	超大规模存储能力		
	数据获取能力	数据源获取能力 获取全量/增量数据能力		
	数据调用能力	结果推送到合适存储引擎能力 数据分析实时度 并发查询数		
	数据集成能力	数据集市集成能力		
	服务能力	SLA灵活度		
创新指数 (能力创新)	外部兼容能力	支持主流云平台(如亚马逊云科技、阿里云、华 为云、腾讯云等)的能力 支持传统数据库(如Oracle、MongoDB、DB2、 Redis、MySQL等)的能力		
一 (月とノ」とり初门	开源支持能力	支持主流开源社区(Hadoop生态和Spark生态)的能力 支持主流BI工具(Tableau、SAS、Zeppelin等) 的能力 支持主流开源框架(Tensorflow、Pytorch、 MXNet等)的能力 支持其它开源组件(Kafka、flume等)的能力		
	数据湖仓一体能力	支持ACID事务 全局数据可见可查 统一的数据开发体系 统一的数据开发体系 数据湖、仓统一的数据集成 多湖多仓的关联计算 数据虚拟化 数据湖、仓统一的数据目录 统一的数据存放 统一的数据调用 加速数据移动 数据开放能力 第三方能力层对接能力 权限管理		

中国:数据管理系列

沙利文市场研读 | 2021/4 中国:数据管理系列

# 评分维度说明(2/2)

本报告设定成长指数(功能成长)、创新指数(能力创新)、基本指数(基本数据分析能力),对竞争主体旗下数据管理解决方案竞争力进行评估

评价维度 一级指标 要点

		加拉外
	可扩展性	架构能力
		数据处理模式的扩展能力
		存储能力的可扩展性
		支持主流计算组件
	安全性	符合个人信息保护法、数据安全法
		各环节安全覆盖数据储存、使用、加密、脱敏等
		权限安全系统
成长指数		网络安全防护能力
(功能成长)	数据生命周期存储能力	原始数据、分析中间结果及分析过程存储的能力
	数据管理能力	针对特定计算引擎优化数据管理能力
		覆盖元数据、主数据、数据模型、数据标准、质量标准、数据安全、数据共享及数据可视化管理等职能情况
		数据湖多代共存
		数据仓库多代共存
		数据仓库内分析优化(如是否基于CBO、RBO优化模型)
	平台及生态综合能力	厂商赋能合作伙伴程度
	数据存放能力 (针对自研产品)	结构化数据存放能力
		非结构化数据存放能力
基本指数	基础设施能力	私有云能力
		公有云能力
		自研服务器能力



# 领导者-亚马逊云科技

亚马逊云科技是中国数据管理解决方案领导者,通过技术创新、灵活的数据管理、云上安全、全球商业实践和较强的平台生态综合实力提供高创新性和高成长性的数据管理解决方案

## ■ 技术创新:智能湖仓架构

- 将AI技术与数据管理功能深度融合,提出智能湖仓架构,运维工作可托管,实现数据管理服务的高可用架构、安全认证和细粒度监控、存储和计算分离和资源自动扩展
- Amazon QuickSight Q(机器学习生成的数据模型可自动理解数据含义和关系)

#### ■ 全球商业实践:专用的数据管理服务

亚马逊云科技坚持数据管理的专用性,通过将其全球的最佳实践,跟中国的国情相结合,定义最适合这些行业的中国最佳实践

#### ■ 灵活的数据管理:高易用性

 亚马逊云科技完成节点配置、软件配置、自动化索引编制及提取、数据隔离及安全、 行业合规、集群大小调整、自动不定修补、报警及检测和硬件维护

#### ■ 云上安全:责任共担模型

- 应用亚马逊云科技负责:主机操作系统、虚拟化层、物理基础设施的安全
- 客户负责:上层操作系统以及相关的安全,负责配置亚马逊云科技提供的安全组件

#### ■ 生态合作:为合作伙伴实现商业价值

 为帮助众多合作伙伴提升自身的差异化竞争能力的平台,亚马逊云科技在技术应用、 实施、解决方案落地等多方面赋能合作伙伴,同时也为客户创造更多的实际价值

#### 智能湖仓概述

#### Amazon EMR: 云大数据平台

- 支持多种开放源代码工具(如 Spark、 Hive、HBase、Flink、Hudi和Presto)
- 通过自动执行耗时的任务(如预置容量和 调优集群)设置、操作和扩展大数据环境

Amazon Lake Formation是一项数据湖服务,可在几天内轻松建立安全的数据湖

Amazon Glue是一项无服务器数据集成服务,简化了发现、准备和合并数据以进行分析、机器学习和应用程序开发的工作,且提供数据集成所需的全部功能

#### Amazon Elasticsearch Service: 一项完全 托管的日志分析服务

 提供开放源 Elasticsearch API、托管 Kibana、与 Logstash 和其他 亚马逊云 科技服务的集成以及内置提醒和 SQL 查 询支持

#### Amazon Aurora: 云原生关系型数据库

• 高性价比 (仅为商用数据库 1/10的成本)

具备高性能和高可用性(采用有容错能力 ▲ 并且可自我修复的分布式存储系统)



## Amazon DynamoDB:非关系型数据库

 具有适用Internet规模应用程序的内置 安全性、备份和恢复以及内存中缓存

> Amazon Athena是一种交互式 查询服务,让您能够轻松使用 标准 SQL分析Amazon S3中的 数据

#### Amazon SageMaker: 一个专为机器学习 构建的功能集

帮助数据科学家和开发人员加速标签、 数据准备、功能工程、统计偏差检测、 自动机器学习、训练、调优、托管、 可解释性、监控和工作流等步骤

代表数据重力

# Amazon Redshift:一个可以分析所有数据的数据仓库

 可使用标准SQL在数据仓库、运营数据库和数据湖中查询和合并 FB级结构化和半结构化数据





代表数据湖



# 领导者-华为云

华为云是中国数据管理解决方案领导者,通过技术创新、极致的数据管理、产品全方位安全、 丰富的行业实践和较强的平台生态综合实力提供高创新性和高成长性的数据管理解决方案

- 技术创新:MRS一个架构实现三种数据湖
  - 离线数据湖:开放格式存储引擎,多样性引擎,支撑多种分析工作负载,实现湖仓一体
  - 实时数据湖:实时集成、批流融合、实时更新与删除,全链路实时实现分析业务T+0时效性
  - 逻辑数据湖:数据虚拟化实现湖内外数据统一访问、协同分析

#### ■ 丰富的行业实践:敏锐的需求洞察

在政务(智慧龙岗,一网通办)、运营商(平滑迁移,保护投资)、金融(跨仓协同提效10倍)及大企业等拥有丰富、成熟的商业实践

#### ■ 全方位安全:从基础设施到数据应用访问

网络隔离(支持多平面网络安全)、主机安全(操作系统内核安全加固等)、应用安全(用户级别的访问控制等)、数据安全(多副本、备份容灾保证)和安全认证(统一认证体系)

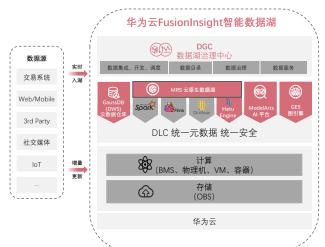
#### ■ 极致的数据管理:满足不同场景下的需求

- 丰富的数据管理职能、数据湖、数据仓多代共存
- 在例如数据的收集、整理、脱敏、分析和管理、安全等都做到了极致

#### ■ 生态合作:全面赋能合作伙伴

- 坚持开放、合作、共贏的云生态,作为"智能世界"的黑土地,帮助合作伙伴快速融入当地生态
- 开放社区贡献(Hadoop社区贡献排名全球第2, Spark社区贡献排名全球第4)

#### 智能湖仓概述



#### MRS云原生数据湖

一**个架构实现三种数据湖**(逻辑数据湖、实时数据湖和离线数据湖)

#### GaussDB(DWS) 云数据仓库

基于Shared-Nothing架构提供开放、全场景分析型数据仓(具备多层级全并行计算引擎和多层级容灾能力)

#### DGC数据湖治理中心

**数据全生命周期一站式开发运营平**台(数据集成、数据开发等功能)

快速构建数据运营能力(支持行业知识库智能化建设、大数据存储及计算分析引擎等数据底座)

#### ModelArts AI平台

面向开发者的一站式AI开发平台(为机器学习与深度学习提供海量数据预处理及半自动化标注、大规模分布式训练等功能且具备端-边-云模型按需部署能力)

## GES图引擎

中国本土首个商用且拥有自主产权的分布式原生图 引擎(提供以"关系"为基础的"图"结构数据的 查询和分析服务)泛用于社交应用、企业关系分析 等具有丰富关系数据的场景



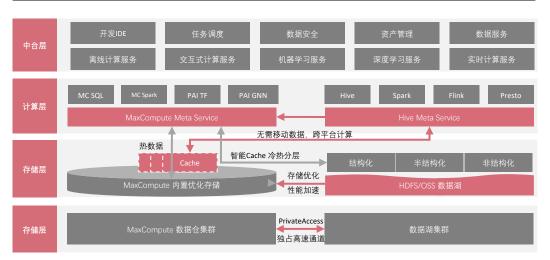


# 领导者-阿里云

阿里云是中国数据管理解决方案领导者,通过技术创新、长足的云原生实践、极致的数据管理、和较强的平台生态综合实力提供高创新性和高成长性的数据管理解决方案

- 技术创新:湖仓一体架构
  - 自创PrivateAccess 网络连通技术,实现整体网络打通能力,快速接入
  - 通过 DB 元数据一键映射技术,统一元数据服务
  - 提供统一的湖仓开发体验,高度开放兼容 Hive/Spark
  - 智能 cache 技术,识别数据冷热度,自动数仓,解决带宽瓶颈
- 十余年沉淀云原生实践:对外赋能数百万家企业进行云原生改造
  - 2009年,阿里首次上线核心中间件系统
  - 2011年,淘宝天猫开始使用容器调度技术,再到推出自研云原生硬件神龙服务器、云原生数据库PolarDB
  - 2019年、双十一、阿里电商核心系统100%上云、这也是全球规模最大的云原生实践
- 极致的数据管理:满足不同场景下的需求
  - 企业级高性能数据仓库,通过云原生服务以更低成本满足业务高弹性和敏捷性需求
  - 弹性资源和EMR集群资源形成互补,两套体系之间进行资源的削峰填谷
  - 基于PAI丰富且强大的算法能力,封装出很多贴近业务场景的算法服务
- 生态合作:城市大脑3.0
  - 将通过城市空间基因库链接农田、建筑、公共交通等全部城市要素
  - 通过人工智能技术、城市大脑可以实现交通、医疗、应急、民生养老、公共服务等全部城市场景的智能化决策

## MaxCompute 湖仓一体概述





沙利文市场研读 | 2021/4

方法论

- ◆弗若斯特沙利文布局中国市场,深入研究10大行业,54个垂直行业的市场变化,已经积累了近 50万行业研究样本,完成近10,000多个独立的研究咨询项目。
- ◆研究院依托中国活跃的经济环境,从数据管理、大数据等领域着手,研究内容覆盖整个行业的发展周期,伴随着行业中企业的创立,发展,扩张,到企业走向上市及上市后的成熟期,研究院的各行业研究员探索和评估行业中多变的产业模式,企业的商业模式和运营模式,以专业的视野解读行业的沿革。
- ◆研究院融合传统与新型的研究方法,采用自主研发的算法,结合行业交叉的大数据,以多元化的调研方法,挖掘定量数据背后的逻辑,分析定性内容背后的观点,客观和真实地阐述行业的现状,前瞻性地预测行业未来的发展趋势,在研究院的每一份研究报告中,完整地呈现行业的过去,现在和未来。
- ◆研究院密切关注行业发展最新动向,报告内容及数据会随着行业发展、技术革新、竞争格局变化、政策法规颁布、市场调研深入,保持不断更新与优化。
- ◆研究院秉承匠心研究,砥砺前行的宗旨,从战略的角度分析行业,从执行的层面阅读行业,为 每一个行业的报告阅读者提供值得品鉴的研究报告。



中国:数据管理系列

沙利文市场研读 | 2021/4 中国:数据管理系列

# 法律声明

◆本报告著作权归头豹所有,未经书面许可,任何机构或个人不得以任何形式翻版、复刻、发表或引用。若征得头豹同意进行引用、刊发的,需在允许的范围内使用,并注明出处为"头豹研究院",且不得对本报告进行任何有悖原意的引用、删节或修改。

- ◆本报告分析师具有专业研究能力,保证报告数据均来自合法合规渠道,观点产出及数据分析基于分析师对行业的客观理解,本报告不受任何第三方授意或影响。
- ◆本报告所涉及的观点或信息仅供参考,不构成任何投资建议。本报告仅在相关法律许可的情况下发放,并仅为提供信息而发放,概不构成任何广告。在法律许可的情况下,头豹可能会为报告中提及的企业提供或争取提供投融资或咨询等相关服务。本报告所指的公司或投资标的的价值、价格及投资收入可升可跌。
- ◆本报告的部分信息来源于公开资料,头豹对该等信息的准确性、完整性或可靠性不做任何保证。 本文所载的资料、意见及推测仅反映头豹于发布本报告当日的判断,过往报告中的描述不应作为 日后的表现依据。在不同时期,头豹可发出与本文所载资料、意见及推测不一致的报告和文章。 头豹不保证本报告所含信息保持在最新状态。同时,头豹对本报告所含信息可在不发出通知的情 形下做出修改,读者应当自行关注相应的更新或修改。任何机构或个人应对其利用本报告的数据、 分析、研究、部分或者全部内容所进行的一切活动负责并承担该等活动所导致的任何损失或伤害。

