

利用图数据平台充分 挖掘数据价值

IDC 信息简报 | 2021 年 6 月



赞助商

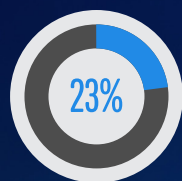
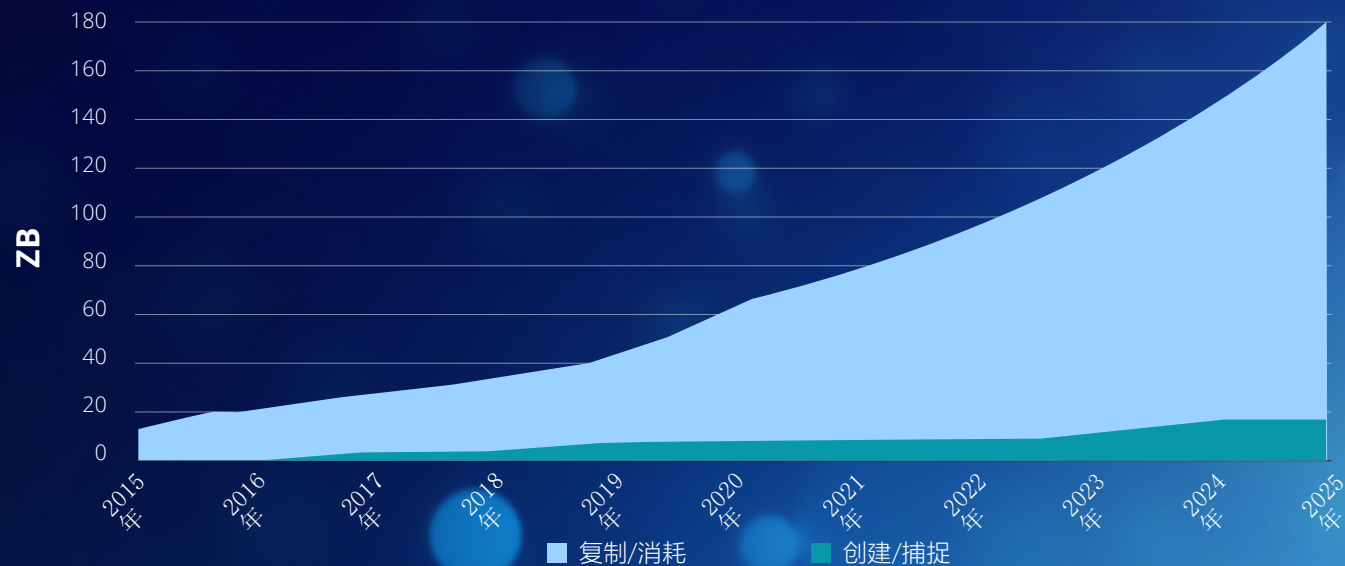


#AP241242IB

数据洪流

据 IDC Global DataSphere 预测，今年全球创建、收集、复制和消费的数据量将超过 80 ZB。在全球数字化、视频、移动和边缘等技术需求的推动下，数据增长势不可挡。然而，数据越来越多也有不好的方面。

IDC Global DataSphere 2020 - 2025



到 2025 年，Global DataSphere 数据将接近 180 ZB，也就是说，2020-2025 的年复合增长率 (CAGR) 为 23%。

来源：IDC Global DataSphere 2021

从数据中提取价值越来越难

注意力让我们能在自己的知识场景下进行数据筛选



到 2024 年, 如果组织无法应对不断增长的数据、分析和决策孤岛, 其出现注意力匮乏的几率将会翻倍, 最终导致无法区分有用信息与干扰信息

知识层次



独立的数据几乎没有价值。要让数据成为对决策者有用的**信息**, **场景**和**相关性**缺一不可。

场景是描述一条数据与其他数据的关系。在没有场景（比如特定行星表面一个世纪内的平均温度增长幅度, 以摄氏度为单位）的情况下, 1.7 这条数据没有任何意义。但场景并不一定代表着**相关性**, 即数据对做出特定决策的用处。如果那个行星是地球, 这条数据便与人类相关。

信息是具体场景中的相关数据, 能影响一个人的决策, 或帮助一个人弄清楚其想要了解的情况。信息帮助人们积累**知识**——关于现实世界的真实且可验证的认知。这些因果知识帮助决策者改变世界并创造**价值**。

要将数据转化为信息、知识和价值, 这个过程离不开决策者的努力——它需要**注意力**和先验知识。但数据洪流使得越来越难以理解相关数据的场景, 更不必说去创造价值。当数据丰富, 但场景匮乏时, 不可避免地会导致个人和组织出现注意力匮乏问题。而组织的注意力匮乏会产生重大不利影响, 既会产生无力感, 又会妨碍区分有用信息与干扰信息。

来源: IDC FutureScope: 2021 年全球智能未来预测——APEJ 的影响

数据存储方式很重要

数字化企业必须将数据转化为切实可行的商业洞见才能创造价值

显然,数据具有“潜在”价值,一些公司能够相较于其他公司从数据中提取更多的价值。要将数据高效地转化为信息、知识和价值,需要考虑很多驱动因素,例如数据驱动的领导方式、公司员工掌握的知识和技能、业务流程以及组织文化。但在从数据到洞见的过程中,另一个重要驱动因素是数据存储技术。从数据中提取价值的过程取决于存储方式、访问的难易程度,以及数据与组织内其他相关数据、信息和知识场景的相关程度。

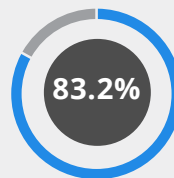
数据库管理系统具有以下三种常见类型:

关系数据库管理系统 (RDBMS) 将数据存储于表格中。当数据之间的关系稳定并且可以在顶层表格定义中捕获时,常常会使用这种类型。关系数据库可以快速处理大量记录,使用的存储空间往往较少。但在处理场景可能随时间变化的动态数据时,灵活性略显不足。

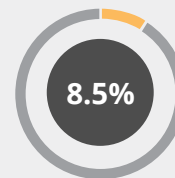
非关系数据库管理系统 (NDBMS) 提供多种非表格形式的替代方法来组织、管理、存储和检索数据。传统上,NDBMS 针对特定数据类型进行了优化,通常是旧式系统,难以处理更新、更动态的工作负载。

动态数据管理系统 (DDMS) 比 RDBMS 更敏捷和高效,最近广受关注,随着其不断发展成熟并克服潜在问题,2016-2019 的 CAGR 增长率达到 83.2%。图数据库就是这样一种动态数据库管理系统。随着图驱动分析和人工智能 (AI) 工具的需求不断增加,对图数据库和图数据库平台的需求也在增加。

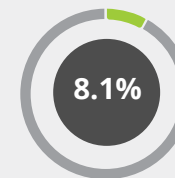
数据库管理系统: CAGR 2016-2019



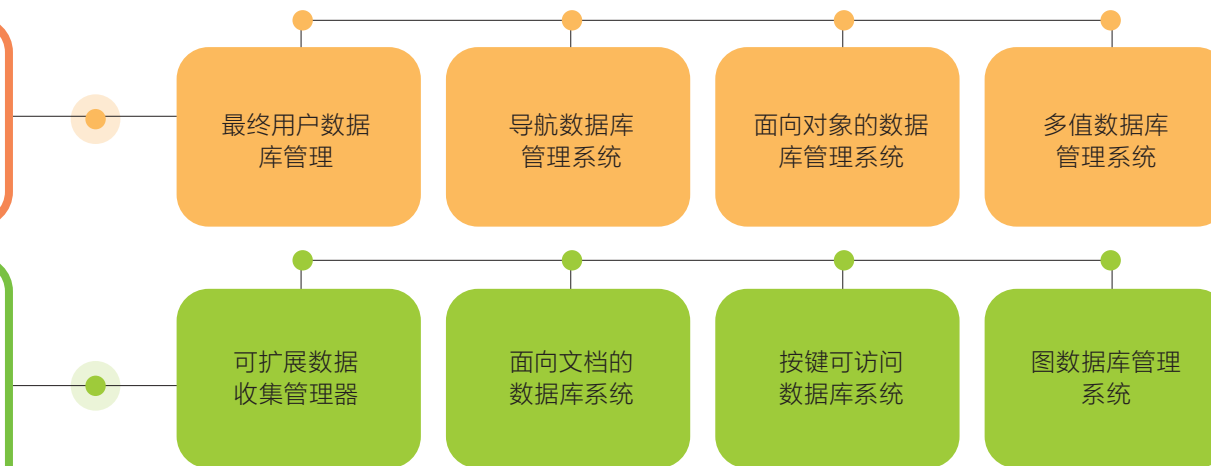
动态数据库
管理系统
(DDMS)



关系数据库
管理系统
(RDBMS)



非关系数据库
管理系统
(NDBMS)



来源: IDC Worldwide Semiannual Software Tracker, 2020

图、图数据库和图数据平台

图数据平台让用户能够在图数据库上开发和运行应用程序



采用图数据平台的驱动因素

图可给企业带来什么好处？



揭示隐藏结构

图可以揭示数据中的隐藏结构，其中数据是未知的，不需要广泛的数据存储，也不需要有关数据组织方式的大量先验知识。

更高的效率

在某些情况下，图数据库管理系统可以比传统 RDBMS 更快、更有效，因为它们能够快速分析模式和关系。图数据库也简化了提取-执行周期，可以更快地执行查询。



广泛的功能

图可以应用于各种问题和应用。其中包括：

- ▶ 发现人、地点或事物之间的关系模式
- ▶ 映射或绘制事物或位置之间的结构或空间关系
- ▶ 从人类语言捕捉语义结构以便系统地理解内容
- ▶ 基于谱系、基于空间和基于执法/情报服务的“已知关联”分析
- ▶ 药理学研究、流行病学和效用网络分析

利用人工智能工具

图很可能在未来的人工智能中发挥关键作用。由于复杂系统中的行动和后果通常会导致数据关系模式发生变化，因此图数据库正在帮助推动机器学习 (ML) 和其他 AI 相关操作实现创新。

图应用 - 为 AI 提供场景

图技术和 AI/ML 具有重要的协同作用

2020 年亚太地区（不包括日本）10 大 AI 应用¹

可由图驱动的 AI 工具在当今 10 大 AI 应用中占大多数。

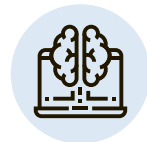
- 1 自动化客户服务代理
- 2 自动化威胁情报和预防系统
- 3 质量管理调查和推荐系统
- 4 数字助理
- 5 欺诈分析和调查
- 6 智能业务创新和自动化
- 7 销售流程推荐和自动化
- 8 诊断和治疗系统
- 9 项目顾问和推荐系统
- 10 自动化预防性维护



到 2022 年, A2000 公司中至少 55% 将在企业范围内利用 NLP、ML 和 DL 等 AI 工具, 在 客户体验、安全、设施和采购等领域实现 55% 的应用²。



将企业数据捕捉为图的知识图谱可以使用 AI/ML 来“理解”数据的场景, 从而实现更快、对场景更敏感的数据搜索。



图技术支持许多重要的人工智能应用, 这些应用依赖于图数据库中已捕捉数据的复杂程度和丰富互连性。例如, 使用智能知识发现工具、推荐引擎和对话式 AI 等等, 用于进行场景搜索、做出更智能的推荐并提供更丰富的洞见。同时, AI/ML 应用会创建更多数据和更复杂的图, 进一步推动 AI 能力。



因此, IDC 认为图数据库正与 AI 同步增长。从 2019 年到 2024 年, 人工智能软件应用正以 **34.1%** 的 CAGR 快速增长。

¹来源: IDC 支出指南, 2021

²来源: IDC FutureScape: 2021 年全球人工智能预测 - APEJ 意义

图应用 - 为 AI 提供场景

按国家划分的 5 大 AI 应用

单位 (百万美元)

	 澳大利亚 (百万美元)	 中国 (百万美元)	 印度 (百万美元)	 韩国 (百万美元)	 新加坡 (百万美元)	其他亚太地区 (百万美元)
1 自动化客户服务代理	28.0	自动化客户服务代理 115.1	自动化客户服务代理 13.2	质量管理调查和推荐系统 8.5	自动化客户服务代理 13.1	自动化客户服务代理 26.0
2 自动化威胁情报和预防系统	27.0	销售流程推荐和自动化 50.0	自动化威胁情报和预防系统 9.4	自动化威胁情报和预防系统 8.4	自动化威胁情报和预防系统 9.6	自动化威胁情报和预防系统 17.7
3 数字助手	18.9	公共安全和应急响应 47.9	质量管理调查和推荐系统 8.1	自动化客户服务代理 7.4	数字助手 7.4	质量管理调查和推荐系统 15.8
4 项目顾问和推荐系统	17.4	智能业务创新和自动化 47.1	数字助手 8.0	自动化预防性维护 5.4	项目顾问和推荐系统 7.3	数字助手 15.2
5 诊断和治疗系统	15.1	诈骗分析和调查 39.0	项目顾问和推荐系统 7.2	项目顾问和推荐系统 5.0	诊断和治疗系统 6.2	项目顾问和推荐系统 13.1

■ 可由图驱动的 AI 工具

来源: 2021年 IDC 支出指南

绘制更大的愿景

图让用户能够检查数据之间的动态、复杂或特殊关系

这对于欺诈检测、反洗钱和供应链可见性等应用特别有用。



检测欺诈和洗钱活动需要了解不同司法管辖区中复杂交易、个人和公司之间的关系。图数据库让用户能够直观地呈现实体之间错综复杂的关系，并提供比传统方法更多的欺诈和洗钱活动线索。图还让用户能够在更广泛的范围内查看关系。在确定特定模式后，图分析工具可以帮助发现类似类型的洗钱或欺诈。得到实际结果（欺诈与否）后，调查人员可以将此图数据输入 ML 模型，以确保预测更加准确。

案例分析

Allianz 借助图技术来发现金融系统中的欺诈活动。检测方法是直观地识别欺诈者试图隐藏的非法联系。由于客户数据存储在图数据库中，Allianz 能够更好地识别和发现风险，避免同一风险重复投保。



供应链是复杂且紧密联系的。只要一个实体出现任何问题，都会对其他产品产生连锁反应。图可以帮助个人从宏观角度直观地检查供应链，识别供应链各个环节之间的复杂关系。图分析工具让企业能够快速搜索大量数据，以确定成本、需求、法规和供应量变化的潜在影响，进而采取缓解措施。



案例分析

Boston Scientific 拥有复杂的全球供应链，从原材料开始，一直到离散产品生产结束。这意味着在审查产品时，如果发现任何缺陷，很难追踪到根本原因。通过使用图技术，Boston Scientific 成功创建出数据模型，帮助他们追踪故障组件并管理到产品。



主数据管理

随着存储的数据越来越多，越来越需要链接各种来源的数据。图可以绘制各种孤岛内数据之间的联系，通过保证数据的一致性和正确沿袭，帮助企业进行数据治理。

主数据管理变得越来越重要。 COVID-19 疫情加速了数字化转型，扩大了组织捕捉的数据量。IDC 将数据定义为三个 D: 分布式 (distributed)、多样化 (diverse) 和动态 (dynamic)。对于被孤立数据困扰的企业，主数据管理是一剂良方。

图可以在共享数据对象 (例如客户、产品和合作伙伴) 之间建立关系，帮助进行主数据管理。 这意味着存储在多个数据源和孤岛的数据可以连接起来，帮助员工更轻松地找到需要的数据。

随着数据不断增长，数据治理变得越来越重要。 图让公司能够更有效地跟踪和直观地呈现数据沿袭关系，同时快速发现数据问题。

案例分析



Lockheed Martin Space 使用图数据库连接所有数据孤岛。通过将数据与各种系统之间的关系存储在图数据库中，他们得以更快地找到问题的答案。Lockheed Martin 表示他们现在能够揭示以前从未发现的联系。



Airbnb 的业务遍布全球，这意味着收集的数据量非常庞大而复杂，管理数据是一项艰巨任务。员工的常见投诉之一是无法找到执行任务所需的数据。Airbnb 使用图数据库将数据生态系统联系起来，帮助员工轻松找到需要的数据，从而缓解了这个问题。



在尝试使用关系数据库解决方案后，UBS 转而使用图来评估数据血缘，并在数据血缘追溯图中描绘结果。这使他们能够更好地了解自己的数据。



基本指南

IDC 向企业推荐以下内容：

最大限度地提高将数据合成到商业价值的能力

在当今充斥丰富数据的世界中，数据越多并不表示企业智能越高。了解将数据转化为组织中的信息、知识和商业价值的核心途径。

确定图数据库和平台如何帮助您应对挑战

数据库在促进组织将数据转换为信息、知识和商业价值方面发挥着重要作用。因此，请仔细考虑不同类型数据库的优势和劣势，以及它们对数据价值创造起支持还是阻碍作用。了解图带来的优势，利用这些优势来提高组织的学习能力。

为您的生态系统选择合适的图数据库合作伙伴

利用图不仅仅需要图数据库。寻找提供图分析平台的供应商，与您当前的生态系统无缝集成。评估他们的整体平台，确保他们的工具符合您当前和未来的分析和 AI 需求。





IDC China

DC中国（北京）：中国北京市
东城区北三环东路36号
环球贸易中心E座901室
邮编：100013
+86.10.5889.1666

idc.com

[@idc](https://twitter.com/idc)

关于 IDC

国际数据公司 (IDC) 是在信息技术、电信行业和消费科技领域, 全球领先的专业的市场调查、咨询服务及会展活动提供商。IDC 帮助 IT 专业人士、业务主管和投资机构制定以事实为基础的技术采购决策和业务发展战略。IDC 在全球拥有超过 1100 名分析师, 他们针对 110 多个国家的技术和行业发展机遇和趋势, 提供全球化、区域性和本地化的专业意见。在 IDC 超过 50 年的发展历史中, 众多企业客户借助 IDC 的战略分析实现了其关键业务目标。IDC 是 IDG 旗下子公司, IDG 是全球领先的媒体出版, 会展服务及研究咨询公司。

IDC Custom Solutions

本出版物由 IDC Custom Solutions 制作。除注明具体供应商发起人, 本文展示的观点、分析和研究结果均源自 IDC 独立进行和发布的更为详细的研究和分析。IDC 定制解决方案以多种格式提供IDC内容, 供众多公司发行。许可分发 IDC 内容并不表示对被许可人或其观点的认可。

版权所有 2021 IDC。未经许可, 不得复制。保留所有权利。

IDC 信息和数据的外部出版

凡是在广告、新闻发布稿或促销资料中使用IDC信息, 都必须事先获得相应 IDC 副总裁或国家区域经理的书面同意。此类申请均应附上所提议文件的草案。IDC 保留以任何理由拒绝批准外部使用的权利。Email: ap_permissions@idc.com