

# 数据驱动型企业

作者: Mark Schwartz, 企业战略师, AWS





## 简介

近来很多人都在讨论数据驱动型企业，以及是否有必要成为这样的企业。但究竟要怎样才能以数据为导向，而且为什么在当今数字环境中这样做如此重要？企业可以采取哪些实际步骤，才能以数据作为其思维和做法的基础？以及数据和数据时代其它优先事项（如业务和技术敏捷性）之间，又存在怎样的联系呢？本电子书将介绍什么是以数据为导向，并举例说明公司如何利用数据来推动其业务发展。我们还会将以数据为导向与敏捷性、数字化转型及持续创新串联起来。

数据驱动型组织致力于将其战略业务决策建立在由数据得出的证据之上，这不仅对严谨性有一定要求，与此同时还需要有能力基于数据本身发现机会，从而催生新的产品或市场。他们还会将数据视为可为自己所用的资产，以优化与客户之间的互动及提高效率。换言之，他们会分析数据，形成决定并且运用数据来为其客户提供服务。数据可以是决策的基础，例如在个性化、动态定价、市场拓展、产品创新或供应链优化等方面。

但直到最近，不少企业发现很难以这样的方式运用数据，因为他们仅在交易的背景环境中孤立地思考数据，随之而来的是，他们将数据锁定在孤立的数据库内，那么在这样环境里对交易处理非常有利，但可能不太适用于开放式分析。我们曾以发货单或订货单的思维模式思考问题：“请以每个 100 美元的价格给我来 20 个小部件”，或者，“请支付 20 个小部件的费用，每个价格是 100 美元”。数据描述行为，而且势在必行，数据也是交易处理的刺激因素或构件。今天，数据的价值已远远超出其在交易中所扮演的角色。

我们应该如何在金融领域考虑数据的价值，并且将价值最大化？

## 数据的业务价值

每条数据都可用于各种分析，促进提升业务效益，然后在从此类分析所得到的结果中凸显其价值。例如，如果企业通过分析其历史交易，从中发现优化其供应链的方式并进而降低成本，那么数据就能在实现成本缩减的过程中扮演重要的角色。因此，数据的业务价值来源于其可用于增加利润或达成使命目标的可能性。

正因为数据的价值超乎交易价值，我们很容易就能举出数据因存在超乎交易价值的额外价值，而被使用的实例。以强生公司 (Johnson & Johnson) 为例，他们使用存储在云中的交易数据，来改善其医学合规性、优化其供应链，并研发新药物。耐克公司 (Nike) 依靠收集客户成就数据，来提升用户在使用 NikePlus 时的数字体验。美国打车公司 Lyft 会收集并保存其所有行程的 GPS 坐标；他们通过分析发现，90% 的行程与来自附近地点的其它行程重叠。这一发现促使他们推出 Lyft Line，该服务允许乘客拼车并享受高达 5 折的优惠。<sup>1</sup>

由于此类使用情况未来可为企业实创造利润（即使目前尚待实现），因此我们可以将数据视作金融资产（虽然在大多数情况下，数据属于非 GAAP 资产）。因此，一家公司所累积的数据可以构成其收购价值中的一项重要因素，或许有助于该公司与其他企业缔结伙伴关系，这些一点都不意外。例如，微软 (Microsoft) 以 262 亿美元的高价收购拥有 4.33 亿用户数据的领英 (LinkedIn)，又或者 Caesars Entertainment Operating Corp. Inc. 在 2015 至 2017 年期间的破产官司中，其债权人坚称该公司在 Total Rewards 客户忠诚度计划中所持有的 4500 万客户的数据价值 10 亿美元，是该公司最珍贵的资产。<sup>2</sup>

将数据当做某种有业务价值的金融看涨期权这一做法很有帮助，因为这样能让我们有机会（但这并非必须）对供应链做出相应调整，或发布新产品。根据数据所显示的新业务的价值多少，我们可以选择是否行使该期权。评估数据资产价值的难点也出在这里：相对于计算预计现金流的投资回报率，评估看涨期权显得更加复杂。所以，企业常常忽略其价值，但正如我在《战争与和平与 IT》<sup>3</sup>一书中所介绍的那样，很多敏捷 IT 交付技巧都能产生此类期权价值。

<sup>1</sup>AWS 案例研究。请参阅 <https://www.youtube.com/watch?v=6AItOfqggek>、<https://aws.amazon.com/products/databases/> 和 <https://aws.amazon.com/solutions/case-studies/lyft/>

<sup>2</sup>两个示例都来自 <https://sloanreview.mit.edu/article/whats-your-data-worth/> 有关 Caesars 破产的详细分析，请参阅 <https://turnaround.org/sites/default/files/11.%20Paper%20-Caesars.pdf>。破产是极其复杂的，Total Rewards 的价值包含在其他资产中，因此最终附加的价值并不明确。

<sup>3</sup>Mark Schwartz, 《战争与和平与 IT: 数字时代的商业领导力、技术和成功 (War and Peace and IT: Business Leadership, Technology, and Success in the Digital Age)》(俄勒冈州波特兰: IT Revolution Press, 2019 年)。



## 数据与敏捷性

价值的产生不仅源于数据本身，还来自于我们用于分析数据并最终达成此类业务成果的工具和流程。当今的数字世界充满了快速的变化、不确定性和复杂性，甚至是所谓的颠覆性，我们需要利用数据来保证业务的敏捷性，并有充分的灵活性，来迅速针对不断变化的环境做出反应。敏捷性使组织能够将快速变化转变为机遇，并且通过对竞争威胁做出快速反应，避免受到阻碍。数字时代的企业认识到，他们需要将早期版本的产品迅速投入市场，并通过从市场持续获得反馈来对其加以改进。<sup>4</sup>

在最近几年，用于构建敏捷性的技术被产品开发流程所采用，例如，敏捷软件开发、开发运营和精益软件开发等。人们将云用于软件和硬件，以便加快 IT 功能的交付。基于团队的组织架构让我们能够调动资源，满足不断变化的需求。所有这些发展都帮助企业，使其流程变得更具敏捷性，

但敏捷的流程只是其中的一部分：公司的数据本身也必须具有敏捷性。公司的数据必须可以轻松应用于计划之外且不断变化的情况中。公司的数据必须触手可及且据实际使用价值。员工必须能够轻松获取相关操作数据的工具，并具有这样做的技能。这种灵活使用工具（包括我们无法预知的新用途）的能力，正是我们实现企业敏捷性所缺失的环节，扣上这一环节后，敏捷企业就能从仅单纯采用敏捷模型的框架及形式的组织区分开来，脱颖而出。数据敏捷性对于业务敏捷性来说必不可少。数据驱动型企业在这两个方面都十分精通。

提升数据的敏捷性是一个全新的概念。如果只是交易数据，我们可以将其局限在高度结构化的数据库当中，而数据库的架构反映了其将被用于此类交易的方式。关系数据库系统（如 Oracle 或 SQL Server 等）是我们的工具，这些工具在交易处理中体现出明显的优势。我们使用数据开展交易，并制作运营报告为这些交易提供支持。

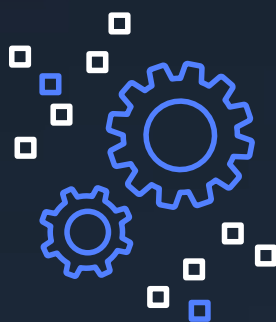
在重视隐私方面，我们通过严格限制，对数据强行实施严格的访问权限，而不是想办法将数据的使用局限于隐私保护的界限以内。我们采用类似于“通过模糊化来保护隐私”的措施，而不是“通过设计来着手保护隐私”。

是的，我们也尝试使用所谓的商业智能 (BI) 系统释放数据以用于临时分析。但目前工具的先进程度已远超 BI 系统所能处理的能力：我们现在有机器学习、一系列用于处理不同类型数据的专用数据库、进行大规模并行处理的算法、大量非结构化数据（如视频和语音）、传输传感器数据流的 IoT 设备，以及……当然，还有海量的数据。借助于这些工具，我们可以将数据从其交易和运营背景中解放出来。

更重要的是，我们已经意识到，实现以数据为导向不仅是一项技术挑战，还是一项组织挑战。**要成为一家数据驱动型组织，企业必须以独特的方式思考如何做出业务决定，以及如何与客户互动。这是企业对数据价值的承诺，也是组织的一种谦卑的态度：“数据比我们更了解情况。”**

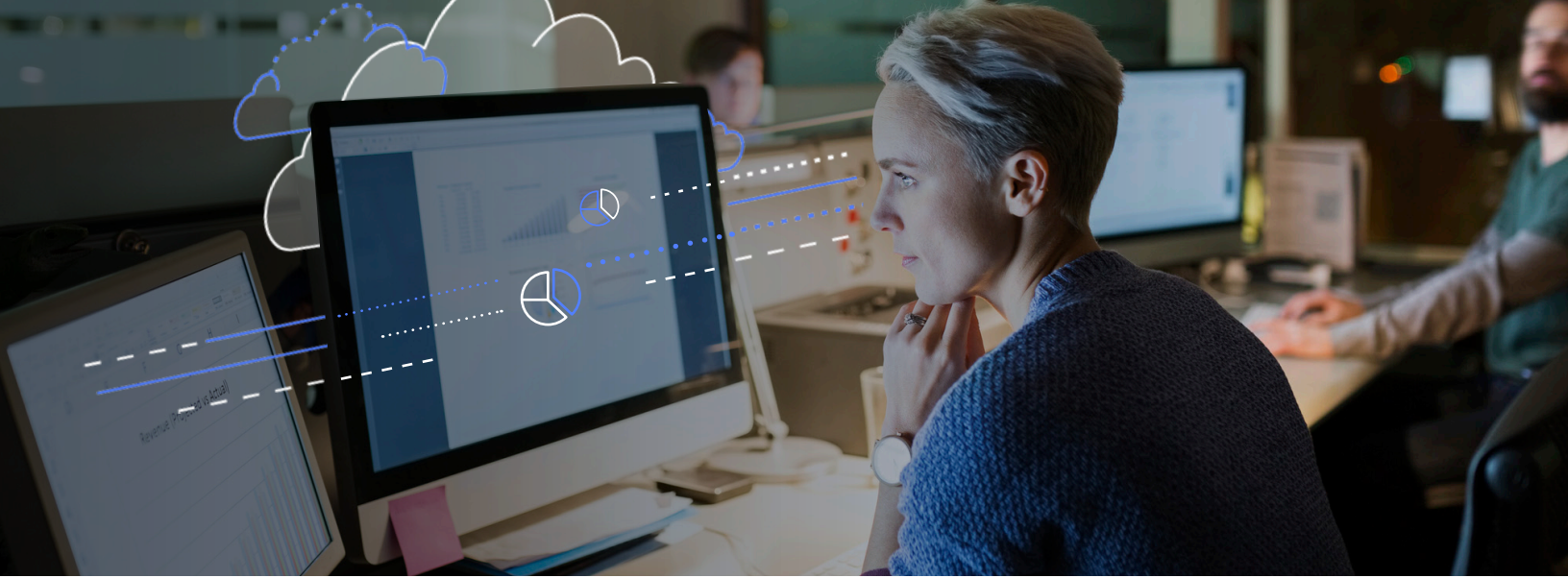
我们要如何做到以意想不到的方式使用我们的数据，也就是说，我们要如何灵活地使用数据，以达到业务敏捷性？我们要如何利用数据提高业务决策的严谨性和创造性？我们要如何改变企业文化，以充分利用此项新的灵活性所带来的优势？

此外，我们还要对数据采取哪些适当的控制措施，始终保护数据的隐私，且同时又能灵活、快速地使用这些数据？



归结起来其实就两个问题：

- 1.如何提高数据的敏捷性？
- 2.如何使用数据提高业务的敏捷性？



## 数据敏捷性

### 如何提高数据的敏捷性？

要实现业务敏捷性，我们需要随时对业务和竞争环境中的意外变化做出反应，并且进行真正意义上的创新，因此，我们将需要以在收集数据时可能无法预知的方式，将数据运用到工作当中。

### 我们的挑战：

- 我们的数据可能被封隔在事务性的关系数据库中，而且采用孤立的存储方式，导致组织不同部分的人员无法获取。
- 我们可能没有适当的分析工具，或者适当的人员在适当的时候可能无法获得这些数据。
- 我们的安全和隐私模型是临时的，因为我们或许从未考虑过使用数据进行探索。最有可能的情况是，我们只是通过尽量加大数据获取难度来加强保护隐私。

### 我们的目标：

- 在满足隐私和机密性保护要求的前提下，最大限度提高数据可用性。
- 打通信息孤岛之间的阻隔，增强整个企业的透明度。
- 为员工提供适当的工具，借助分析技术最新进展，支持员工以计划之外的方式探索数据。
- 确保具备专业知识，从而能够严谨且具创造性地解读数据。

在 [Analytics without Limits: FINRA's Scalable and Secure Big Data Architecture](#) 这篇博文中，美国金融业监管局 (FINRA) 的首席信息安全官 John Brady 将这些目标简练地总结为“他希望降低好奇心的成本”。他指的是最广义的成本概念，其中包括利用数据进行推断所用的时间，以及将数据转化成可用数据的风险。FINRA 的业务在于仔细观察日均逾370 亿笔交易的金融市场，从中寻找欺诈行为的蛛丝马迹。由于没办法每次都提前了解欺诈行为的规律，他们必须依靠其分析师的专业洞察分析，才能锁定可疑的行为。他们的任务与好奇心息息相关：他们要求分析师充满好奇地对数据进行检查，从而筛查出欺诈行为的规律及诱因。FINRA 的 IT 团队的任务则是，降低分析师在好奇心的驱使下，投入精力对直觉进行探索的成本。

Brady 的想法被全组织及各个职务所采用。营销人员可以轻松通过探索数据发现意想不到的消费者购买活动规律吗？运营部门能否通过仔细分析解读数据，找到需要进行绩效优化的地方，或诊断运营流程中出现的问题？财务部门能否通过探索数据，构思出提升绩效的新方法，或者对数据进行细分解读，以助于机构高管尽快决策？IT 主管能否测试其关于‘如何严谨但又锐意创新地优化云支出’的这一假设？

好奇心是创新与改进的推动力量。敏捷的数据让员工能以跟得上思维的速度，自由地探索头脑里迸发出的各种想法、直觉、假设和猜想，并且以数据为基石，推广新的想法。

要提升数据敏捷性，企业需要解决如何获得以及获得哪些数据，如何保护这些数据，如何以及在什么条件下提供数据，还有他们拥有哪些工具和技能可以处理这些数据等问题。



# 1

## 获取数据



要迅速使用数据，首先我们必须拥有数据。考虑到会将其用于未知的用途，我们收集的数据必须远超出我们知道如何使用的。这也是“大数据”的意义所在。值得庆幸的是，在云中存储数据的成本低，而且在进一步降低。因此，我们可以运用业务流程生成大量数据，并将其用于分析。例如，物联网 (IoT) 应用通常包含传感器，将数据点流快速地传输到云，使企业可以立即加以分析或存储用于未来分析。企业现在还可以处理更多数据类型，如：视频、文本和语音等。企业有极大可能会以新颖有趣的方式，使用此类全部信息。

举例来说，GE Oil and Gas 在其输油管中使用他们称为“清管器”、类似于 MRI 的设备，来收集有助于找到管道基础设施潜在问题的信息，信息大小超过 750 TB。Hudl 收集了大约 10 PB 的视频和其它数据，供体育教练和运动员进行回顾检查。Peloton 从锻炼周期收集数据，进行分析，以便为客户提供见解。而 Airbnb 每天也在云中积累大约 50 GB 用于快速分析的数据，使用 Amazon Elastic MapReduce (EMR) 工具允许对大批量数据进行快速地并行分析。<sup>5</sup>



# 2

## 存储数据



在获取数据以后，我们必须存储，作为分析所用。传统上，我们会根据我们所预期的数据在交易时被使用方法，将数据以结构化格式进行存储。例如，在数据库中，我们可能有独立的“订购数量”字段和“单价”字段。我们会收集数据以填写这些字段，通过将其保存到数据库的适当空白位置，对其进行归档，因为我们知道将二者相乘便得出总价。通过将数据置于此类模式中，我们使其可应用于交易，但我们可能丢失很多对分析十分有用的数据。这就是关系数据库模型。

在过去几十年里，此类关系数据库的使用占统治地位，因为非常适合通过业内早已熟悉的方式（即“订单数量乘以单价”），高效处理大量传统交易数据。但如果您处理的是非交易数据，或操作海量互联网交易，又或管理无法被轻松存储到预设定“数据字段”的数据，那么现在有了针对量身定制的更佳替代方法。

例如，Amazon Timestream 是专门为管理时间序列数据（如由工业传感器在很长一段时间内生成，或通过长时间追踪市场活动而生成的数据）而设计的数据库；Amazon Quantum Ledger Database 适用于在区块链中使用的数据类型（必须属于可采用类似于密码学技术对其历史记录进行验证的数据）；以及被设计用于表示复杂联系与关系（如社交网络）的 Amazon Neptune。企业不再受限于将其强行置入到关系模型。

更好的是，（对于敏捷性来说）将用于未确定分析的数据，可被存储在被称为‘数据湖’的灵活存储库中，而且每条数据会采用收到时的原本格式被存储。数据湖的好处在于可用于对其进行分析的工具：这些工具让您可以混合异构信息、结构化与非结构化数据、来自于不同组织孤岛中的数据，以及其它大量数据。目前的工具可以应用机器学习算法和统计分析，而且这些工具也能够兼容自然语言文本、视频和语音。

换言之，数据库可以满足企业在掌握数据将被使用的所有方式前加以存储的需求。我们可以将不同业务孤立渠道的数据，大批量保存到数据湖，然后对其进行整体分析。我们可以快速设置将数据从新收购公司导入到数据湖的方式，从而清楚理解被收购公司的运营情况，而且我们还可以将所导入的数据与我们自己的数据进行整合。实现所有这一切的神奇之处在于：(1) 低成本存储，(2) 此类工具兼容松散结构化的异构数据，以及 (3) 可使用特定服务以高带宽、非同步的方式，将数据推送到数据湖（以原来的格式，将数据发送到数据仓库，数据会迅速到达数据仓库而无需等待，这有点像电子邮件）。

# 3

## 提供数据



提升数据敏捷性的下一步就是在有帮助的时间和场合提供数据。（注意，我说的不是有需要的时间和场合。这里要讨论的是敏捷性和创新。）目前经常被用到的模型是一种自助预配式。如果分析师充满好奇，就可以启用一系列工具和数字子集进行分析，而不用申请或等待其他人提供。由此得来的自由，让分析师可以抓住一连串的灵感或“思路”，而不是采用断断续续，且会破坏创造性（或者您也可以说增加好奇心成本）的方式进行操作。云在其间扮演重要的推动角色，因为云使用户可以调配并使用全新的工作环境，然后在不再需要的时候将之舍弃。同时，还简化了防护措施的部署，对隐私加以保护（见下方关于此方面的详细信息）

# 4

## 提供工具



员工通常可以通过自预置模型，从其所在的数据驱动型企业，轻松快速地获取适当的分析工具，正如上所述。公司可使用一系列的软件和服务：如果您想要对数据执行传统的结构化查询，例如，您可以基于数据湖中的数据设置数据仓库，或者您也可以预置工具，让您可以直接对数据湖执行旧式的 SQL 类型的查询。

但今天，可能性变得更多。例如，您可以使用建模工具对数据进行可视化，而且您还可以构建场景并确定其后果。当今的分析革命都和人工智能及机器学习有关，为我们能对数据展开的操作，铺开了一条创新道路：预测结果、找出异常、分类数据、分析情绪、发现规律、引导机器人……等等。

例如，Capital One 正在使用机器学习来侦测欺诈行为，而同时又能维持高水平的客户服务。T-Mobile 通过机器学习改善其客户服务，机器学习会预测哪篇文章对客户最有帮助，并快速将其提供给客服代表。Sky News 在报导英国王室婚礼时，利用 AWS 机器学习，为电视观众识别并找出人群里的名流。而 Formula 1 赛车、美国职业棒球大联盟以及美国全国橄榄球联赛，都使用机器学习，来优化运动赛事观众的经验。<sup>6</sup>

要应用机器学习，您要根据更早之前的数据集训练模型，然后将这个模型，应用到新的数据集，并对其进行观察。在 AWS 中，机器学习的一般方法分为三种：(1) 使用预训练的模型，如 Amazon Rekognition（已经过预训练，可识别图像中的对象），或 Amazon Lex（已经过训练，可理解自然语言中所表达的意图）；(2) 使用 Amazon SageMaker 训练与应用您自己的基于任何一种用于机器学习的常用算法的模型；或者 (3) 如果您的员工精通机器学习，使用您自己的算法和训练方法，直接操作针对机器学习进行过优化的 Amazon 基础设施。

借助于上述工具，企业能促使员工发挥创造力，并找出运用数据的新方式。

# 5

## 提升技能



从您的数据当中提取价值的下一个重要的因素，是确保您的员工具有适当的技能……以及好奇心。这也是为什么数据科学家在当今如此受欢迎的原因。确实，即使在统计方面技能或经验不足的人员，现在也能找到各种工具。但要切实做到严谨地利用大多数数据，相关人员必须充分理解如何根据数据进行正确的推理。

举个简单的例子，我们缺乏统计经验的人员有过度依赖平均值的倾向，但有些时候需要观察值的全面分布，才能获得重要的见解。记得在我担任 USCIS 的首席信息官时，发生过一件事：当时我们正在探索如何缩短处理特定类型应用的时间。我们开发控制面板以追踪处理的平均时长，但我们所做的每次更改对指标的影响似乎都微不足道。我们没有意识到的是，有少量的应用会引起国家安全或欺诈忧虑，而这些应用所需的处理时间要长得多，因此导致偏离平均值的现象。我们对这些时间毫无控制能力。虽然我们的改进办法已应用到大多数案例中，但由于平均值严重偏移，我们无法观察到应用后的影响力。当我们最终意识到问题所在并开始监测时，比方说，第 85 个百分位的完成时间，我们观察到，我们的变更对大多数案例产生巨大的影响。我们有数据、工具和访问权限……我们所缺少的只是正确推理的技能。

如果数据以（甚至无意地）误导性的方式呈现，以数据为导向的决策也可能站不住脚。在其著作《定量信息的可视化显示》中，Edward Tufte 介绍了数据是如何因不恰当呈现方式，而被扭曲或遮蔽。<sup>7</sup> 对于想严谨使用数据的企业来说，确保其具备正确的分析与呈现技能，以及正确数据的重要性，再次被凸显出来。

# 6

## 提供保护



在可以将数据用于新颖的用途之前，如满足好奇心，我们必须保护数据的隐私和机密。数据驱动型企业会“通过设计来着手保护隐私”，刻意地根据计划和预测的方式，来设置保护的机制。他们会确保其已经考虑到什么需要保护，以及设置自动化的做法，以便最终获得速度和灵活性。实际上，最近欧盟颁布的《通用数据保护条例》(GDPR) 也要求通过设计来保护隐私。

云可以提供众多设置自动化访问控制的工具，让您能够以十分精细的级别，设定员工对于应当拥有的数据访问权限。追踪数据的来源和有效性，对其进行加密或模糊化，以及基于逐字段或逐记录的原则，限制访问权限的方法有很多。换而言之，您可以指定员工可以访问哪个客户的数据，以及查看哪条与这些客户关联的数据。Amazon Macie 甚至可以通过机器学习，来确定数据湖中的哪些信息是个人可识别信息 (PII)，并追踪该类信息如何被使用。或者，您还可以选择仅在聚合级别管理数据，或对信息进行遮蔽或匿名处理。在提供如此灵活性的情况下，每家数据驱动型企业必须根据所保管数据的类型，负责任地做出关于隐私的决策。

在使用企业所持有的海量信息的过程中，还有很多其它挑战。其中一项挑战，即如何准确地关联来自不同 IT 系统的关于个人的数据，在像美国这样没有专门的全国身份证制度的国家或地区来说，尤其如此。数据也有可能是不准确的，不仅因为在数据录入时可能出错，而且收集数据的 IT 系统也有可能存在局限性。例如，有些 IT 系统只支持姓氏和名字，在遇到姓名中有更多组成部分的人员时，就有可能发生不准确的情况。<sup>8</sup>

无论如何，数据驱动型企业的目标是提供数据，实现严谨而准确的决策和持续的创新。所需要采取的行动包括收集与存储数据以灵活地在未来使用，提供数据及合适的工具，避免与其他使用人员发生冲突，通过设计确保隐私和机密的安全，培养员工做出有效推理，并且解决可能导致错误决策的数据卫生问题。这是提升数据敏捷性的意义所在。

# 如何使用数据提高业务的敏捷性？

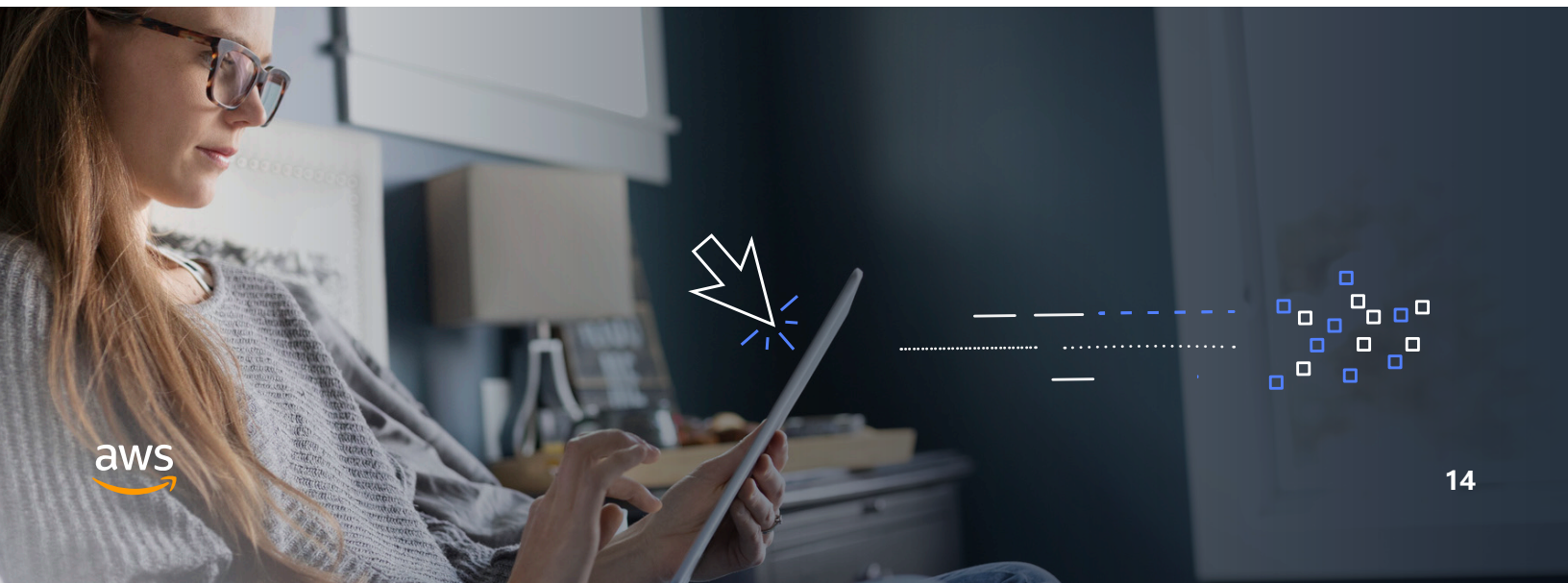
通过尝试新的想法，获得反馈，调整方向，然后重复这些做法，数字时代的敏捷业务得以维系。这个快速反馈的方式，让公司能够在使用前，测试想法、（低风险、高速且低成本地）进行创新并降低投资风险。这种方法使公司的产品与其服务的市场高度契合，确保该公司在适当的时间，以适当的方式，解决适当的问题。

## 快速反馈

在这种意义上来说，反馈并非指问客户他们是否喜欢新功能或产品。更常见的情况是，数据驱动型企业会使用定量反馈（此类反馈的收集方式为观察客户的实际行为），或监控市场行为或其它指标的变化。

例如，公司经常通过 A/B 测试来改善其网站的使用体验；亦即，尝试使用两种不同的设计方案（通常一种为当前使用的版本，另一种为他们考虑采用的新设计）。公司会向部分客户展示 A 版本，向另一部分客户展示 B 版本。他们会收集关于客户活动的数据，分析是否与自己关心的结果存在某种关系，以及是什么样的关系。如果他们要决定使用绿色或红色的按钮，以便使点击次数最大化，那么他们可以向有些用户展示绿色按钮，有些展示红色按钮，然后观察哪一个的点击量更高。Expedia 和 NetBix 这两家公司就定期开展 A/B 测试，利用大量来自云数据仓库的数据。<sup>9</sup>

<sup>9</sup> <https://www.youtube.com/watch?v=k8PTetgYzLA>.





通过反馈进行学习与调整的好处远超过单纯的 A/B 用户界面测试。例如，针对新的产品创意，可以通过打造“最小化可行产品”，然后进行测试，这是公司可以用于收集该产品是否能成功，或者如何改进才能成功等信息的最小型、最简单的产品版本。营销策略、促销和替代技术等，所有这些都可以通过试用和衡量进行测试，以减少不确定性。这样做的关键就是，收集数据，然后将其用于分析。

Eric Ries 的著作《精益创业》对使用最小化可行产品和快速反馈的技巧进行了描述。<sup>10</sup> 根据 Ries 的理论，一家初创企业会在特定的时候产生两种假设：关于他们的产品方案如何为客户创造价值的价值假设，以及关于该公司将如何才能拓展其市场，也就是吸引客户使用其产品的发展假设。最小化可行产品是能为初创企业确认或推翻此类假设，提供信息的最小型产品，企业可以借此进行更改，并再次投入市场测试。

这套实践方法不仅适用于初创企业或新产品开发，而且该方法已经成为组织（包括大型企业）基于所学到的经验教训改变方向，从而实现业务敏捷性这一过程中的核心。如果某企业考虑开发新的 IT 系统供其员工使用，他们会假设该 IT 系统将如何在其商业案例中实现提议的业务成果。他们应该对该假设进行测试，并根据显示的数据做出更改。

因此，敏捷的实践需要数据：要学习和适应，企业必须收集数据，了解其新举措的影响，并将其用于为此类举措提供充足的信息。另外，敏捷性还要求企业能够感知到其业务环境的变化，以便于做出适当反应，从而将业务成果最大化。数据驱动型企业不仅能提升其数据的敏捷性，还能利用数据为其自身的敏捷性提供支持。

## 文化和流程变更

从这个意义上来说，以数据为导向需要与众不同的决策方式；对于很多组织来说，这是一次非常深刻的文化转变。在过去，我们可能制定详细的计划，以可用的数据分析选项，然后仅根据可用的数据选择可能会产生最佳结果的选项，这就是我们的决策方式。在数字世界，我们拒绝接受制定计划时仅有的数据。我们会通过设计实验来获得更多数据，然后将那些数据用于我们的决策过程。我们通过生成新数据来消除不确定性。

其中一个例子就是我们在 USCIS 时设计的 IT 治理方法。我们并没有编写冗长的规格需求文件，然后交给技术人员去实施，我们只是提交了一项业务目标。例如在一个案例中，我们注意到一位娴熟的案件处理人员（“状态验证人员”）每天可以处理大约 70 个案件，但我们的业务目标是使其高于这个数字。在另一个业务案例中，我们发现有些纸质文件在不同处理地点之间传送期间丢失了，而我们想杜绝这样的问题。

针对这些目标中的每一项，我们首先会创建显示有关键指标的控制面板：每天的案件数量，或丢失的文件数量等。我们会创建一支由业务运营人员和 IT 技术人员组成的跨职能团队，然后要求他们优化指标，而不是编写一份需求文件。我们为他们提供工具，以便快速对 IT 系统和业务流程进行更改，然后在控制面板中加以监控。他们尝试小型的增量更改，并每天监控结果。依据所见的数据变化，他们可以确定下一步该怎么做，以便将结果最大化。而管理层则可以决定是否继续为相关举措提供资金，或者将资金用到其他地方。其结果是可快速创造价值的以数据为导向的低风险、轻量级治理流程。

这将引出另一个要点：通过透明度强化问责制。在大范围内提供数据，我们使团队进度保持透明。因此，监管机构可以时常反思其投资决定，增加或减少投资，重新定义目标，或者完全停止投资行为。结果是衡量成功的唯一标准，而且能很迅速地得到结果。但这些结果必须有数据作为支撑。



## 发现模式

数据可以提升敏捷性的另一个方面，就是感知环境中的变化并找到规律。例如，机器学习可用于侦测并对异常情况做出反应。我们会用历史或常规数据，对机器学习模型进行训练，使其习惯于什么是“正常”，然后将模型应用，才能找出不正常的活动。例如，此项技巧可被用来侦察出欺诈交易或黑客对网络的入侵。或者找出工厂生产线上行为异常的设备，并在其实际发生故障前，进行维修或更换。

当收集大量数据时，我们可能会发现过去没有意识到的事物之间的关系。社交媒体公司会构建大型的人员关系数据库。国土安全部可能会发现，他们正在调查的恐怖分子嫌疑人曾经与确认为恐怖分子的人住在相同的地址，这样一来，他们就可以在下次遇到该人员时，提出相关问题。大量欺骗性的移民申请可能被证明全部都出于同一个移民律师之手。在这种情况下，我们已不再局限于简单地使用数据来处理交易：我们现在可以发现此类交易之间那些又重要又存在具体意义的关系。但还是老问题，我们不知道究竟可能会发现怎样的关系；而且灵活性，以及好奇心是从数据获取价值的关键。

要引述关于利用数据“监控事件”的更多例子，数据点的存在可被用来证明发生过的活动，例如，当审计跟踪记录自动创建时。对活动进行跟踪，审计人员或许可以验证合规性或调查不当活动。区块链常被用于存储证明活动发生的数据，例如，双方之间的资金转移，或有关方批准相关合同。通过使用自动化保护措施和审计数据建立合规性，企业经常可以避免会降低敏捷性的重量级合规流程。

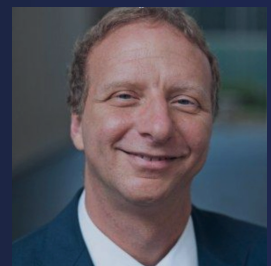
当然，在使用数据支持业务敏捷性的过程中，还存在很多挑战。正如我们在上文中谈到的那样，从数据中获得适当推理是一种必备技能。数据并不是总能告诉我们要采取什么行动：我们必须进行解读并做出正确的决定。通常，我们要在误中和误漏之间权衡，例如，如果使用数据发现异常交易以便确定潜在的欺诈行为，就可能需要承担跳出过多异常交易的风险，而这样又给客户制造了烦恼，或者标记出过少异常交易，导致犯罪分子得以蒙骗过关。数据集的规模越大，出现无意义模式或重要的模式被大量潜在联系所掩盖的可能性也变得越高。干扰会随着信号的增加而累积。

## 结语

数据驱动型组织会使用数据促成严谨的决策过程，以及提供可用的数据激励创新并为客户创造价值，从而在工作中运用数据来优化业务成果。若数据被锁在僵化的框架或孤岛中，或者难以取得，就会成为业务敏捷性的绊脚石，使公司无法对机遇做出反应或使产品无法快速被投放到市场。更糟糕的是，当业务无法通过使用数据促进其流程和投资，就会放弃与自己尝试服务的市场产生重要联系，或错过可能有助于其举措取得更大成功的反馈。而在另一方面，数据驱动型组织会通过数据获得敏捷性，并利用此敏捷性使数据变得更有价值。

### 作者简介

Mark Schwartz 是 Amazon Web Services 的战略师，也是《The Art of Business Value》、《A Seat at the Table: IT Leadership in the Age of Agility》和《War and Peace and IT: Business Leadership, Technology, and Success in the Digital Age》的作者。在加入 AWS 之前，他曾担任过美国公民及移民服务局（国土安全部的下属机构）的首席信息官、Intrax 的首席信息官，以及 Auctiva 的首席执行官。他拥有沃顿商学院的 MBA 学位，以及耶鲁大学的计算机科学理科学士和哲学文科硕士学位。



Mark Schwartz,  
企业战略师, AWS

[参阅关于 Mark Schwartz 的更多内容 »](#)