



2019 年  
**商业  
智能  
趋势**



## 目录

### 01 可解释型 AI 日益崛起

随着企业越来越依赖于人工智能和机器学习模型，如何确保它们值得信赖？

### 02 自然语言让数据更加人性化

NLP 系统的进步使所有人都能与数据自然交流。

### 03 结合上下文分析数据来为行动提供指导

商业智能平台已进化到能够结合操作上下文来分析数据。

### 04 数据协作增进社会福祉

公共组织与私营组织携手努力，推动“数据福祉”运动。

### 05 道德准则跟上数据发展的步伐

按照《一般数据保护条例》(GDPR) 等法律规定，领导者纷纷评估数据伦理实践的未来。



- 06 **数据管理功能融入现代商业智能平台**  
受管控的数据管理实现了数据与业务的对接。
- 07 **讲述数据故事成为企业的新语言**  
数据见解的发现和传达如今是一项团队工作。
- 08 **企业在分析采用方面更加明智**  
如果企业领导者更关注的是参与度而不是采用率，会发生什么？
- 09 **数据大众化让数据科学家的地位不断提高**  
数据科学家提升软技能来推动企业变革。
- 10 **加速进行的云数据迁移推进现代商业智能的采用**  
数据正以前所未有的速度向云端迁移，这促使企业重新思考其数据策略。

# 01

# 可解释型 AI 日益崛起

随着企业越来越依赖于人工智能和机器学习模型，如何能确保它们是值得信赖的？

人工智能 (AI) 的强大潜力表明，机器可以通过自动化决策来增强其仿人类的理解力。Tableau 市场情报总监 Josh Parenteau 以另一个视角解释了人工智能和机器学习将如何发挥作用：“帮助发掘那些之前未被发现的见解。”Gartner 研究表明：到 2020 年，“85% 的首席信息官将通过购买、构建和外包等方式试点人工智能项目。”但是，随着企业越来越依赖机器学习模型，人类如何才能确保这些建议是可信的呢？

许多机器学习应用程序目前还没有以“知其然之后，知其所以然”的方式理解决策和建议背后的算法和逻辑，因此，AI 项目的试点企业有理由对人工智能的广泛应用感到担忧。正如剑桥大学机器学习领域的高级研究员 Adrian Weller 在这篇文章中所述：“透明性通常被认为是现实世界有效部署智能系统（如机器学习）的关键”。这种观点的产生有多种原因——主要是为了确保模型按照设计宗旨运作，或者是为了与用户建立信任，以便他们根据预测做出决策时有十足的把握。

**85%**  
的首席信息官将通过购买、构建和外包等方式试点人工智能项目。

对透明度的需求促进了可解释型 AI 的崛起，这种实践要求理解透明的观测结果并将其转换为机器学习模型。当某个模型得出一项结论时，决策者们希望能够通过进一步的询问来了解它为什么作出此结论，可信度有多高，不同的参考数据会让结论发生什么变化。这其实与领导者在做出关键决策时询问人类专家的情景非常相似。正如 Tableau AI 产品经理 Richard Tibbetts 所说：“当 AI 和机器学习提供的答案无法得到解释时，决策者持有怀疑态度是很正常的。分析和 AI 应该对人类的专业知识和理解能力提供辅助，而不是完全取代。”

在很多组织中，尤其像金融服务和制药公司等关注风险的组织，业务领导者会要求数据科学团队使用更易于解释的模型，并提供关于模型如何构建的文档或审计跟踪。由于数据科学家的工作是向业务用户解释这些模型，因此他们日益依靠商业智能平台，以交互式方法来探索和验证结论。

最终，企业接受了人工智能和机器学习的价值。但要在企业中产生颠覆性影响，人工智能必须得到信任。它必须以尽量简明的方式论证其结论，并能动态回答上述后续问题，一切都是为了让人们能更好地理解他们的数据。

“

**分析和 AI 应该对人类的专业知识和理解能力  
提供辅助，而不是完全取代。**

Tableau AI 产品经理 Richard Tibbetts

# 02

## 自然语言 让数据更加人性化

NLP 系统的进步使所有人都能与数据自然交流。

自然语言处理 (NLP) 将计算机科学和语言学结合起来，帮助计算机理解人类语言背后的含义。现今，商业智能供应商提供了一种可视化的自然语言界面，用户不需要深入了解商业智能工具，就可以自然地与他们的数据进行交互，像平常一样思考并提出问题。

在现代商业智能的背景下，自然语言被用来促进分析会话。所谓分析会话，是指人们围绕其数据与系统所进行的对话。系统根据对话的上下文，理解用户的查询意图并深化对话，从而创造更自然的对话体验。例如，

当对数据产生后续问题时，人们不需要重复完整的问题来进一步询问或澄清一个模棱两可的问题。您可以让商业智能工具“找出加利福尼亚州附近发生的大型地震”，然后输入一个后续问题“德克萨斯州附近呢？”，这一次没有提到地震这个词。机器学习使系统能够根据企业的数据和用户提出的问题类型，逐渐获取更深入的领域知识。

“分析会话的关键特征之一是避免陷入对话僵局，即能够提出问题，得到结果，且不偏离最初的问题。”Tableau 自然语言处理团队研

“

**通过自然语言，可将所有类型的技术带给更广范围的受众。自然语言减少了技术壁垒，让您无需学习软件，甚至不必学习如何进行分析，您只要拥有相关业务上下文即可提出恰当的问题。**

Tableau 高级产品营销总监 Stephanie Richardson



发经理 Vidya Setlur 解释道，“每个人都渴望能深入数据并得到见解。自然语言则是实现这种对接互动的一种重要方式。”

自然语言还可让用户基于数据可视化来提出问题：Tableau 软件工程师 Ryan Atallah 表示：“假设我用商业智能工具问了一个有关疾病爆发的问题，并得到可视化结果。那么我可以问‘橙色尖峰表示什么？’”这是一个后续问题，但它与我的数据无关，而是在询问数据可视化的图例。”并且，当目前的可视化在后续问题的上下文中没有意义时，它将提供一个替代方案。

自然语言改变了人们对数据提问的方式。当用户可以像人际交流那样与可视化进行交互时，数据分析便为普通用户敞开了大门，而不再是传统上数据科学家和高级分析师的专属领域。用户不受分析技巧的限制，而只受

限于自己问题的广度。它还允许高级用户在更短的时间内回答更丰富的问题，并为其他人提供更具吸引力的仪表板功能。随着自然语言在整个商业智能行业的成熟，它将打破各个组织采用分析的障碍，进一步将数据植入工作场所文化的核心。

据估计，到 2023 年，  
自然语言生产市场的  
规模将增长到

**8.253 亿美元**

(MarketsandMarkets)

# 03

## 结合上下文分析数据 来为行动提供指导

商业智能平台已进化到能够结合操作上下文来分析数据。

数据工作者希望并需要让他们的数据和操作位于同一位置。任何数据工作者都应该能够留在业务流程和工作流的上下文中，而不是在一个位置执行分析，然后在另一个位置采取操作。商业智能平台利用移动分析、嵌入式分析、仪表板扩展程序（也称为外接程序）和 API 等功能将核心业务操作、工作流和流程合并，从而满足这一需求。因此，具有指导意义的分析将加快技术和非技术角色的决策过程。

这些功能可让数据工作者分析数据，并在发现见解后采取行动，而无需转换环境。嵌入式分

析是聚合见解和行动的一个例子。嵌入式分析将数据和见解放在人们当前的工作环境中，这样他们就不必前往另一个应用程序或共享服务器。分析可嵌入到内部门户（如 SharePoint）或其他常用的应用程序中。

例如，企业可将分析嵌入到客户关系管理（CRM）软件中，比如 Salesforce。销售人员可以在上下文中查看有价值的客户数据，比如一段时间的产品偏好或支出，这可能有助于与客户更好地沟通，或改善要对客户采取的后续步骤。

“

**为了让分析结果具备可行性，我们需要确保在恰当的时间以易于理解的方式向相关人员提供正确的信息。**

Automated Insights 战略联盟负责人 Peter Benson

仪表盘扩展从另一个方面实现了这种融合。对于把大部分时间用在分析平台上的人来说，扩展程序可以将需要访问的其他系统直接引入当前仪表盘。现今，用户可以在不离开其分析工作流的情况下执行操作。

例如，依靠分析来监视工单队列情况的 IT 经理可以使用仪表盘扩展程序来编辑案例信息或直接在仪表盘内执行操作，而无需切换到工单系统。这使得他们能够保持工作的连贯性，在已经花时间熟悉的界面上实现更快的操作。

嵌入式分析和仪表盘扩展程序将操作和见解一同引入平台和工具中，而移动分析则将这些功能带给用户，无论他们身在何处。顾问可以在现场利用客户数据，而维修技师可以利用物联网 (IoT) 数据实地修理设备。

将分析放在上下文中的行为具有更大的影响，因为它有助于为特定的业务或行业定制分析。分析和操作的融合将缩短从见解到决策需耗费的时间和精力。这将使数据在业务工作流的应用更加广泛，同时鼓励更多的人将数据纳入日常决策。

到 2022 年，**50%**  
的数字业务技术平台  
项目将能够在相关活  
动与业务结果之间建  
立联系。(Gartner)

# 04

## 数据协作增进社会福祉

公共组织与私营组织携手努力，推动“数据福祉”运动。

数据改变了组织的运作方式，包括非政府组织 (NGO) 和非营利组织。现今，“数据福祉”运动在私营组织和公共组织中都呈爆炸式增长。Gartner 研究表明，“社交媒体去年对数据福祉的提及率增加了 68%”，原因是公众进一步认识到数据如何对社会产生积极影响。

诸如跨国电信公司 Orange 等私营公司正在设立项目，利用数据驱动型见解来进一步促进

社会公益事业。通过 Orange 公司的 OPAL 项目，该公司与当地政府合作成立了一个治理委员会，以规范数据的收集、匿名化和保护方式。这让他们能够以安全的方式与社会福祉组织共享详细的通话记录。在塞内加尔，这些数据最近被用来基于短信使用情况评估识字率，帮助社会福祉组织围绕扫盲项目做出资源配置决策。

从历史上看，非政府组织和非营利组织没有足够的资源投资建立复杂的数据基础结构或雇用庞大的数据工作者团队。现在，有了云计算的成本效益和灵活性，这些组织无需大量的本地投资即可开发复杂的数据环境，从而为更多数据驱动、具有社会影响力的方案的实施铺平道路。

云计算平台用于组织间的共享与协作，以实现共同目标，这是数据联邦产生和增长的一个实践范例。例如，Hutch Data Commonwealth<sup>1</sup> 是一个多学科的团队，其任务是帮助福瑞德哈金森肿瘤研究中心 (Fred Hutchinson Cancer

Research Center) 的研究人员通过创新的数据科学工具、基础设施功能和合作来加速他们的研究。”在一篇 GeekWire 文章中，Hutch Data Commonwealth 首席信息官兼执行董事 Matthew Trunnell 分享了云技术如何成为“科学数据共享的基础”。“云成为我们聚集与协作的地方，成为每个人为了共同利益而聚集在一起的地方。”数据联邦依赖与其他研究机构和技术供应商的合作关系，并以数据为其使命的核心。

无论是公共机构和私营企业的项目还是数据联邦计划，这些合作都需要以信任为基础。这些组织

正在评估成功伙伴关系的关键要素，包括涉及共享数据的法律影响和治理标准。这包括评估隐私风险，围绕共享个人身份信息的行为灌输保护意识。

纽约大学坦顿工程学院 (NYU Tandon School of Engineering) 的 Governance Lab 行动研究中心最近发布了一份报告，该报告主要关注私营企业和政府公共机构之间共享社交媒体数据所面临的挑战<sup>1</sup>，但其原理也适用于其他多种数据共享伙伴关系。GovLab 预计，更多的组织将任命数据管理员来推动数据协作，同时确保：设立正式的

“

**数据联邦让各企业能够在内外部安全共享数据，同时也能保护所有数据提供者的隐私。**

Tableau Foundation 的全球负责人 Neil Myrick

去年，社交媒体对“数据福祉”的提及率增加了  
**68%。**

流程来响应数据请求；采取适当的系统对特定类型的信息进行过滤和排序；应用适当的方法来保证发布的数据符合公众需求。

像福瑞德哈金森肿瘤研究中心那样，在正确的管控下访问大量不同的数据源，可以产生颠覆性的影响。尽管在这些大规模的合作项目中挑战仍然存在，但“数据福祉”运动已经证明了数据共享能够带来互惠利他的结果。技术的进步、数据素养的提高以及对协作的关注，为解决一些全球最困难的问题创造了有利环境。

<sup>1</sup><http://datacollaboratives.org/static/files/social-media-data.pdf>

# 05

## 道德准则跟上 数据发展的步伐

按照《一般数据保护条例》(GDPR)等法律规定，领导者纷纷评估数据伦理实践的未来。

数据隐私成为越来越突出的话题，消费者对共享个人数据也比以往任何时候都更加谨慎。这影响着企业进行数据创收、数据收集和数据共享的方式。随着 GDPR 等新规的出台，企业在日常业务实践中围绕数据道德和数据隐私展开了重要对话。这体现在以下方面：

**道德准则：**许多社会角色已经被职业道德规范所约束，包括法律行业、医疗行业和会计行业。随着每个业务领域的数据持续激增，企业开始评估如何将相同的原則应用到数据分析实践中。正如 [Gartner](#) 在一份报告中写道，

“数字商务时代模糊了技术和业务之间的界限”，数据现在已是战略拼图中不可或缺的一块。如今，越来越多的企业依靠数据在各个部门和角色中制订业务决策，这就意味着，更多的人参与到了数据的使用和共享当中。

为此，企业领导者，尤其是首席数据官 (CDO)，正领导制定全公司数据实践的內部指导方针，以此作为数字转型计划的一部分。事实上，根据 2017 年 [Gartner 首席数据官调查](#)，“从 2016 年到 2017 年，认为道德是自身责任的一部分的首席数据官数量增



认为道德是自身责任的一部分的首席数据官数量增加了  
**10**  
个百分点 (Gartner)

加了 10%。“这些道德准则将成为未来基础结构、管控和人员配置决策的框架。

**业务流程的变化：**企业正在对整个数据生命周期（从收集到分析）进行批判性思考。这为评估整个企业的数据管理策略提供了机会，有助于确保遵守法规和内部道德准则。这种评审过程不是一次性事件。正如埃森哲咨询公司 (Accenture) 在其通用数据道德准则报告中所述，“管控工作应当稳健可靠，为所有团队成员所熟知，并定期接受审查”，以适应企业的成长和变化。

数据道德不局限于数据收集或数据管控。它同样适用于数据的解析和依据数据采取的行动。现代商业智能平台为许多人打开了数据分析的大门，更多的角色将有责任遵守数据道德准则。Teknion Data Solutions 公司的高级顾问 Bridget Winds Cogley 建议，任何分析数据或交流见解的人士都应“考虑是否存在偏见，实际情况有没有清晰地呈现”，“这些数据的限制有没有被充分认识，能不能贴合当前问题”。随着越来越多的人成为数据工作者，数据道德将成为数据素养工作的核心部分，影响人们在个人和专业背景下处理数据的方式。

“

**道德实践帮助实践者退后一步，从道德视角对相关情况进行评估。最重要的是，数据道德可作为我们工作中的管控措施，让我们明白如何从个人及专业角度面对各种两难困境。**

Teknion Data Solutions 高级顾问 Bridget Winds Cogley

# 06

## 数据管理融入 现代商业智能平台

受管控的数据管理实现了数据与业务的对接。

随着数据源变得更加复杂、多样和众多，数据管理如今在现代商业智能部署中显得尤为重要。随着越来越多的工作人员使用数据来推动决策，企业必须确保其数据本身及其在分析中使用时的准确性。

企业已经转向数据管理，以应对更广泛的数据访问带来的数据管理和管控挑战。数据管理囊括了一个企业捕捉、清理、定义和排列

不相关数据的方式。此流程架起了数据与其实际应用之间的桥梁。

企业已经在将数据定义集成到分析工具的技术上花费重金，用来帮助分析数据，旨在消除团队和企业之间的模糊性。相应地，数据管理工具和流程(比如数据目录和语义管控)正在与商业智能平台融合，将数据与业务环境联系起来。

“

**在数据管理过程中，先要确定所需数据源，再将这些数据放在相关业务上下文中，让业务用户能够与数据进行交互，了解这些数据并使用它们来创建分析。**

Tableau 高级产品营销经理 Mike Hetrick

数据目录可充当数据源和公用数据定义的企业业务术语表。像数据工程师和数据管理员等主题专家可以为数据源和字段添加描述和定义，以作为标签来方便查找，甚至还可以添加有用的数据质量指标——包括对可信内容认证的通知，或者对数据资产的维护或折旧。

普通用户不需要知道数据在数据源中的具体位置，但是他们需要了解数据在现实世界中表示什么内容。例如，内容的分析师和使用者常常需要验证一段数据的来源（也称为谱系分析）。如果数据集发生变化，数据工程师和数据管理员需要分析下游影响，了解这如何影响他们管理的表或计划所关联的资产。将数据目录和商业智能平台结合有助于简化上述所有任务，提供使用指标来快速识别访问最频繁的数据源和仪表盘。

尽管数据目录非常必要，但在语义管控领域，可能存在比元数据管控更好的提升机会。语义不仅有助于连接数据的上下文，还有助于关联分析操作的意图，例如通过映射同义词来连接“订销量 (order size)”和“数量 (quantity)”等命令。这使得数据工作者能够以新模式与数据进行全方位交互，并迅速得出新见解。一种方法是通过自然语言交互，在此过程中，商业智能平台理解涉及多个查询的语义层，例如“突出显示最高值、最低值和平均值”。

随着这些技术和流程的不断融合，数据管理和语义将为余下的分析体验提供更坚实的基础。这将统一数据生态系统中更多不相干的组件（如清理和下游分析），并为表、联接和数据模型提供更强大的机器生成的建议。最终，数据管理的进步将使工作人员不仅可以在分析过程中对数据提出问题，还能转而对其业务提出问题。

到 2020 年，数字数据将以年复合增长率 (CAGR) 达 **42%** 的速度增长 (IDG)

# 07

# 讲述数据故事 成为企业的新语言

数据见解的发现和传达如今是一项团队工作。

无论我们的自动化程度有多高，无论我们的数据集有多大，也无论我们的运算有多机敏，如果不能将结果传达给其他人，所作的数据分析就没有意义。这时，数据可视化就派上了用场。数据可视化是一种语言，现今的分析师都应当知道如何以可操作且简单易懂的方式向决策者传达信息。这一技巧，结合分析师对其在数据中发现见解所采取的步骤的分享能力，常常被定义为“讲述数据故事”。

讲述数据故事是分析过程的一个关键要素。分析在变化多端的工作场所文化中占据主导地位，而这一不断变化的文化正在完善讲述数据故事的定义。随着越来越多的组织建立了分析文化，分析师讲述数据故事的方法更多的是围绕数据进行对话，而不是为单一的结论而争论。这些分析文化也促进了数据素养工作，旨在教导人们从发现见解的那一刻到最终达成业务决策的过程中真正理解数据，并参与分析会话。

“

**作为受众，我们需要做到自愿了解相关信息且能够解读数据，因此有必要掌握一定程度的专业知识。如果超出我们的理解范围，设计师就有责任向我们阐明这些数据的含义是什么。**

VisualisingData.com 网站创始人 Andy Kirk

根据 Dresner 2018 年的一份市场研究，**75%** 的受访者发现，讲述数据故事对于他们的商业智能计划至关重要或非常重要。

VisualisingData.com 网站的数据可视化专家兼创始人 Andy Kirk 分享了数据可视化的 7 个身份。其中之一是通信员，即“一位从根本上关心任何项目中涉及的所有人际关系的人（可能是主管、利益相关者和受众）”。Andy 解释说明了“所有的可视化工作，至少从传达信息的角度来说，必须以受众为中心。”数据工作者需要理解受众从可视化数据中形成结论的过程。与此同时，受众有责任掌握必要的专业知识来解读数据，并且必须“自愿了解相关信息”。

讲述数据故事的这种转变也体现在数据可视化趋势中。通过滚动或多页仪表盘展现的长篇讲述形式已经越来越常见，这种形式让分析师能够展现其逐步得出结论的过程。这些方法使分析师能够展示其分析进展，并强调他们在数据和假设结论中遇到过的见解。下一步则是围绕这些见解创建一个开放式对话。这为来自不同角色或部门的人留出了空间，以便引入额外的业务上下文，而且有助于在做出业务决策之前集思广益。

随着越来越多的组织创建 workflow 并建立专注于分析协作的团队，讲述数据故事将继续渗透到工作场所文化中。这种方法正在影响企业利用数据汇聚、传达和测试创意的方式。

同样，随着越来越多的人理解如何解读数据并解释他们的分析过程，还有助于提升业务的潜在影响力。



# 08

## 企业在分析采用方面更加明智

如果企业领导者更关注的是参与度而不是采用率，会发生什么？

商业智能计划通常会明确定义开始和结束日期，并且常常在向用户推出之后就被认为“已经完成”。但是，仅提供对商业智能解决方案的使用权不等于真正的采用。首席数据官们首先会重新评估采用商业智能在迈向现代化的战略转变中扮演怎样的角色，因为真实价值不是由您所部署的解决方案来衡量的，而是在于您的工作人员如何利用解决方案提升业务。

不要轻易断定每个人仅因为可以访问商业智能平台就能从中获得价值，这种假设实际上会阻碍通过分析来实现真正的进步。Tableau 市场情报总

监 Josh Parenteau 认为，根据许可证数量衡量投资报酬率，“无疑会忽略学习、成长和实现更大成功的可能性。”<sup>2</sup> 如今，企业领导者关注的是数据和分析是否正在改变整个企业的决策方式，而不是单纯地采用数据和分析。例如，如果把商业智能平台从员工那里拿走，会影响他们日常工作中作出决策的方式吗？

在您的手机上下载一个应用程序并不一定意味着您在使用它，同理，每月打开一次报告并不意味着它会推动任何行动或产生任何影响。企业领导人正在对鼓励参与的项目进行评估，比如内部社区和用户群体。这些以前被认为是草根项目的努力将被视为企业商业智能策略的基本要素，它们可帮助用户更快地上手、实现自助式服务，并迅速得到问题的答案。采用则是水到渠成，促使企业领导者增加投资，帮助扩展社区。

摩根大通集团 (JPMC) 旗下卓越中心团队在 IT 部门的带领下，利用这种模式部署数千名分析师，发展其用户社区。该集团举行全天会议来分享数据可视化和数据管理的最佳做法，副董事长兼商业智能创新领导 Steven Hittle 称之为“数据诊疗会”。这些会议只是众多活动的一部分，所有活动均旨在激发角色和部门之间的互动和对

话，以帮助 JPMC 将其现代商业智能平台扩展到三万多人。

随着这些内部社区将工作人员部署到商业智能平台上，企业可以开始委派分析职责并创建新的用户推动者。这将最终减少通常留给 IT 部门的繁重的维护和报告工作。更多的内部推动者将开始出现，他们扮演着主题专家的角色，不仅宣传最佳做法，还让大家在数据定义上保持一致。毫无疑问，所有这些运动都将使更多的人使用商业智能软件，并从中获得价值。最重要的是，您的工作人员会变得更高效，您的组织也会更有竞争力。

**60%**  
的首席信息官计划在  
未来 12 个月增加分析支出。

(IDG 首席信息官技术调查：  
2018 年技术优先要务 (CIO Tech  
Poll: Tech Priorities 2018))

<sup>2</sup> <https://www.tableau.com/zh-cn/about/blog/2017/10/three-reasons-your-business-intelligence-adoption-has-stalled-77448>

“

**我们需要重新思考衡量商业智能优点的方式。这不仅关系到谁拥有访问权限，而且还关乎人们如何利用分析结果来帮助制定决策。这才是采用。**

Tableau 市场情报总监 Josh Parenteau

# 09

## 数据大众化让数据科学家的地位不断提高

数据科学家提升软技能来推动企业变革。

如今，数据科学家这一职业炙手可热。

LinkedIn 发布的《2017 年美国新兴工作岗位报告》中指出，“自 2012 年以来，数据科学家的职位增长了 650% 以上”且“数百家企业都在招聘这些职位”。随着“机器学习工程师、数据科学家和大数据工程师成为最热门的新兴职业”，此类候选人才越来越多。

但是，随着更多的部门和职位都需要与数据打交道，企业人员的数据素养也有了全面提高，也出现了更多的公民数据科学家。根据 Gartner 的定义，公民数据科学家“能够创建

或生成采用高级诊断分析或预测和规范功能的模型，但本职工作并不属于统计和分析领域”。这些人不会取代数据科学家，但他们正在成为提出并测试假说的关键合作伙伴。

这改变了数据科学的定义，也模糊了传统数据专业知识和业务领域知识之间的界限。嘉信理财集团 (Charles Schwab) 统计分析经理 Sonic Prabhudesai 分享了为何“越来越多的业务人员了解如何处理数据，而数据科学家则越来越熟悉业务的内部工作方式”。

“

**如今，要成为一名数据科学家，就必须熟知统计建模和机器学习。区别在于，在这一领域工作的人是否能够以简单可行的方式传达他们的发现。**

嘉信理财集团 (Charles Schwab) 统计分析经理 Sonic Prabhudesai

自 2012 年以来，  
数据科学家的职位  
增长了  
**650%**

(LinkedIn)

如今，数据科学家需要掌握先进的统计和机器学习知识，同时他们也需要具备业务方面的战略头脑，包括对行业的深入了解。“数据科学的工作不仅仅是数字运算：其职能是应用各种技能，解决一个行业的特定问题。”印孚瑟斯 (Infosys) 首席全球数据科学家 N. R. Srinivasa Raghavan 博士解释道，“数据科学家需要深入了解将要应用其见解的领域。”

对于算法和模型的输出，只有使用它们在正确的上下文中帮助解决正确的问题时，它们才有效用。在流程的开始阶段，这意味着与利益相关者共同确定并改进问题陈述和假设，并使他们持续参与整个工作流。在工作流的最后，这意味着知道如何以相关且可操作的方式将结果传达给业务合作伙伴。

“如今，要成为一名数据科学家，就必须熟知统计建模和机器学习。”Sonic 表示，“区别在于，在这一领域工作的人是否能够以简单可行的方式传达他们的发现。”数据科学家将在如何将结果应用于业务方面发挥核心作用，而不只是交付结果。

利用自助式分析工具，数据科学家和高级用户都可以探索并更好地理解他们的数据。

这激发了能够指导余下分析和最终分析的见解，从而提升业务。

<sup>3</sup> <https://medium.com/@sonicmsba/how-to-build-an-effective-business-context-for-data-analytical-problems-cb02906341cd>

# 10

## 加速进行的云数据迁移 推进现代商业智能的采用

数据正以前所未有的速度向云端迁移，这促使企业重新思考其数据策略。

革新数据策略通常意味着要重新考虑数据的存储位置。越来越多的企业看到了将数据迁移至云端的优势，包括以较低的总拥有成本提高灵活性和可扩展性。事实上，Gartner 研究表明，“预计到 2018 年，公有云服务市场将增长 21.4%，总额达到 1864 亿美元。”<sup>4</sup>云服务让企业更容易捕捉和集成不同类型的数据。无论是全面云方案还是混合解决方案，这都意味着将所有储存数据从一个高度结构

化的本地仓库环境转移到一个扩展性更好、灵活更高的基础结构中。

这为我们带来了“数据引力”的概念，这个概念表明服务和应用程序受到数据所在位置的吸引<sup>5</sup>。随着越来越多的企业加速将工作负载转移到云端，这种数据引力也将分析过程拉到云上。“随着企业纷纷迁移到 Google Cloud，我们发现这些企业的领导者开始反思他们的

<sup>4</sup> <https://www.gartner.com/newsroom/id/3871416>

整个数据分析策略，并重新审视云平台在改善业务和增加收入方面的作用。”Google Cloud 产品管理总监 Sudhir Hasbe 解释道。

这种引力转移背后的驱动因素是延迟和吞吐量：延迟是执行一个动作所需的时间量，吞吐量是每个给定的单位时间内执行某一操作或实现某一结果的次数。当数据、应用程序和服务紧密结合在一起时，延迟和吞吐量就会降低，从而提高效率。当然，当数据储存在云端时，这些应用程序和服务将开始跟进。

当企业评估他们更为广泛的数据策略时，他们也会重新考虑其分析模型，从传统商业智能平台转向现代商业智能平台。McKinsey 指出，当企业接触到云基础结构和系统时，云的价值自然显现出来，“这不是一次性的战术决策，而是整体战略的一部分，用以追求数字转型”。<sup>6</sup>

传统的商业智能依赖于 IT 部门解决问题，这造成了解决方案的瓶颈，并导致分析与业务环境的分离。同样，为支持这种企业报告模式，传统商业智能的部署通常基于本地模型构建，很不灵活。

与此相反，云分析提供了多种优势，包括有机会考虑新的部署模型，而且企业领导者也渴望利用这些机会。

这包括向一线员工推出移动仪表盘，这样他们无需通过防火墙就能访问数据。云服务还支持与合作伙伴或客户共享安全仪表盘，在内部流程之外建立单一事实来源。

尽管并非所有企业都准备将全部数据转移到云端，但许多企业正在试验混合解决方案，以利用数据源多样化的优势。因此，在评估现代商业智能平台时，企业都会看它们是否支持未来过渡至全面的云分析。

预计到 2018 年，公共云服务市场将增长 21.4%，总额达 **1864 亿美元**

(Gartner)

<sup>5</sup> <https://www.techrepublic.com/article/how-data-gravity-both-hurts-and-helps-cloud-adoption/>

<sup>6</sup> <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/cloud-adoption-to-accelerate-it-modernization>



“

**我们发现这些企业的领导者开始反思他们的整个数据分析策略，并重新审视云平台在改善业务和增加收入方面的作用。**

Google Cloud 产品管理总监 Sudhir Hasbe

 + a b | e a u<sup>®</sup>