

O'REILLY®

Compliments of
IEEE®

AI之梯

解密AI的挑战

The AI Ladder

Rob Thomas

REPORT

AI之梯

解密AI的挑战

Rob Thomas 著

Tim O'Reilly 序

Beijing • Boston • Farnham • Sebastopol • Tokyo

O'REILLY®

The AI Ladder

by Rob Thomas

Copyright © 2019 O'Reilly Media. All rights reserved.

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (<http://oreilly.com>). For more information, contact our corporate/institutional sales department: 800-998-9938 or corporate@oreilly.com.

Acquisitions Editor: Rachel Roumeliotis

Development Editor: Nicole Tache

Production Editor: Kristen Brown

Copyeditor: Octal Publishing, LLC

Interior Designer: David Futato

Cover Designer: Karen Montgomery

Illustrator: Rebecca Demarest

August 2019: First Edition

Revision History for the First Edition

2019-08-23: First Release

The O'Reilly logo is a registered trademark of O'Reilly Media, Inc. *The AI Ladder*, the cover image, and related trade dress are trademarks of O'Reilly Media, Inc.

The views expressed in this work are those of the author, and do not represent the publisher's views. While the publisher and the author have used good faith efforts to ensure that the information and instructions contained in this work are accurate, the publisher and the author disclaim all responsibility for errors or omissions, including without limitation responsibility for damages resulting from the use of or reliance on this work. Use of the information and instructions contained in this work is at your own risk. If any code samples or other technology this work contains or describes is subject to open source licenses or the intellectual property rights of others, it is your responsibility to ensure that your use thereof complies with such licenses and/or rights.

This work is part of a collaboration between O'Reilly and IBM. See our [statement of editorial independence](#).

978-1-492-07310-9

[LSI]

目录

序	1
AI（人工智能）之梯	3
介绍	3
人工智能是新的电力	3
人工智能的简要定义	4
人工智能的挑战	5
人工智能之梯	7
现代化企业的基础架构：让数据为人工智能和多云 的世界做好准备	9
收集：让数据更简单且易于访问	10
整理：奠定业务就绪的分析基础	11
分析：以可信任和透明的形态构建和扩展人工智能	14
注入：让人工智能运用于实践	16
结论	18

序

多年以来，我意识到好的名字可以帮助大众更好地理解一个新的重要技术概念。在IBM Think大会期间一次喧嚣的鸡尾酒会上，当Rob Thomas第一次告诉我“AI之梯”这个想法的时候，我意识到这是一个类似“开源软件”、“Web 2.0”、“大数据”、“创客运动”那样的好概念。它们不只是一个标签，而是一张地图，可以帮助大众探索和理解之前被认为的“未知世界”。

现在每个人都在谈论“AI（人工智能）”，但是大部分的企业并不知道如何把人工智能应用于实践。他们仅知道需要一些看起来像魔术的东西。但是没有一家人工智能供应商可以像魔术师从帽子里面拿出一只兔子那样来帮助把人工智能给到他们。那些已经成功（采用人工智能）的企业都是沿着类似Rob的“人工智能之梯”那样的路径发展来的。首先，这意味着企业需要明白他们正在尽力解决什么业务问题。然后企业需要把数据准备好，这并不只是针对企业传统的业务和数据。人工智能不光和企业已有的数据相关，还可能需要创建和购买与业务问题相关的数据。这些数据可以被用来训练人工智能模型，从而让模型能够更好地理解和应对真实世界里的数据。为了构建现在的语音和图像识别能力所需要的数据是海量的。因此企业需要一个新的数据基础架构，能够支持不同的数据源。

企业还需要开发人员和业务人员都获得新的技能，同时也需要改变企业文化来适应这种新的工作方式。正如著名的人工智能教科书合著者

Peter Norvig在首届O'Reilly AI大会上所讲的：“人工智能开发者的工作流程和软件开发人员的工作流程迥然不同。”

企业也需要放弃对之前许多工作的假设。很常见的情况是，新技术的第一轮发展仅仅只是对于已有事物的稍加改进，例如所谓的“没有马的运输”。只有在采用了新技术之后，企业也才能意识到他们需要针对这些新的能力作出多少业务模式的改变。这也就是为什么每个新技术革命都会孕育出新的商业领袖而顽固派却常常落伍的原因。

这个收集、整理、分析并最终向企业注入人工智能的过程可以被比喻成一把梯子。通过把原本不可能一跃完成的事情变成一系列的台阶，梯子可以帮助人爬到他自己没法到达的高度。理解了梯子的含义可以帮助企业评估他们对人工智能的准备程度。正如人不能一跃到达屋顶，用一把残缺的梯子也没法帮人爬上去。这个人工智能之梯的比喻并不完美，因为与爬梯子不同，企业对于人工智能的准备并不是一个线性的过程。但是了解人工智能之梯的这些步骤可以帮助企业发现他们的优势和差距。

这看起来是个简单的想法。确实，它是很简单，就如同“开源”或“移动优先”。这些事后看起来都非常明显的想法，却是经过大量的努力和挫折之后才获得认同。虽然开始的时候没有获得认可，但倡导这些简单想法的企业成功了，而没有这样去做的企业则落后了。

本报告介绍了一个路线图。它能帮助那些没有多年人工智能研究积累和雇用几百人博士团队的企业受益于计算机领域的下一个重大进步。

前美国国防部高级研究计划局人工智能项目主管、现匹兹堡大学信息科学学院院长Paul Cohen说：“人工智能的机遇是帮助人类对复杂的交互系统进行建模并管理它。不要错过这个机遇。”

—— Tim O'Reilly, O'Reilly Media创始人和CEO

AI（人工智能）之梯

介绍

人工智能（AI）是我们这个时代最伟大的机遇之一。得益于海量的数据和机器学习领域前所未有的进展，人工智能拥有着到2030年为全球经济增加16万亿美元的潜力。人工智能迫使所有行业的企业都进行转型，并将改变人类的工作方式。根据Gartner的研究，到2021年人工智能增强技术（把人和人工智能结合）将会提升62亿小时的人工生产力。尽管这些预测令人鼓舞，但人工智能被采用的速度却慢于预期。据报道，81%的商业领袖并不理解人工智能需要的数据和基础设施。因此本报告旨在为企业高管和一线专业人士提供一套统一的、指导性的方法——人工智能之梯，来帮助企业成功地实施人工智能。

人工智能是新的电力

人工智能（AI）是我们这个时代最伟大的挑战和机遇之一。它有望改变人类的工作方式、企业的运营模式以及整个产业界的转型模式。人工智能的创新不仅仅只是节省费用，它们实际上可以帮助企业预测并塑造未来成果，它可以让人来从事更高价值的工作；它可以实现决策、流程、体验自动化；人工智能还可以帮助重新思考新业务模式。最终，这一切都意味着收入的增长。

然而，人工智能却经常被描述成某种神秘的事物。像一个魔术黑盒子，人们只能看到它变魔术，但对其内部的运作机制却知之甚少。大众贬低人工智能是专家们才能掌握的用来炫技的东西。在这样的环境下，人工智能充满了神秘主义的气息。它有着宏大的承诺，且超越了凡人的世界。人工智能是新的类似于电的发明。

当电首次被发现时，它被认为是巫术，是一种让观众无法了解它来自于哪里以及如何被生成的魔法。所有的新发明都经历了类似的演进：发现，探索，应用，并最终普遍存在。

现在我们发现自己处于类似的情形。我们已经理解了人工智能的力量，但是尚未完全发现如何释放它的潜能。现实是：人工智能不是魔术，应用它是艰难的工作。没有魔法棒可以被挥舞着来解决企业的低效率问题，而且仅凭技术本身是不够的。

为了更好地铺垫本报告的上下文，让我们先简要地定义什么是人工智能，并讨论在企业级别上实施人工智能时常见的挑战。

人工智能的简要定义

人工智能是一系列技术的通称。它们让机器从数据里进行学习并根据所学进而采取行动，而不是机械地根据程序员写的指令行事。机器学习是软件工程学的一个分支。它几乎总是一个更大的传统软件系统的一部分。如我在演讲里调侃的，如果是用Python写的东西，我们叫它“机器学习”，而如果使用PowerPoint写的东西，我们叫它“人工智能”。

人工智能是目前语音识别、图像识别和自动驾驶等方面进步的推动力。它让手机里和家里使用的语音唤醒助手成为可能，同时在用户关怀、社交媒体和网络安全等领域也扮演着重要的角色。

不过对于商业而言，人工智能可以被定义成从根本上改进三件事的方法：预测、自动化和优化。首先，人工智能和预测紧密相关。企业和组织希望能在宏观和微观两个层面来预测他们业务的发展。其次，有些业务流程很重要但非常耗时且常常需要人工完成，自动化它们具有极大的业务价值。这可以解放员工来专注于高价值、更具创造性的工作。最后，人工智能也和优化相关，无论是优化路径和物流，还是优化市场推

广的花费，或者是优化云环境安装的配置。人工智能是一种工具，可以更快且更大规模地改进人类的决策制定，同时具有增强每个员工工作能力的潜力。

是时候揭开人工智能挑战的神秘面纱了，了解这些挑战能帮助企业成功地把人工智能的能力引入他们的业务。

人工智能的挑战

目前的现实是每个行业、各种规模的企业都在艰难地引入人工智能技术。他们所面临的挑战可以归纳为五类。

首先是缺乏（对人工智能）的理解。很多企业仅仅是因为“人工智能解决方案”的日益流行而纷纷加入并寻求实施“人工智能解决方案”。他们假定这些方案可以解决任何的问题。从本质上讲，人工智能代表了一整套强大的软件和数据工程技术，它可以从大量的非结构化数据里获取洞察。但这并不意味着人工智能是可以解决任何问题的魔法杖，它必须被应用于它适合解决的问题。因此，企业首先要做的事情是理解他们试图去解决的业务问题，问出正确的问题，并确定人工智能是否是达成他们业务目标的合适方法。

企业面临的第二个挑战是如何处理好他们的数据。数据是人工智能的基石和燃料。训练机器学习模型需要好的数据，只有这样才能让融合人工智能的业务流程真正起作用。目前有三种不同类型的数据问题。

缺乏数据

企业需要开始着手收集自己的数据，并从第三方获取额外的数据，同时让整个企业都可以访问到这些数据。

数据太多

缺乏数据会阻碍人工智能的采用，同理，太多的数据也会。当企业有非常多的数据，并且分布存储在不同的环境和数据库里，这马上就变成了一个数据工程问题。在这种情况下，企业需要收集并整理好数据以供人工智能使用。

坏数据

“垃圾进垃圾出”的思想在人工智能时代依然有效，这和它最初被创造出来的计算机编程早期的情形是一样的。问题是，尽管商业领

袖们会把提升数据的使用列为高优先级的事情，他们中的60%会面临管理数据质量的挑战。和人工智能相关的大部分工作都涉及到数据的清洗。

企业面临的下一个挑战是缺乏相关的技能。人工智能要求即使是最有经验的软件工程师都要重新学习如何编程的知识。在现有的软件开发形式里，程序员编写出计算机需要执行的内容。但在基于人工智能的项目里，程序员把训练数据喂给机器学习算法，让算法自己从数据里去学习和构建出能代表需要完成任务的数学模型。当模型再看到真实世界里的数据时，它就能够识别出在训练数据里看到的同样的模式。然后，模型的输出与传统的流程化程序相结合来作用于它识别到的内容。训练过程是非常耗时的，常规的持续集成（CI）和持续发布（CD）的流程并不容易被应用到这个过程中去。

由于人工智能技能的稀缺，且对它的需求非常旺盛，所以真正的挑战是不容易雇用到有经验的人员。因此，无论企业员工的技能水平如何，非常重要的一项是让正在被构建和使用的技术对企业里的每个人都能可用。

下一个挑战是信任。确保人工智能的推荐或决策过程完全可追溯是非常重要的。这可以保证企业能审计模型和相关的训练数据的血缘关系，以及审计每个人工智能推荐的输入和输出。随着在更多的应用里使用人工智能，企业需要能够理解他们的人工智能应用作出的决策。对于类似于金融和医疗这样需要遵循欧盟《通用数据保护条例（GDPR）》和其他相关条例法案的特殊行业，广泛地采用人工智能面临着更高的门槛。为了能够在生产环境里使用人工智能应用，企业必须能够解释它们的产出结果。

当企业把人工智能模型投入生产环境后，重要的是要打开人工智能模型“黑盒子”，制定出策略来持续地监控模型上线后的产出。没有这个策略，企业就会对于人工智能应用的行为一无所知，不知道人工智能在做什么、多久会被使用一次、输出是什么，以及是否有源于模型的训练数据产生的歧视。作家Cathy O’Neil把模型可以被编码进歧视的这种状况称为“数学杀伤性武器”。

最后一个挑战是关于如何改变企业的文化和业务模式，以便更好地利用新技术带来的机遇。就如同很多企业没能拥抱互联网和移动革命一

样，他们通常不愿意深度思考他们的业务模式和商业流程如何才能全面拥抱人工智能带来的新机遇。人工智能不仅只是改进现有的业务流程，它让企业可以重新思考所有的流程，并做出之前不可能的事情。

人工智能让企业能从海量的数据里获取洞察。但这并不意味着让传统的商业分析师继续去钻研同类型的电子表格来做决策。它意味着企业有了可以可靠地完成的全自动的“工人”。由之前分析师转型的数据科学家需要成为这些“工人”的管理者，评估它们的工作。并在出现问题时进行干预，但不一定要推翻它们的决定，而是回头来改进训练它们的数据，从而让它们可以在下一次能成功完成任务。同样重要的是要能理解人工智能转型并不是完成单个的业务项目，而是要转变整个企业的商业文化——创建一个不断迭代、试验的新文化。

人工智能之梯

如前所述，人工智能不是魔法。企业如果希望运用人工智能的潜能，他们需要使用来自多源的数据，支持最好的工具和框架，并在多种环境里运用模型。然而有81%的商业领袖并不理解人工智能所需的数据和基础设施是什么。

绝大部分人工智能应用的失败都是源于数据准备和整理的失败，而不是模型本身。人工智能模型的成功需要企业先能成功地收集和整理数据。

人工智能算法纵然精妙，但它无法抵消数据[架构]匮乏所带来的问题……糟糕的数据就是残废。

——《MIT Sloan管理评论》

企业需要一套经过深思熟虑和精心设计的方法，尤其是在现在的混合云和多云环境里。他们需要一种现代的、开放的设计，能保证企业可以灵活统一地和任意云上的开源系统一起工作。使用统一、指导性和公开的信息基础架构，企业可以对其数据架构进行现代化改造，从而让数据为人工智能和多云环境做好准备。

简而言之，没有信息基础架构就没有人工智能。人工智能信息基础架构是企业整理和结构化数据的基石。它可以帮助消除数据孤岛，避免数据封闭，并允许任意地点的敏捷化运行。同时，随着为人工智能设计的信

息基础架构的成型，企业就能够使用一个统一的方法来自动化和管理它的数据和人工智能应用的生命周期，从而最终可以用可信任和透明的形态让人工智能应用于实践。

由IBM提出的人工智能之梯（见图1-1）可以帮助企业了解他们在人工智能旅程中所处的位置，以及作为一个框架来帮助他们决定需要关注的方向。它为企业提供了业务转型的规范原则，包括四个需要重点关注的领域：如何收集数据、整理数据、分析数据，并最终把人工智能注入进企业中去。

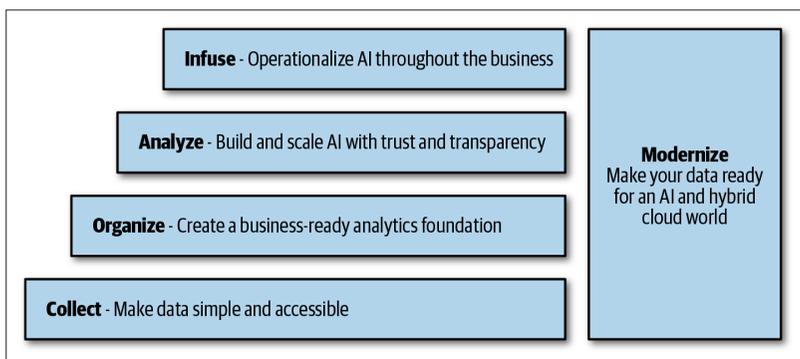


图1-1：人工智能之梯，一个企业通过连接数据和人工智能来实现业务转型的规范性战略

无论企业处于人工智能旅程的哪个阶段，把一个人工智能战略分解成小块（或梯子上的横档）可以作为一个规范性原则。无论数据存位于何处，通过统一数据收集、整理和分析，这个策略能让企业简化和自动化把数据转化成洞察的过程。通过使用人工智能之梯的规范性框架，企业可以构建一个高度受控、高效、敏捷和面向未来的人工智能方法的基石。

下面让我们看看构成人工智能之梯的四个主要步骤的细节。

1. 收集：让数据变得简单且易于访问

无论数据位于何处，把每种类型的数据都收集起来，并形成应对持续变化的数据源的灵活性。

2. 整理：奠定业务就绪的分析基础

把所有的数据整理进入一个可信任和业务就绪的基础里。其中内建了数据管理、保护和合规的功能。

3. 分析：以可信任和透明的形态构建和扩展人工智能

用智慧的方法来分析数据，充分利用人工智能模型来赋能企业获得新的洞察，并作出更好更智能的决策。

4. 注入：让人工智能应用于实践

借助于预测、自动化和优化，把人工智能应用到企业的多个部门和各种流程中。

不管企业现在处于什么样的阶段，人工智能之旅依然是很困难的工作。只有技术是不够的，这就是为什么人工智能之梯的每一步都很重要。下面的章节会对每一步都作出详细的讲解。不过在实际走出第一步并开始收集数据之前，企业需要对他们数据基础架构进行现代化改造。

现代化企业的基础架构：让数据为人工智能和多云的世界做好准备

在我们这个语境下，“现代化”意味着为人工智能构建一套信息基础架构，从而给整个企业和组织提供多种选择和灵活性。为了既能满足当下的需求，同时保证在未来具有竞争力，企业需要一个高效、敏捷的数据基础架构，能在上面可靠地构建应用。这会是一个混合云和多云的平台。

混合云和多云的平台是数据基础架构的未来。通过使用容器，混合云与多云平台可以让企业受益于在任何云环境（公有云、私有云、本地机器）里的数据和应用。随着企业为人工智能和多云世界进行架构改造，他们将会发现为了在整个组织内扩展人工智能的影响力而需要的“组装要求”会更少。

在大部分企业中，数据都是分散在不同的环境里，如公有云（包括多个供应商）、私有云和传统的本地机器部署。混合多云平台可以帮助解决很多企业面对的数据分散的挑战，即信息分布存储在多个孤岛、数据库

和云上。如果企业正在这些环境里部署和应用人工智能，那么他们就不得不去使用这些云供应商的人工智能工具。这就是当下所出现的“供应商锁定”的问题。它会抑制创新和阻碍企业想扩展人工智能的努力。混合型平台可以提供必要的基础，支持企业大规模构建、部署和管理人工智能模型所需的所有能力。

最后，考虑到人工智能自身不断变化发展的特性，企业要通过协作流程连接起任何加入的贡献者来自动化人工智能的全生命周期。通过采用一个敏捷的、云原生的平台，企业将能成功地让他们的数据为人工智能做好准备，引入开源软件，并把人工智能注入到自己的各个团队中。

在企业对数据基础架构进行现代化改造后，就能判断出他们在人工智能旅程中的位置。正如在上文指出的，人工智能之梯有四级：收集、整理、分析和注入。下面我们会一一介绍。

收集：让数据更简单且易于访问

在企业改造完他们的基础架构后，他们必须让数据的使用更简单且易于访问。长久以来，企业的数据都是被束缚在记录系统里，被僵化的平台、分离的业务功能和数据类型所隔离。这一切导致了数据孤岛的产生，让数据的访问变得很困难，进一步使得从数据中获取真正的分析洞察变得很困难。随着企业业务的发展，挑战会变得更加严峻。这种情况下，人工智能给企业带来转型的可能性几乎没有，因为洞察的好坏取决于数据的好坏。

下面是一些企业如何让他们的数据更简单且易于访问的案例。

使用所有类型的数据

如果一个零售企业目前只能访问到它自己过去的交易数据，这意味着它的人工智能模型只是使用一个数据源进行了训练。但是如果这家零售企业能把其结构化的交易数据和其他来源的非结构化数据（比如社交媒体或实时的点击流，或两者都有）进行结合，会发生什么？这样的话，这家零售企业就可以构建出人工智能模型，指出客户上周都买了什么，客户对商品的感受如何，以及他们当下正在看什么商品，这就能从更立体的维度来检视企业的业务。

投资人工智能注入的数据管理工具

数据库曾经只是一个记录系统和数据存储仓库，但这样的时代已经过去了。现在，从前端用户界面里使用自然语言和语音识别，到App的选择，一直到最后端的数据管理层，我们希望在现代技术栈的每层都能引入智能。一个人工智能赋能的数据管理平台可以提供给企业一整套的能力，来满足来自不同供应商、使用不同语言、针对不同地点和不同类型数据的查询。这意味着企业要能构建出合适的基础平台，保证能访问到他们的所有数据，并确保平台也能随着数据的增长而增长。

一个混合的数据管理策略不仅能解决当前数据质量相关的挑战，还能提供对更多数据的访问，推动更智能的人工智能。

例如，一家制造企业正在进行数字化转型。它不仅仅希望重构客户成本，还希望能降低总拥有成本（TCO）和IT成本。从性能指标的报告到企业收入和利润表现的评分，这家企业有PB级别的数据生成。其旧的数据库管理软件的成本是不可持续的。因此，这家企业决定把他们的数据库系统迁移到一个更灵活和高费效比的平台。通过使用一套混合的数据管理系统来收集数据和提供更好的数据访问支持，这家制造企业已经降低了数据存储空间，提升了性能，并有更快的交易响应时间。不管是支持分析平台还是交易平台，这家企业现在不光有了一个可靠和可用的数据库系统，更重要的是它还获益于预期的总拥有成本（包括数据库、平台和管理的服务成本）的降低。

现在，基于这个坚实的数据基础，这家企业的IT部门可以提供切实可行的业务报表，而不是让大家自己从系统里去拉数据，再试图去搞清楚怎么处理数据。这个基础也帮这家企业打下了继续向前去实施人工智能和机器学习解决方案的地基，从而能继续完成它的数字化转型。

整理：奠定业务就绪的分析基础

云计算和移动革命已经加速了数据（结构化和非结构化）产生的速度。但是很多企业并不知道他们实际拥有什么数据，都存放在哪里，需要什么流程来使用，以及如何针对数据相关的法律法规进行合规。

在人工智能之梯的这一级，企业必须要考虑三个关键问题。

首先是要解决数据质量问题。企业需要判断他们的数据是否业务就绪，这意味着数据已经经过了清洗、完整且合乎规范，可以被用于构建人工智能的模型。如果数据不是业务就绪的，对于包括数据科学家、分析师和业务人员在内的数据消费者来说，需要寻找、理解并把数据导入生产环境去使用将会是一个持续的挑战。

根据Forrest《预测2019：人工智能》报告，采用人工智能的企业里，60%的决策人员把数据质量列为尝试交付人工智能所面临的最大的挑战。现在企业要花费80%的时间来为数据的生产化使用做准备，这就形成了阻碍业务敏捷性、竞争力和盈利能力的瓶颈。一个经过恰当设计和控制的数据湖可以消除数据的不一致，解决数据重复问题，并为用户创建单个版本的标准数据。通过在数据湖里对数据进行管理和控制，企业就能够发现需要的数据并对数据建立信任。

接着企业必须要很好地整理和记录他们的数据。要理解为什么这一步很重要，可以用图书馆来类比一下。如果图书馆仅仅只是一个堆满了几千本书的房间，它对于普通读者而言就没什么价值。图书馆之所以有价值是因为它们组织编写了图书目录，从而让用户能通过多种方法发现信息，例如可以找到某个作者、类型和主题的全部书籍。在每次有新书加入（或旧书移除）时，这个目录都要更新。数据也一样。企业也需要有关于他们的数据的目录，以便提供数据来源是什么，谁拥有它，与业务上下文相关的元数据等信息。当通过集成或复制新数据源来提供给用户新数据后，如果没有适当的记录过程（就如新书出版并出现在图书馆里一样）和更新的数据目录，数据就很难或不可能被了解、信任和使用。

最后，企业必须要管理数据，确保只有获得授权的用户才能访问。许多企业低估了不良数据治理可能造成的陷阱。例如，一家大型银行发现他们的一个销售经理和她的团队擅自修改了数据，最终给银行造成了数百万美元的损失，对这家银行而言这个教训太惨痛了。因为没有合理地管理自己的数据，从而让这家银行成为过于宽泛的数据访问授权的受害者。就如企业必须控制数据的访问权一样，他们也必须控制数据本身。

在无法确保使用的数据是可信、完整和一致的情况下，领导者很难对人工智能产生信心。数据是人工智能的“面包和黄油”。数据质量直接影响到人工智能的产出结果，因为人工智能的性能好坏取决于数据的

好坏。基于糟糕的数据，企业是没法产出可信和透明的人工智能应用的。企业需要对其数据进行清洗、组织、编目和治理，并确保只有那些应该访问它的人才能访问它。

下面是一些企业通过整理它们的数据从而形成可信的和业务就绪的分析基础的详细案例。

专注于数据准备和质量（DataOps和人工智能）

DataOps是数据运营的代名词。它是一套方法论，可以协调人员、流程和技术，从而能随着业务发展的速度向数据用户提供持续的、业务就绪的数据管道。它可以促进数据提供者和消费者之间的合作，创作出自服务的数据文化，并移除阻碍大规模提升业务敏捷性、竞争力和盈利能力的瓶颈。

治理数据湖

一个经过恰当设计和控制的数据湖可以消除数据的不一致，解决数据重复问题，并为用户创建单一版本的标准数据。通过在数据湖里对数据进行管理和控制，企业能够找到所需的数据并对数据建立信任。这样就让企业能轻松地构建人工智能模型，并从它们信任的数据湖里提取出可行动的洞察。

“现代化”改造应用

企业希望确保其数据安全并符合法规要求，以保护客户和用户。同时企业还希望能够通过在内部环境或在云环境中部署和更新测试数据环境，从而能用更少的时间和更低的成本部署更高质量的应用。

确保数据隐私和合规

现在企业担负着两个相互竞争的业务目标：获取利润和遵守不断演进的法律法规。履行保护隐私和个人数据的义务至关重要，这就需要企业能够发现和区分业务里面不同类型的数据。

提供360度信息驱动的洞察

一个企业里会有多来源的客户和产品数据，因此就需要有一个数据管理方案来构建一个单一、可信赖的信息视图，同时把实时数据作为一个关键资产进行使用。

分析：以可信任和透明的形态构建和扩展人工智能

当企业能够收集到数据，并把它们整理成一个可信任和统一的视图后，它就可以构建和扩展人工智能模型到整个业务线。这让企业从所有的数据（不管数据位于哪里）里获取洞察，并使用人工智能来改变它们的业务，从而在竞争中处于明显的优势地位。

阶梯上的这一级也有它自己独特的挑战。仓促的投资会带来工具、人员和流程等方面的严重问题。单点性质的解决方案会在系统集成、维护和支持等方面导致复杂性，并最终带来技术和成本上的挑战。对数据科学家需求的激增使得雇用和留住他们成为一个新的挑战。另外，缺乏一个统一的分析平台也常常会导致不好的投资收益比（ROI），并让那些寻求更积极影响的领导感到沮丧。在人工智能能被全面地扩展和自动化整个企业之前，企业需要对人工智能有更多的信任。

为了从零开始构建人工智能模型并扩展到整个组织，企业需要有覆盖整个人工智能生命周期的能力和设施。这包括：

构建

在这个设施里，企业构建他们的人工智能模型。可以把这种地方理解为艺术家的工作室或是木匠的工作台。这是企业构建和训练用于预测的模型的地方。在人工智能生命周期的这一阶段，对企业而言，至关重要的是确保使用正确的算法来构建模型。

运行

模型被构建后，就需要作为某个应用或业务流程的一部分被部署到生产环境里。模型被部署上线后，就开始决定索赔申请的结果、定价决策等。需要时还可以再次对模型进行训练。

管理

在模型被构建和部署上线运行后，核心问题就变成了如何对它用可信任和透明的形式进行扩展。为了解决复杂和多样的模型构建和运行的环境，企业需要一种工具。这个工具不仅要能管理这些环境，还能解释模型的预测是怎么产生的，并能够在企业范围内扩展模型。随着这个管理工具到位，企业就能够追溯谁修改了模型，模型

是什么时候被部署上线的，以及模型的血缘关系是怎么样的。通过跟踪这些内容，企业就能确保他们的模型没有偏见，同时模型都可被解释且透明。

在当今充满各种法规、《通用数据保护条例（GDPR）》和数据隐私法的世界里，企业参与人工智能的行为都受到严格的审查。企业需要管理他们人工智能的整个生命周期，从而能够面向消费者或是其他业务伙伴解释他们的系统是如何作出决策的，以及原因是什么。例如一家银行需要能够告诉消费者他们贷款申请被拒绝的原因是什么，以及消费者需要做什么才能改变这一决定。

如果不能先在人工智能应用中建立信任和透明性，企业将永远无法真正地在整个企业的工作流程中扩展其人工智能。下面是一些例子，展示了为什么企业需要了解人工智能的工作内容。

保险

保险承保人可以使用机器学习方法来对索赔申请进行更一致和准确的风险评估，从而确保对客户的公平。同时出于合规和业务目的，他们也要能解释人工智能的建议。

电信

与运维团队合作，数据科学家可以构建机器学习模型，并有信心向电信运营商建议主动性的设备维护。

医疗

医疗行业的企业为了保持合规性，必须跟踪和解释整个工作流程里人工智能的决定。并能智能地发现和纠正歧视，从而改进结果。

欺诈

近年来，欺诈行为持续增加。欺诈很难被发现，因为用于欺诈检测任务的数据非常复杂且通常是非结构化的数据。另外，如果数据是孤岛性存在的，企业就很难获得对欺诈行为的360度视图。这都会导致误报警或者未报警，可能给企业带来数十亿美元的损失。

使用预测性洞察、实时分析、更精细的建模技术和自动化技术，企业能够使用人工智能来帮助他们解决上述问题。通过投资一个混合的多云平台，企业就能够在一个高度受控和安全的环境里实现这一切。

注入：让人工智能运用于实践

在人工智能之梯的“注入”这个阶段，人工智能被运用到整个企业实践中。这意味着企业要把人工智能引入各个部门和多种流程中（从薪资、客户关怀到市场推广，再到进行预测、自动化和优化），从而帮助提升企业的业务能力和范围。在很多情况下，这就要求企业发展出全新的涉及到每个部门的工作流程和业务流程。可以想像一下按需叫车的运输公司和自动化仓储里面的分配网络，这些系统都有人工智能在里面帮助完成复杂的人与机器的合作，从而能以更快的速度和更大的规模完成非常复杂的任务。

下面是一些案例，介绍了不同行业的企业是如何把人工智能融合到他们的业务里面的。

客户关怀

人工智能助手可以访问数百个客户触点的数据并进行“学习”，还能把这些数据通过一个统一的方式共享给其他系统，从而让客户获得一致的体验，并且能感觉到这个企业是“懂我的”。一个正确运行、人工智能赋能的对话机器人可以提升净推荐值。因为它能快速解决问题，全年无休地工作。另外不管是通过网络聊天、短信、类似于Facebook这样的社交媒体，还是直接给客户打电话，它能用客户感觉最舒适的渠道触达客户。

知识工作者

人工智能可以被用来检查和比较海量的文档，从而能够理解它们并找出人类可能从来没有注意到的关联、洞察或是不一致的地方。例如，使用人工智能后，采购人员或律师就能花更少的精力更准确地研究几十亿页的合同和法律文件，节省下大量的时间和金钱。

市场推广

人工智能非常善于从海量分散的数据里发现洞察，找到人类分析师从来没有注意到的联系。现在，创新型的市场团队里都有数据科学家，因为市场团队面对着海量的数据，他们需要有能力使用人工智能来分析这些数据。从预测消费趋势，优化市场推广策略，到交付可以打动用户的、个性化的广告，这一切都正在改变市场团队运作的方式。

供应链

使用人工智能，企业现在能够理解他们的运营数据。能以前所未有的洞察来制定更清晰的决策，并可以在一些破坏因素影响业务交付能力之前避开它们。最终，这将能够提升整个供应链的效率，增加ROI并获得更好的客户满意度。

人力资源

企业的成功依赖于他的员工。通过使用人工智能，企业可以更准确地清晰地预测哪些员工可能会流失，或者帮助为空缺的岗位找出合适的候选人。它还能被用来帮助确定哪些员工更应该升职加薪。

保险

使用人工智能，保险公司的索赔批准流程可以更加一致和自动化，能更准确地评估风险，并确保客户获得公平的结果。这就可以让保险公司的员工去关注更高业务价值的任务，同时让他们也能解释人工智能的决策是如何做出的。

企业财务规划

通过投资人工智能赋能的财务分析解决方案，企业可以发展出下一代的CFO。他们能运用预测性分析技术，解锁之前藏于数据中的价值。这就让企业可以摆脱掉那些人工进行的、易出错的工作。

这些例子展示了，随着企业不断地从数据中获取更多的价值和转变业务，人工智能能带来的影响。

例如，半个多世纪以来某大型独立能源公司一直是全球石油和天然气的引领者。它成功的秘密是什么？它招募和培养了一批世界上最好的工程专家。这个方法帮助这家企业在一些最遥远的海区建设了这个星球上最大的一些设施，并能安全地把生产出来的能源运输到全球，从而让下一代还有能源可用。

这家能源公司知道它必须利用好他的最佳员工的知识。通过使用人工智能，这家企业可以从它过去30年积累的密集和复杂的工程数据里提取出有意义的洞察，并把这数十年的知识给到每个员工。这一切都在帮助这家企业更快地回答困难的问题，从而帮助员工在复杂的项目里做出基于事实的决策。

虽然这只是企业如何把人工智能融入到核心业务的一个例子，但同样的方法可以在任何行业被复制，因为它从让人工智能起作用开始。无论企业的规模和所处的行业，这个方法用全新的方式来自动化和优化流程，并运用人工智能来解锁新的业务价值。

结论

在当今这个混合多云的世界，越来越清晰的是如果企业想成功发挥人工智能的能力，他们就必须更新改造数据基础架构，让信息能被收集、存储、整理、访问和分析。

当企业寻求采用人工智能时，记住三件事情很重要。首先从问题开始。无论企业是刚刚开始他们的人工智能之旅，还是已经在路上，请务必回头去分析和寻找企业试图去解决的核心业务问题。在大部分情况下，从需求和痛点开始找。例如，人工智能怎么能帮企业构建卓越的、个性化的客户体验？

其次，没有IA（Information Architecture，信息基础架构）就没有人工智能（AI）。企业需要一个现代化的信息基础架构来从所有需要的来源获取数据，并让各个团队的用户都可以访问。这个基础架构还能支持动态地构建和部署人工智能模型，并简化和统一跨云环境的数据和人工智能服务。人工智能之梯就是旨在帮助企业构建信息基础架构，并最终达成企业的人工智能目标。

最后，人工智能不是魔法，而是辛苦的工作。它需要合适的工具、方法论和思维模型，来克服企业所面临的各种障碍（数据、技能和信任），让企业最终真正地拥抱人工智能并把它运用进实践。

人工智能是我们这个时代最大的机遇。不过社会上还有一些顾虑，认为人工智能会替代人类的工作。现实是：人工智能不会替代管理者。相反，使用人工智能的管理者会替代不使用的那些人。

作者介绍

Robert D. Thomas是IBM数据与人工智能总经理。他总揽IBM的产品投资策略、销售和市场推广、研究实验室和全球软件产品的开发。

Rob领导的主要产品包括Watson、Db2、Netezza、Cognos、SPSS和InfoSphere。在加入IBM的软件部后，Thomas先生担任了多个职位，承担的职责也不断增加，包括销售、产品工程师、销售和市场、以及总经理。他还监督了IBM价值25亿美元的四次收购。

Thomas先生在范德堡大学获得经济学学士学位。在佛罗里达大学攻读MBA期间，Thomas先生从事证券研究，并学习了应用经济学、金融学和金融分析。

Thomas先生出版了两本著作：《大数据革命：农民、医生和保险代理人能教会我们什么大数据的模式》（John Wiley & Sons, 2015）和《技术公司的终结》（自出版）。这些著作可以指导各个行业的商业领袖们如何应对数字化颠覆。现在，他主要通过个人博客（robdtomas.com）写作。