

## 第一章

# 3D数据科学简介

你看过电影《The Matrix》吗？能想象成千上万绿色字符在黑色背景上不停流动的画面吗？这些字符代表了一个计算机程序中的“世界”，程序有一套规则和算法来决定事物如何运作。在不牵强基本原理的前提下，我们3D数据科学的设定与这个理念相差不远。想象拥有创建现实世界现象精确数字环境的能力。通过现有工厂，你可以研究引入机器人和传送带的新虚拟装配线的影响。在医疗领域，你可以模拟不同医疗方案对患者虚拟模型的效果。基于城市模型，你可以预测自然灾害的影响，并提出最佳紧急响应方案。数字化我们的世界能带来巨大收获。但我们从何开始？在3D数据科学中，我们生成并操作描述环境（如你的城市）的电子表格。然后我们结合跨学科技，提取有价值的见解，设计能够增强三维基础的工具。这通过实现与现实的准确数字连接，彻底改变了我们的决策能力。最棒的是？你不需要博士学位就能开始。基本的好奇心和一点编程技能（想象非常基础的Python）就足以让你投入这个令人兴奋的领域。本章奠定了我们探索3D数据科学的基础。我解释了维度在科学中的重要性，以及空间AI如何利用几何学模仿大脑分析3D数据的方式。我们探讨了3D数据科学的定义及其快速发展的原因。随后，我介绍了3D数据科学的模块化工作流程，并讨论了相关挑战，包括3D数据采集，

预处理、增强、标注数据集的稀缺性、计算需求、模型构建、可解释性、性能和稳定性。为了确保对这些概念有深入的理解，我包含了一个动手练习，我们将探索和操作来自OpenTopography存储库的点云数据。

## 3D 数据科学简述

科学是关于我们周围世界提出问题、收集证据以回答这些问题，并利用这些证据来发展可测试和完善的解释的过程。它基于观察、测量、实验和逻辑推理。当它利用3D数据时，就形成了一个迅速发展的领域：3D数据科学（图1-1）。

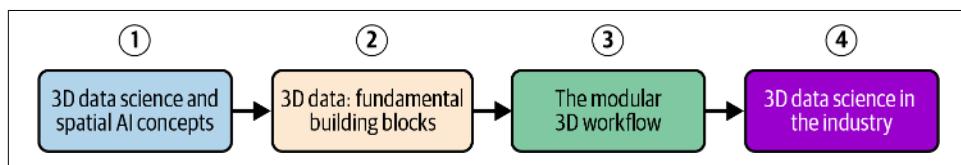


图1-1. 3D数据科学的四个核心组成部分

这个领域架起了我们三维世界与从数据中提取知识和洞见的科学之间的桥梁。它涵盖了一系列技术，从处理复杂的3D模型到制作令人惊叹的可视化，揭示隐藏的模式。

要开始，我们需要了解那个三维世界。下一节将详细解析什么是维度以及是什么让它成为3D。

## 尺寸与3D数据科学

“维度”这个词在你心中唤起了什么？你是否将其与科幻联系在一起，比如某个虚无世界的黑暗维度？虽然与虚构毫无关系，但在科学的语境中，特别是三维数据科学中，我们探讨的维度却有着完全不同的含义。

本质上，维度是某事物的可测量方面或特征。从数学的角度来看，维度指描述一个系统或对象所需的变量数量。从3D数据科学的角度，维度通常指空间范围。现在，让我给你一些例子。

一维对象可以用一个变量来描述，比如沸水的温度或本章的页数。二维数据指的是以两个维度表示的信息，通常使用x和y坐标。利用这些x和y坐标，可以创建由像素组成的数字图像，这些像素是

类似于地理地图，这些地图利用纬度和经度坐标在平面区域上描绘地点和特征。

延伸这一逻辑，我们能够理解我们所处的现实世界被认为是三维的，空间中的点由 x、y 和 z 坐标定义。这就是为什么当我们把维度与我们的视觉感知系统联系起来时，本质上是指将物体与其上下文环境联系起来的三维空间。例如，一把椅子是一个三维物体，因为它有长度、宽度和高度，这使我们能够将它与椅子的图片区分开来，图片是二维的表现形式。



我们可以进一步推进维度的概念。一个n维的物体或空间可以用n个变量来描述。虽然很难直观地理解，但像高维空间这样的数学概念常常在物理学和几何学等领域中被探讨。例如，弦理论假设宇宙可能存在于10维或11维中。但当然，随着维度数量的增加，我们为描述物体和系统增加了更多的复杂性和可能性。

维度成为了让我们固定所经历现实的基础。借助3D数据科学，我们获得了一种全新的方式，试图模拟大脑处理这些“3D数据”的分析部分。所以，假设我们能以某种方式获得观察到的三维形状的虚拟复制品，它可以作为我们进行科学的研究的一些数据基础。空间人工智能为数字世界开启了令人难以置信的脑启发式可能性之门。

## 空间人工智能：从现实到虚拟

空间人工智能是一种强大的工具，它使机器能够感知、推理和与世界互动，方式类似于人类的空间智能。它就像赋予计算机一双能够超越平面屏幕的眼睛，理解深度、距离和方向的细微差别。它由以下基础组成部分构成：

### 知觉

机器如何“看见”三维世界？传感器如LiDAR、摄像头和深度传感器捕捉空间信息。

### 推理

机器感知了世界，它们如何理解这些信息？算法和技术使空间推理和决策成为可能。

### 行动

机器如何将它们对三维世界的理解转化为有意义的行动？准确的数字环境构成了坚实的基础。

让我们进一步拆解这个问题。三维形状主导着我们的物理世界。如果有人让你描述你所在的房间，你很可能会将你的环境描述为由3D形状和几何体组成的集合。当我写下这些文字时，我将键盘视为主要的三维对象，并勾勒出每个键的形状，这些形状与变形的立方体紧密相关。当我抬头时，我的屏幕静静地矗立在我面前一米远的位置，位于我当前对办公室的局部空

当你分解这种定位和观察行为时，你会惊讶于我们的脑袋能做些什么。A. M. Treisman 在1980年用简单的术语解释了人类视觉感知的复杂机制：

当我们睁开眼睛看到熟悉的场景时，我们会立即对可识别的物体形成印象，  
这些物体在空间框架中有条理地组织起来。<sup>1</sup>

对于未受损的人类来说，它通常是我们认知决策系统所依赖的主要信息来源。我们的脑袋会延展自己并快速适应新环境，保留视觉中最关键的信息。如果我们深入观察，大脑每秒只接收三个“图像”。然后，这些图像会被整理并与已有的先验知识结合，创造出我们所体验的现实。<sup>2</sup>

这种高效机制使我们能够在红灯时刹车或阅读并理解文字。我们的视觉还可以适应不同的注意模式，例如取向注意，它通过不完全处理周围环境来节约能量；以及发现注意，它在“大脑从记忆中收集数据以全面理解场景”时以“慢模式”运行。

借助先进的计算能力和非物质化技术，现在可以虚拟复制复杂的过程。虽然复制某些操作具有挑战性，但研究人类如何与环境互动可以帮助我们更好地理解可有效使用的界限和机制。让我举个例子。

图 1-2 描述了物理世界与数字世界之间的关系，特别是在数字孪生的背景下。该过程始于现实世界，传感器捕捉各种物理对象或系统方面。这些传感器可以是从拍摄建筑立面的摄像头到记录温度的温度计。捕获的数据代表物理世界的现状/建成状态。然后处理这些捕获的数据，用于创建物理对象或系统的数字复制品。

---

<sup>1</sup> A.M. Treisman 和 G. Gelade, “注意的特征整合理论”，认知心理学，卷136，第97–136页，1980年。<sup>2</sup> F. Poux, “智能点云：结构化三维智能点数据”，博士论文，列日大学，2019年。

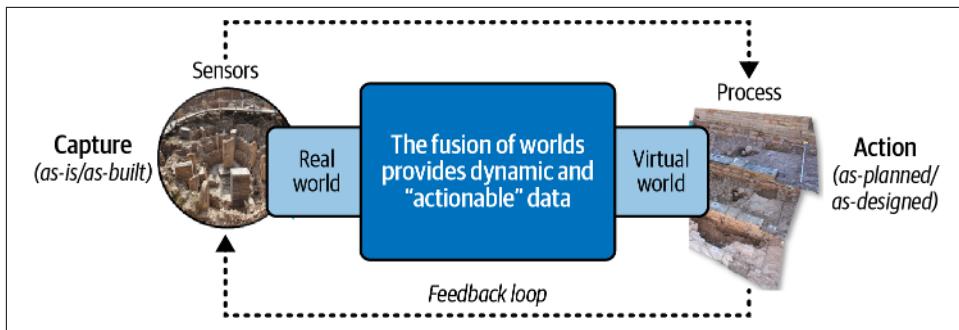


图1-2. 在数字孪生背景下从现实到虚拟的过程

这个数字孪生体存在于虚拟世界中。融合过程允许将现实世界中传感器的数据连接起来，这些数据可以反馈到数字孪生体中。模拟或模型可以更新，以反映现实世界中不断变化的条件。这在两个世界之间创建了一个反馈循环。

即使在分子层面，构成粒子也使用特定的几何形状，赋予它们所描述的内容以功能意义。这个迷人的特性主要可以通过3D数据科学来研究：这是数学、计算机科学和3D数据专业知识的完美结合（图1-3）。它利用这种协同效应，将3D机器学习、3D数据处理以及几何与空间分析推向新的前沿。混合起来，我们获得了一个强大的新领域，具有广泛的应用，并且与最前沿的工作流程密切相关。这是一个强大的研究领域，正通过3D机器学习、3D数据处理、几何和空间促进空间人工智能的进步。



3D数据科学提供了一个探索先进技术并帮助解决现实世界问题的绝佳机会。3D数据应用广泛，对熟练专业人才的需求日益增长，使其成为未来稳固的职业选择。3D数据科学推动创新和创造力，帮助创造沉浸式虚拟体验、新产品设计以及先进的人工智能系统。通过掌握这一领域，您可以助力重要研究，解决复杂问题，并影响科技的未来。第二章回顾了数学、计算机科学和3D数据专业的主要概念，以帮助拆解主要组成部分。

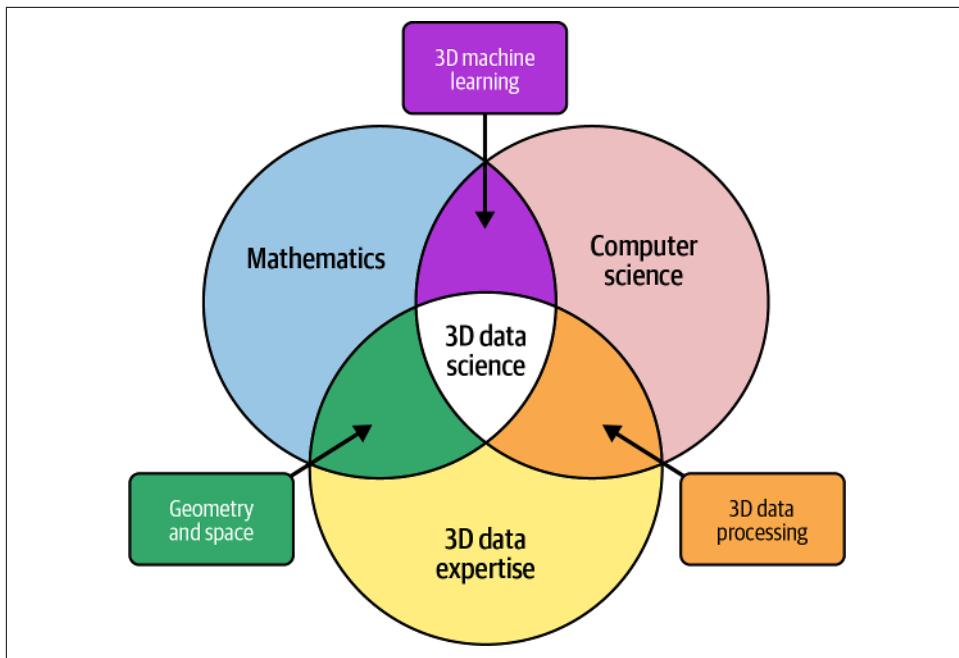


图 1-3。3D 数据科学是数学、计算机科学与 3D 数据专业知识的交叉领域

现在让我们转换一下话题，进入具体内容：建立3D数据科学的工作流程。

## 3D数据：基本构建模块

3D数据科学的主要目标是定义“数字环境”的边界。本质上，我们希望创建一个我们所居住世界的复制品，作为我们数字应用的坚实基础。这个对现实世界的精确数字表示将使智能代理能够以类似人类行为的方式执行任务。一旦这些基础建立起来，我们就可以连接到通过语音命令、操作和其他方式触发的决策元素。

第一步是要理解我们有一个基础，这个基础是真实世界。然后，我们以这个基础为出发点，创建我们所称的低级数字模型（图1-4）。它本质上是我们所观察到的几何描述。通常，大多数工作流程止步于此。但是，要达到所谓的“智能系统”或可以进行操作和决策的数字环境，你需要更进一步，创建一个高级数字模型。这个模型可以用于推理服务。

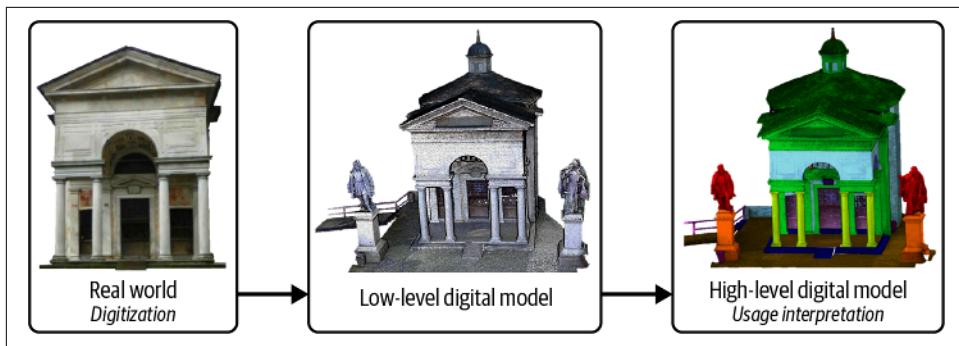


图 1-4。从现实世界到高级数字模型

这在实际操作中意味着什么？我们需要在几何基础上注入额外的信息，具体来说就是语义和拓扑。以图1-4为例，可以知道模型中哪一部分属于柱子，哪一部分属于屋顶，雕像的用途是什么，所有这些元素如何相互作用，它们彼此之间的关系是什么，如果其中某个元素出现故障会发生什么，或者涉及到什么物理原理。

我们的目标是创建一个简单的程序，利用现有的数据和信息存储库来生成场景的语义表示，整合概念及其含义。我们如何实现这一目标？在这种情况下，空间传感器充当我们的眼睛，捕捉数字空间资产，并利用现有知识（无论是直接还是间接）将其提炼为语义表示。可用性往往是在开发在线认知感知时遇到的第一个难题。我们的基于记忆的结构经过优化，可以快速访问所需信息。然而，用计算机复制这一过程极其复杂，创建一个适用于所有情况的通用解决方案是一项重大挑战。

第二个在虚拟认知决策系统时的瓶颈是创建语义表示。收集领域知识并将其附加到空间数据上，可能会由于数据类型、来源和表示的不同，导致复杂的集成和挖掘挑战。

所以，基本上，你想要将一切都按原样建模，或者基于它进行某种模拟。但我们不要先搭建带顶的系统。让我们先铺设砖块，然后再用砂浆把它们固定，只有在优化工具以加快建造速度之后，我们才能实现更稳健的3D数据科学工作流程。首先，从概念层面来说，我们需要结合导致准确语义表示的三个主要元素：几何、语义和拓扑。

# 几何学、拓扑学与语义学

理解几何学、拓扑学和语义学之间的相互作用对于从三维数据中提取有意义的信息至关重要。每一个概念都为空间信息的结构和意义提供了独特的见解。

## 几何学

几何学包含了对空间属性的研究，根据大小、位置、方向和比例来描述三维形状。为了更好地说明这一点，我们以一栋房子为例（图1-5）。在此背景下，几何学定义了房子的物理属性，包括其尺寸、形状和空间布局。

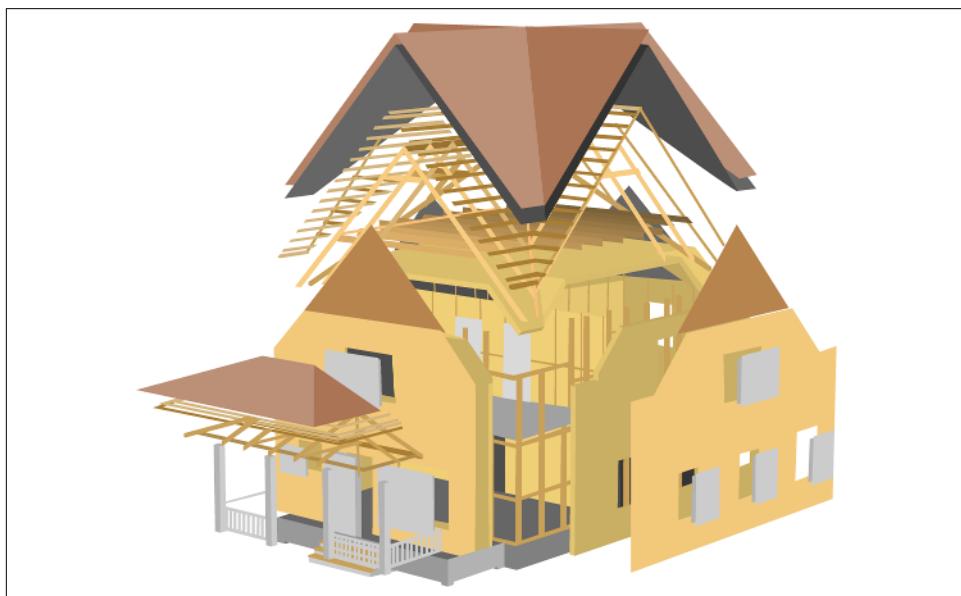


图 1-5. 一个几何房屋的示例：注意模型中需要表示的许多不同平面和表面

它允许提取可量化的方面，例如墙壁的长度和宽度、屋顶的高度以及不同表面交叉的角度。房屋的几何形状可以用多种方法描述。例如，您可以生成许多离散点，这些点采样在房屋结构元素的表面上。或者，您可以看到由点、线和面定义的多边形组合，这些多边形描述了物理占地面积。您还可以结合几个基本的3D形状（见图1-6），然后变形并组合它们以获得最终的3D几何形状。

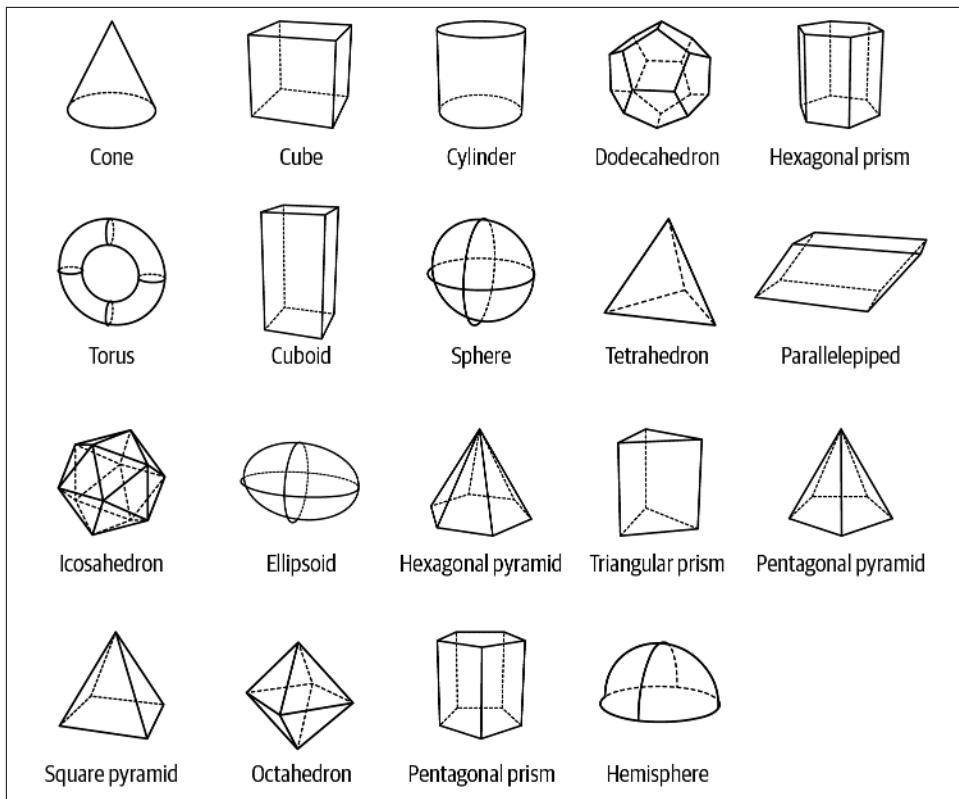


图1-6。可组合和变形以创建复杂3D几何体的基本3D形状

总体而言，获取几何形状通常是3D数据科学工作流程中的第一步。但如果仅停留在这一步，就无法收集到复杂决策场景所需的深刻见解。这就是为什么一旦你获得了可靠的几何形状，就需要定义拓扑结构。

## 拓扑学

拓扑学关注空间的内在属性，这些属性在连续变换下保持不变，无论是拉伸、弯曲还是撕裂。它涉及诸如连通性、邻接性、包含性和连续性等概念。

这可能更难理解，所以让我用我们的房屋示例（图 1-7）来尝试说明。拓扑帮助我们理解其组成部分之间的关系。例如，它识别出墙壁连接形成房间，门口连接房间。拓扑使我们能够分析房屋中不同元素的空间关系，从而提供关于结构布局和连通性的见解。

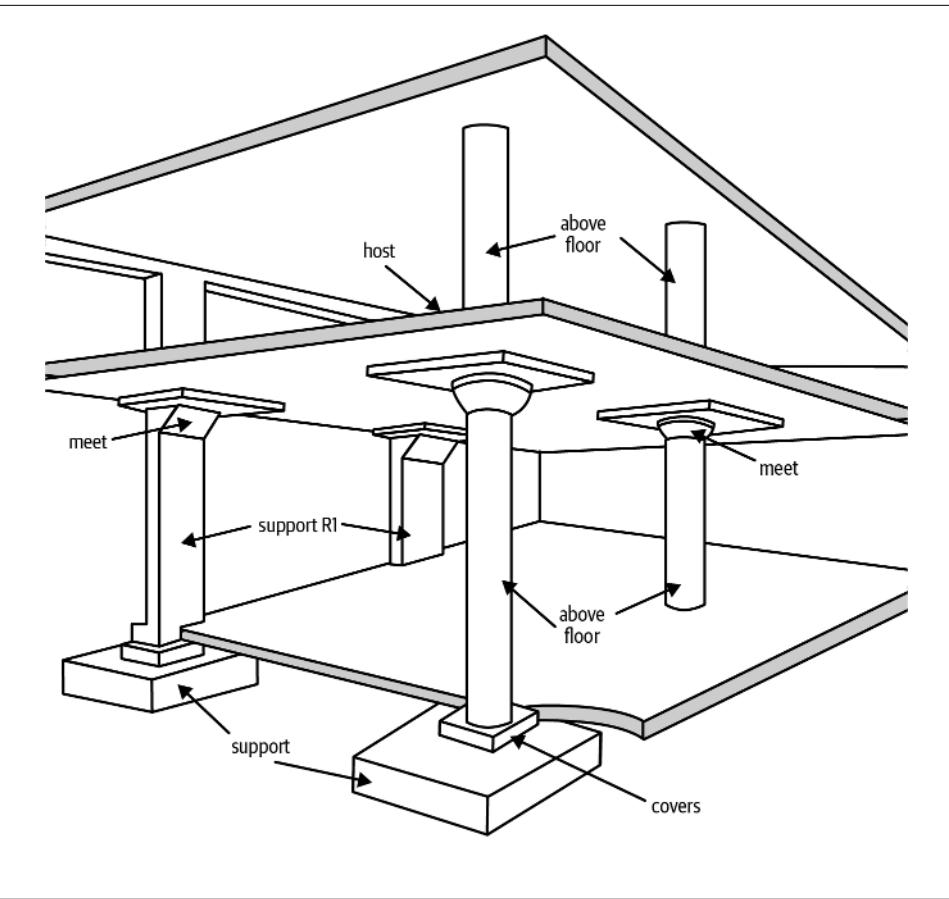


图 1-7。屋内几个元素之间存在拓扑关系：柱子支撑楼板，楼板立于支撑物之上，柱子排成一列

这是一种在三维几何场景中对感兴趣的对象进行情境化的方法。这在实践中听起来很棒，但我们如何为我们操作和拓扑定义的实体赋予意义呢？我们怎么知道某样东西是地板，而另一样是柱子呢？答案是，有了语义。

## 语义

语义指的是与数据集中的元素或结构相关联的意义、解释或重要性。它涉及分配标签、属性或分类，以为数据提供上下文和理解。同样，在房屋的背景下，语义使我们能够描述和分类其组成部分（图1-8）。

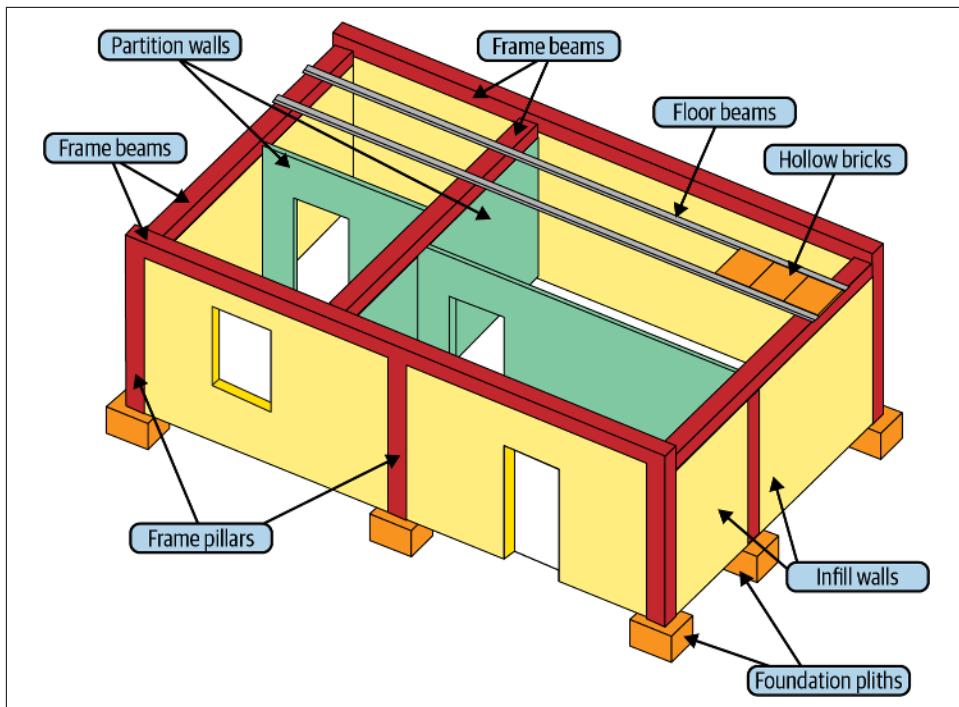


图 1-8。概念标签识别场景或对象的主要组成部分，我们以后可以将其与更深层的知识源关联起来。

语义使我们能够识别并标记您房子中的厨房、卧室和客厅。这个额外的信息层超越了几何表示，提供了对房子各部分用途和功能的更高层次理解。

自然地，这些支柱中的每一个都可以被独立考虑。然而，要开发最强大的3D数据科学解决方案，结合几何、拓扑和语义是一个必要的组成部分。

## 整合几何学、拓扑学与语义学

几何学、拓扑学和语义学之间的协同作用对于全面的三维数据分析至关重要。通过结合这些概念，我们能够对空间数据有一个整体的理解。例如，在三维城市环境中，几何学提供有关物理结构的信息，拓扑学定义这些结构如何相互连接，而语义学赋予诸如建筑物、道路和公园等各种元素以意义。

但是我们甚至可以描述我们认为的基础对象，比如椅子。我们可以将这个对象细分为其组成部分，并定义每个部分如何相互作用。

其他。在图1-9中，我们有一个靠背、一个座位和四条腿（腿1到4），它们位于厨房中。

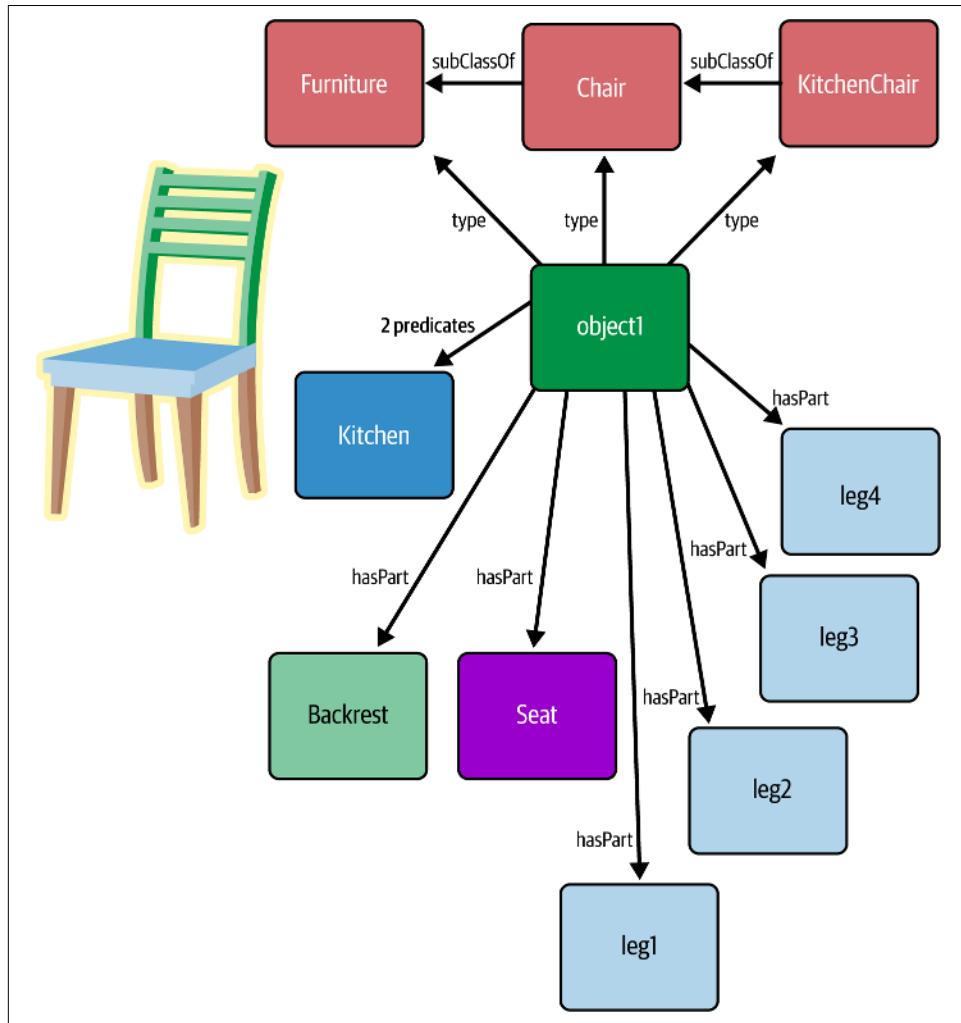


图 1-9。一把椅子，其几何形状、拓扑结构和语义描述

既然我们已经了解了我们在寻找什么以及为什么寻找，让我快速介绍一个经常构成研发解决方案基础层的核心数据模型：3D点云数据。

# 3D点云简介

实体和对象可以在不同的细节层次上进行描述。例如，一个点可以代表一座城市、一位人物或一个国家。然而，在3D数据科学领域，数据点通常通过一组传感器收集的。虽然单张栅格图像或视频流在某些情况下可能有用，但它们可能仅提供模拟我们3D视觉感知所需的深度信息。因此，通常需要更全面和多样化的数据基础。

现实捕捉设备能够以点云的形式捕获全面的三维空间信息，点云是一种由{X, Y, Z}坐标组成的空间集合，附带表示所记录环境的属性，这些属性基于传感器的优势和局限。

合适的硬件、软件和方法选择的格局已经成熟到可以创建现实世界的数字复制品，范围从微观级别到国家级别，如图1-10所示。

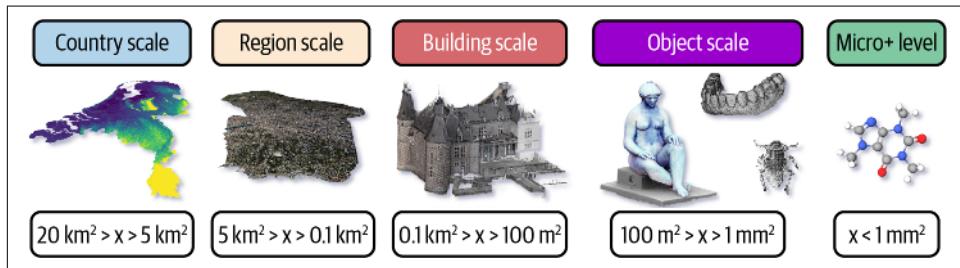


图 1-10. 不同数据集在各种尺度下的多尺度点云

获取点云变得更快且更易获得，得益于低成本的解决方案。然而，支持这些硬件进步的软件尚未跟上步伐，仍面临大数据的五个V挑战，即体量（volume）、速度（velocity）、价值（value）、真实性（veracity）和多样性（variety），如图1-11所示。

为了简单地解释这个概念，当我们使用不同的传感器和方法收集数据时，会得到多样化且体积庞大的点云数据集。这带来了处理速度和精度方面的挑战，因为我们需要确保数据的可靠性，并能够将其转化为有用的信息。因此，验证数据的真实性并及时高效地提取可操作的洞见是至关重要的。

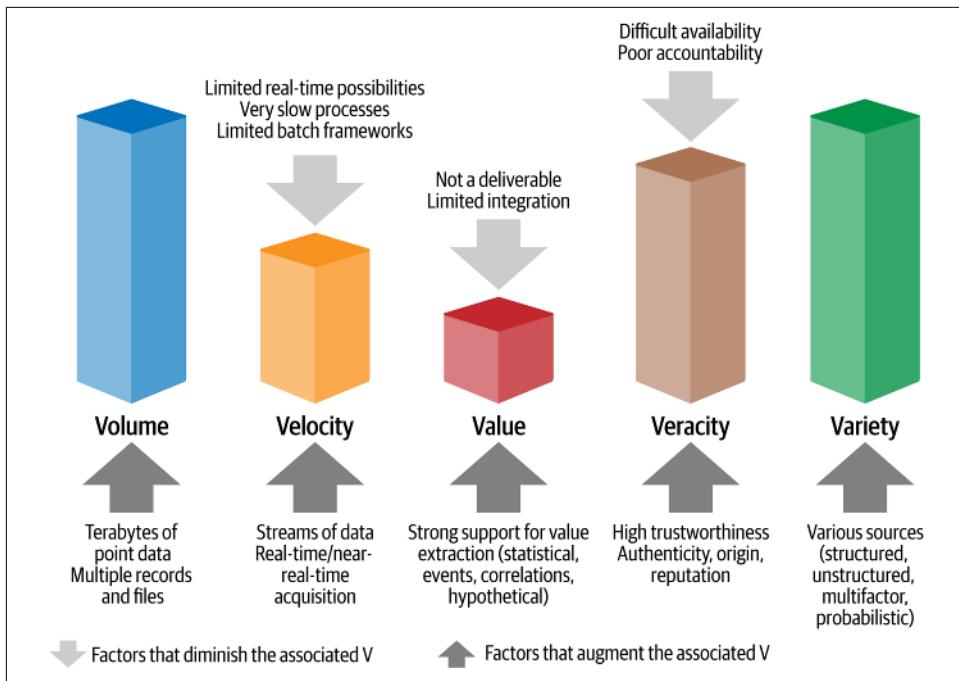


图 1-11. 点云背景下大数据的五个V特征

值得注意的是，依赖点云数据集的3D数据科学工作流程通常是特定于应用的。该工作流程遵循从收集相关数据到创建交付成果的经典步骤。虽然数据采集步骤可能特定于可用的传感器和平台，但点云作为交付成果正变得越来越受欢迎，并且现在是许多行业的首选。在这种以任务为导向的场景中，点云作为空间参考，供专家用来创建其他交付成果。因此，它是项目与现实最直接的联系，提供准确的真实世界信息，以基于数字现实而非过时或解释性信息来促进决策。

专家在许多生产操作中手动处理点云数据以提取有用信息。这个特定过程耗时且通常容易出错。3D数据科学的一个核心思想是将尽可能多的专家知识整合到数据本身中，以自动赋予空间实体语义和拓扑意义。这不仅节省时间，减少错误，还提高了此类操作的整体效率。



集成组件以允许在3D点云模型上实现自动化决策场景，指的是由Poux等人在2016年首次提出的智能点云概念。如果这个概念激发了你的好奇心，你也可以参考公开获取的“智能点云”论文，以构建3D智能点数据。

此外，将点云转换为特定应用的交付成果往往非常耗时，并且通常需要大量的人工干预。由于我们需要处理的信息量和复杂性持续增长，人工专家有效管理这些信息变得越来越具有挑战性，尤其是在处理分散于各个项目利益相关者和平台之间的冲突数据时。

因此，将大型点云数据转化为更高效的流程是实现可持续系统的关键。这为新一代产品和服务打开了大门，有助于决策制定和信息提取。在下一节中，我们将深入探讨高效的工作流程，以避免任务特定且不可持续的系统。

在继续之前，有一点重要说明。我有意用三维点云来展示三维挑战，以提供一个直观的表现形式。但当然，和任何三维数据科学工作流程一样，单独观看视觉效果而不结合三维数据模态，是无法展现完整情况的。在本章剩余部分，我将提供初步的工作流程定义，为各种应用奠定稳固基础。

## 3D数据科学模块化工作流程

你是否曾想过研究与开发如何过渡到现实世界的应用中？在构建生产就绪系统或解决工程挑战时，理解3D数据科学中工作流程的重要性至关重要。如果你回看图1-1，这是四个阶段中的第三阶段。

确实，拥有一个用于管理特定顺序中重复过程和任务的系统，有助于我们理解整体情况，同时清晰地划分每个步骤。在3D数据科学中，工作流程通常具有相同的结构，但根据应用需求提供不同的顺序或组合。“应用”一词非常重要，因为无数的工作流程可以用于数据科学应用。根据您的具体目标，您可能需要朝着特定方向前进。

然而，我建议采取一种更高级的方法，识别可以更广泛和整体使用的工作流系列。通过这种方式，您可以选择任何这些工作流并根据您的需求进行定制。有八个属于此类别的已定义工作流系列（图1-12）。

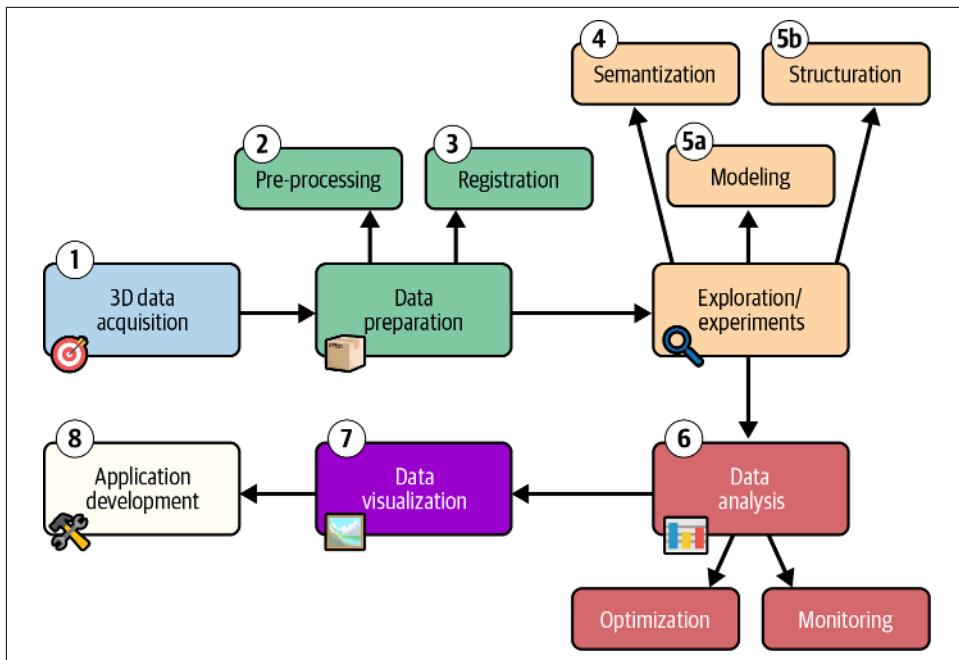


图 1-12. 3D 数据科学工作流程的高级视图

让我们探讨3D数据科学工作流程中的每个关键组成部分，以及它们如何协同工作以创造有价值的解决方案。

## 数据采集

获取这种额外数字表示的第一步是拥有一种捕捉它的方法。这就是我们基本上（但不限于此）依赖3D传感器的原因，这些传感器使我们能够捕捉物体和场景的三维形状，有时甚至还能捕捉其他属性。现在，3D传感器并不是一个单一的实体。它们有各种不同的类型，每种都有其优缺点。如果你想全面了解这些系统，我写了一份补充指南，你可以在3D数据科学资源中心找到。它特别关注立体视觉、结构光、飞行时间（ToF）和LiDAR传感器。



我们还有其他获取数据的方法，这些方法不是源于现实世界的场景，而是源于纯粹的创造力，如模拟和视频游戏。第2章和第3章详细介绍了获取这些数据集的明确方法，或如何最好地获取所需信息。此外，第18章将展示如何使用生成式AI工具进行3D数据集的生成。

本质上，这意味着捕捉三维数据使我们能够创建现实世界物体和环境的精确表示。我们可以通过使用数据采集向量和方法来评估给定的对象（例如城市），从现实世界物体创建三维数据表示。然后，处理数据并创建三维表示。

这些方法对于创建准确的数字环境至关重要，这些环境为城市规划、基础设施管理和灾害响应中的关键应用提供基础。结合三维数据提高了我们依赖准确且可信的数据表示进行决策的能力。此外，将这些数据整合到数字环境中可以生成极为逼真的现实资产和系统的复制品，从而提高模拟和分析的效率。

## 3D数字孪生

还有一个关于3D数据表示的重要概念需要讨论：3D数字孪生。这个术语你可能听过或使用过，但理解它涵盖的内容非常重要。我更倾向于避免模糊的定义，所以与其简单地称之为“虚拟表示”，不如来详细解析它的真正含义。3D数字孪生是现实世界中物理对象或系统的动态数据驱动副本。这个副本不仅仅是静态的3D模型；它会不断更新实时数据，反映其物理对应物的状态和行为。可以把它看作一个有生命、有呼吸的数字镜像。

这种实时连接至关重要。想象一下风力涡轮机。这个涡轮机的三维数字孪生不仅仅显示其形状和位置。真实涡轮机上的传感器会向数字孪生传递关于其转速、温度、振动水平，甚至周围风况的信息。这使我们能够远程监控涡轮机的性能，预测潜在故障，并优化其运行——所有这些都在数字领域内完成。

“3D”方面也至关重要。虽然数字孪生可以存在于二维甚至纯数字表示中，但三维方面允许空间理解和分析。我们可以可视化涡轮叶片与风的相互作用，模拟不同负载下的应力点，甚至在虚拟环境中规划维护操作，然后再应用到现实世界中。这种空间背景带来了巨大的价值，尤其是在工程、制造和城市规划领域。

现在，我们如何实际构建这些数字孪生？这是一个多阶段的过程，首先是对物理对象进行高度准确的3D表示。接下来，我们将这个3D模型集成到我们的3D数据科学管道中。最后，我们需要一个平台来可视化并与数字孪生进行交互，从而提取见解并做出明智的决策。

因此，可以将数字孪生视为特定物理实体或系统的动态、数据驱动的虚拟表示，能够与其现实世界的对应物持续同步。数字环境虽然也是虚拟空间，但其代表的是更广泛的背景或环境——例如工厂车间、城市或自然景观——并不一定与某个物理对应物保持实时的、个性化的连接。

这有助于澄清整个概念，并解释为什么我在本书中特别使用“数字环境”一词，避免使用不准确的炒作分类。

现在，我们准备好好看看来自3D数据采集技术的3D数据集在实际中的样子了。让我们通过一个六步流程，利用图1-13所示的开放数据存储库进行演示。

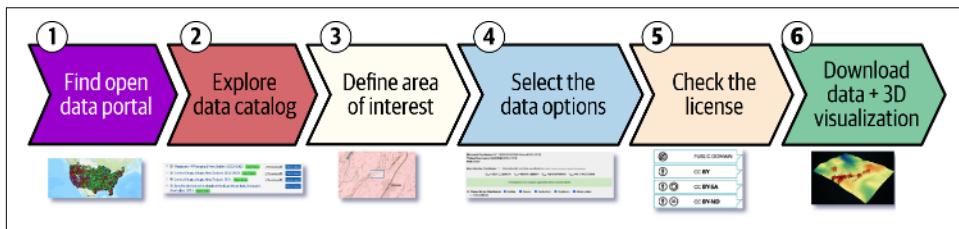


图1-13。一个数据整理工作流程的示例，该流程利用开放数据仓库，而不是依赖现场3D数据采集。

第一步是确定一个开放数据门户。这个例子使用开放数据，因为对于这些目的，我们不想现场进行数据采集，这给了我们一个机会来评估哪些数据是公开可用的。让我们从一个获取三维点云数据集的绝佳资源开始：Open Topography。

从该网站，我们可以浏览数据目录（步骤2），并从中选择我们想要探索的内容。以本例为例，选择Manawatu - Whanganui，新西兰2022-2023。在步骤3中，选择一个小区域以访问点云数据。接下来，选择生成3D点云网页浏览器的选项，并取消生成TIN。最后，勾选使用许可证（CC BY 4.0）和数据引用3，提交我们的任务并获取我们的3D结果。如果满意，我们可以下载我们的数据集和许可证以作说明。

既然我们已经拥有了3D数据集，接下来让我们考虑整体流程中的下一步。

3 Horizons Regional Council, Toitū Te Whenua Land Information New Zealand (LINZ) (2023). Manawatu -Whanganui, 新西兰 2022-2023。由 Landpro 收集，OpenTopography 和 LINZ 发行。访问时间：2023-11-14。



本节旨在确保您能够快速访问实用事项。为了避免在此阶段超出我们的认知能力，我们将继续进行工作流程的下一阶段。您将在第2章（软件）、第3章（方法和工具）以及第18章（结合空间人工智能的3D生成型人工智能）中找到更多关于3D数据采集的方法和工具。

## 预处理

在3D数据科学工作流程中，进行数据预处理步骤是必不可少的。需要干净、无噪声且易于管理的数据集作为基础，以提取重要特征进行后续处理，如配准、分割和分类。

这个过程分为六个主要类别（图1-14）：数据分析、数据清洗、数据缩减、数据转换、数据丰富和数据验证。预处理管道几乎系统地专注于解决噪声降低和缺失数据问题。

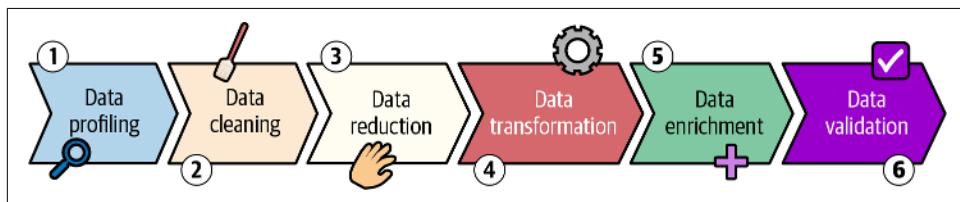


图 1-14. 数据预处理工作流六个主要组件的高级视图

我们将在第5章和第6章更详细地探讨每个步骤，但首先，让我非常简单地概述每个步骤。数据分析很直接：它涉及对手头数据的深入理解。数据清理则侧重于处理噪声和缺失数据，特别是在恶劣天气条件下收集数据的场景中。数据降维包括抽样策略和降维技术，而数据转换包括归一化、缩放和特征选择。

这些步骤可以通过各种策略实现，利用编程语言。正如你所知，本书将重点介绍Python。让我们运行一些代码来揭开数据预处理的神秘面纱。为了不设置本地环境（我们将在第3章介绍），我们可以使用云解决方案，比如Google Colab。



让我来谈谈Google Colab，借此类比一下大家熟悉的东西——云端的一台完全配置好的工作站。它类似于通过网页浏览器访问的计算机，预装了数据科学工具，如Python库、Jupyter笔记本界面，以及处理单元（GPU、CPU和TPU）的访问权限。本质上，这就是Google Colab。它允许你在浏览器中编写和执行Python代码，利用硬件加速，无需本地安装或资源管理。这对于尝试新想法、共享交互式分析或者着手需要大量资源的任务，非常方便，无需大量前期投入。可以把它看作你的按需编码环境。我喜欢用它进行不需要复杂3D可视化的快速演示，这正是本节内容的情况。

要跟随操作，你可以打开并执行我创建的Colab上托管的Python脚本，First 3D Steps，或者如果你已经非常精通Python，可以使用以下代码：

```
# 导入第一个库import numpy as np# 导入第二个库import matplotlib.pyplot as plt# 定义立方体顶点的坐标vertices = np.array([[-1, -1, -1], [1, -1, -1], [1, 1, -1], [-1, 1, -1], [-1, -1, 1], [1, -1, 1], [1, 1, 1], [-1, 1, 1]])# 定义立方体的边edges = np.array([[0, 1], [1, 2], [2, 3], [3, 0], [4, 5], [5, 6], [6, 7], [7, 4], [0, 4], [1, 5], [2, 6], [3, 7]])# 创建图形和坐标轴对象fig = plt.figure()ax = fig.add_subplot(111, projection='3d')# 绘制立方体顶点ax.scatter(vertices[:, 0], vertices[:, 1], vertices[:, 2])# 绘制立方体边for edge in edges:    ax.plot(vertices[edge, 0], vertices[edge, 1], vertices[edge, 2], color='black')# 设置坐标轴标签ax.set_xlabel('X')ax.set_ylabel('Y')
```

```
ax.set_zlabel('Z')

# 显示图表plt.show()

这里，我们简单地创建了一个由顶点（点）和边（点之间的线段）组成的
三维立方体，然后在欧几里得空间中绘制它。之后，我们可以通过平移和
缩放来变换立方体：

# 将立方体平移 (1, 1, 1)
vertices_translated = vertices + np.array([1, 1, 1])

# 将立方体按2倍缩放
vertices_scaled = vertices * 2

# 创建一个图形和坐标轴对象fig = plt.figure()
()ax = fig.add_subplot(111, projection='3d')

# 绘制原始立方体顶点
ax.scatter(vertices[:, 0], vertices[:, 1], vertices[:, 2], label='Original Cube')

# 绘制翻译后的立方体顶点
ax.scatter(vertices_translated[:, 0], vertices_translated[:, 1],
           vertices_translated[:, 2], label='平移后的立方体')

# 绘制缩放后的立方体顶点ax.scatter(vertices_scaled[:, 0], vertices_scaled[:, 1],
vertices_scaled[:, 2],label='Scaled Cube')

# 设置坐标轴标签ax.s
et_xlabel('X')ax.set_y
label('Y')ax.set_zlabe
l('Z')

# 添加图例ax.legend()

# 显示图形plt.show()
```

在前面的例子中，你应该获得原始立方体的蓝色角点、缩放立方体的绿色
角点和移动立方体的橙色角点。变换扩展了平移和缩放的情况，我们将在
本书的其余部分深入探讨这一方面。一旦这一点得到很好的约束，我们通
常会进入数据丰富阶段。它与所谓的特征提取相关。我们的目标是从数据
中提取数值，以丰富整体理解。

这些值是特征，它们需要满足三个主要属性：

- 它们需要具有信息量。推导出不能提供比我们已有信息更多内容的东西是没有用的。
- 它们不必是多余的。如果可以的话，我们必须将其限制在提供实质性信息的内容，而不是非常有限的信息。
- 最后，他们需要在向下传递的过程中扮演促进者的角色。

让我来举例说明一个经典的3D物体分类任务：你想知道一个3D物体是椅子还是桌子。你需要从数据中提取一些信息作为特征，以便对你的模型有帮助（例如，大小、沿z轴的点分布、高度）。各种方法都存在，无论是作为特征工程阶段的一部分（参见第6章和第7章），还是包含在主要3D机器学习架构中（涵盖第12章到第16章）。

在这个阶段，我们已经可以注意到理解目标应用更大目标的重要性和影响。不同的应用，如结构偏差分析、语义分割、分类和3D建模，要求不同的预处理和特征计算水平。最后，我们进行数据验证阶段，以确保我们满足先前设定的具体标准。



处理数据时通常需要额外的步骤。这些步骤可能包括准备一个提供更详细信息的兴趣区域、创建加速结构数据、计算预特征以及将3D数据集转换为对我们的目标有用的形式。因此，重要的是要记住，工作流模块的序列化在实际中可能会有所不同。

3D数据预处理在每一章节中都有广泛应用，深度取决于最终目标。一旦我们实现了将初始杂乱数据转化为精炼且有价值的数据集的目标，我们就可以准备进行数据集之间的配准挑战。

## 注册

当处理来自不同视角的多个数据集时，将它们对齐到一个统一的参考框架变得至关重要。无论是将同一对象的多次扫描合并，还是将从不同角度拍摄的照片汇集，配准工作流程都能实现整体分析和理解。此外，数据配准还处理数据不完整的情况，并需要能够对内部的组成部分进行平移、旋转和缩放。

一个典型的场景是，如果你有两个来自不同视角的“扫描”，并且你想将它们都带入一个视角（例如，图1-15）。在这种情况下，我们想要

合并两个来自摄影测量资产的扫描。我们首先提取一些感兴趣的特征，然后应用一种利用这些特征的配准技术，将两个数据集对齐到一个共同的参考框架中。

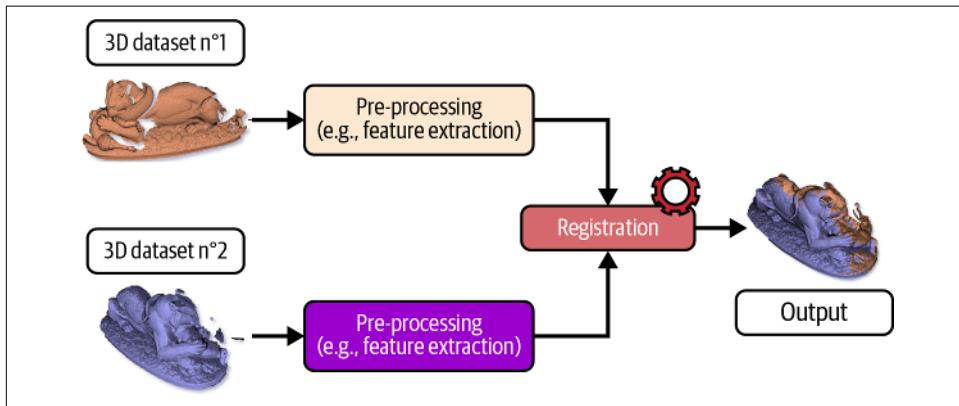


图 1-15。数据注册过程示例

我们称这项任务为数据配准，其目标是将两个局部坐标系统转换到一个公共参考框架，使数据集能够协调对齐。这也指代了跨平台配准，即你拥有不同的视角、不同的采集尺度以及不同的数据类型。

然而，这个过程要复杂一些，可能涉及数据校准或数据融合的子任务，这将扩大本书的范围。对于有兴趣的人，这些内容在3D资源中心有深入的讲解。



数据注册是创建准确数字环境的基础。第6章提供了一个完整的解决方案，包含一个从全局到局部的注册方案，使用Python实现，您可以复制并扩展到自己的项目中。

现在，让我们继续讨论3D数据科学工作流程中最重要的模块：语义化。

## 3D数据分类（语义注入）

我们在3D场景理解框架中通常涵盖的任务是语义化，我们将其称为分类。这一步的主要目标是理解输入数据并解释传感器数据的不同部分。例如，我们有一个户外场景的点云，如高速公路，由自主机器人或汽车采集。

分类的目标是确定场景的主要部分。换句话说，我们需要知道点云的哪些部分是道路，哪些部分是建筑，或者人类在哪里。从这个意义上说，它是一个整体的类别，旨在从我们的传感器数据中提取特定的语义信息。然后，我们希望在不同的粒度上添加这些语义。让我用更多细节描述图1-16中表示的每种技术：

### 3D物体检测

3D物体检测技术，或称边界框检测技术，用于将物体包围起来。它们是许多应用的重要组成部分。基本上，它们使系统能够捕捉物体在世界中的尺寸、方向和位置。因此，我们可以将这些3D检测应用于现实场景中，例如增强现实应用、自动驾驶汽车或通过有限的空间/视觉线索感知世界的机器人。漂亮的3D立方体包含着不同的物体。但如果我们要微调物体的轮廓呢？

### 3D语义分割

通过语义分割技术，我们能够解决最具挑战性的任务之一：为属于感兴趣对象的每个基本单元（即点云中的每个点）分配语义标签。本质上，3D语义分割旨在更好地描绘场景中存在的对象。如果你愿意，可以将其视为增强版的3D边界框检测。这种方法提供了每个点的语义信息，使我们能够深入分析。然而，仍然存在一个限制：我们无法直接处理每个类别（类）中的不同对象。针对这一点，我们是否也有相应的技术？

### 3D 实例分割

是的！这被称为3D实例分割。它的应用范围更广，从自主系统中的3D感知到地图绘制和数字孪生中的3D重建。例如，我们可以想象一个库存机器人，它能够识别椅子，计算椅子的数量，然后通过抓住第四条腿来移动它们。实现这个目标需要区分具有相同语义标签的不同语义标签和实例。把实例分割想象成语义分割的超级加强版。

3D 几何体和语义在此阶段被提取并附加到实体上。接下来，让我们考虑结构化和建模模块。

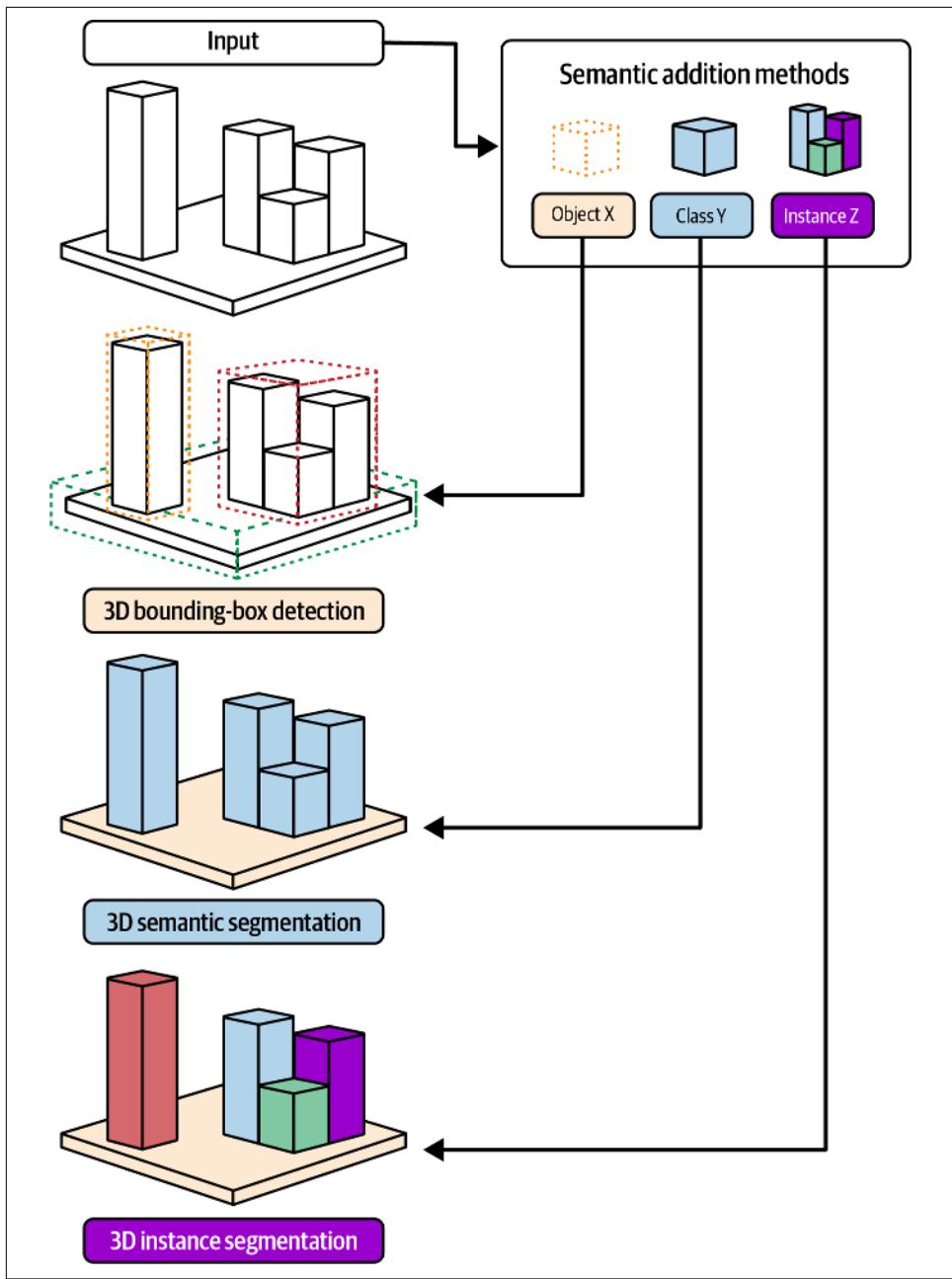


图 1-16. 三种主要的语义化技术：3D 对象检测（边界框检测）、3D 语义分割和 3D 实例分割



语义化模块是新兴AI方法重点关注的对象，展现了3D机器学习与3D深度学习方法之间的相互作用。书中有很大一部分内容专门介绍这些主题，以确保您掌握最前沿的理解和处理这些复杂任务的方法。您可以在第9章以及第12至18章中找到基于原理、算法和AI方法的技术和Python解决方案。

## 结构化/建模

结构化3D数据集常常被忽视，但这是一个关键步骤。它确保数据以一种便于操作和理解的方式组织。同时，它还确保您能够创建与模拟或数据试图捕捉的基本现象完美契合的连贯模型。我们可以利用结构化技术，如构建分层树等，以促进更快的处理和分析（图1-17）。

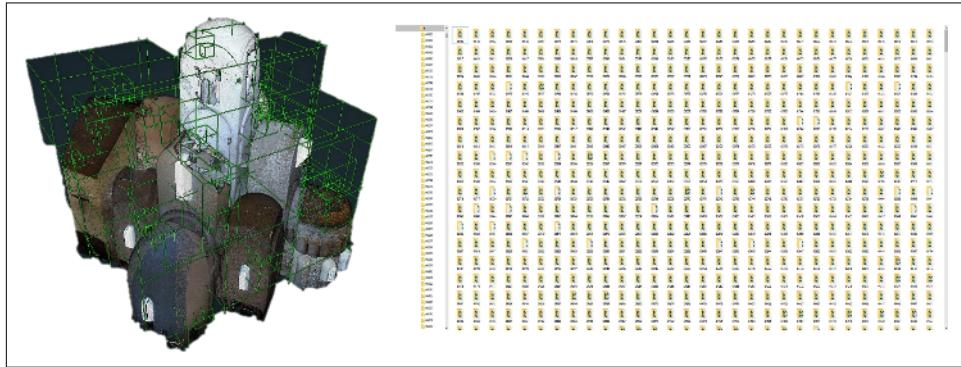


图 1-17. 基于 3D 点云构建八叉树的示例及其相关的外存储文件系统

确实，我们经常需要在我们提供的数据结构和其直接底层之间切换。当我们处理三维数据时，我们拥有一组几何体，需要通过带有提取语义指针的拓扑数据结构将它们连接起来。我们需要克服这个障碍，通常我们可以实现一个数据结构，能够连贯地组织几何、语义和拓扑这三大支柱，以适应某个应用或多个应用。

在此阶段，如果我们保持模块的线性进展，我们将拥有一个具有合理三维几何结构、语义和拓扑的数据集。因此，我们拥有了进行分析步骤所需的所有元素。



在处理3D几何时，我们经常借助加速结构，如k-d树、二叉树和八叉树，以加快处理速度。第四章将介绍如何使用Python和方法论来最佳地构建您的3D数据集。然后，第十章通过提供先进的3D建模技术来补充第四章的内容。

## 3D数据分析

一旦结构清晰，分析工作流程便开始发挥作用。在这里，目标是准确描述数据，理解趋势，并做出推论。这个阶段构成了整体工作流程的重要部分，因为它揭示了隐藏在大量数据中的洞见（图 1-18）。



图1-18。图书馆三维点云数据集分解为其主要组成部分的示例

例如，我们可以通过分析物体之间的空间关系、它们的几何形状及其语义含义，提取有价值的见解，从而更好地理解底层的3D数据。这些分析任务使我们能够识别模式、检测异常并预测未来趋势，这在各个领域都非常有用。

此外，通过利用3D几何和拓扑，我们可以执行传统非3D数据科学方法无法实现的复杂分析。最终，从3D数据中提取有用信息的能力为创新和解决问题开辟了新可能，使分析任务成为3D数据科学工作流程中的关键组成部分。



分析任务与应用紧密交织。然而，我们常常依赖一个标准的子工作流程，该流程利用模型进行推理和预测。由于这是我们在3D建模中遇到的主要挑战之一，我将在第8章详细介绍这一阶段。此外，我们通过第9章扩展了3D数据分  
析的应用，重点关注3D形状识别。

## 3D数据可视化

构建可视化和渲染流程有助于更深入地理解数据。可视化工作流程使人们更容易掌握复杂的信息，通过增强有效展示数据的能力。这是将其他工作流程模块联系在一起的重要组成部分。当你思考某件事时，你是否觉得自己会创建某种视觉表示？如果我们谈论稀有物种，比如独角鲸（Monodon monoceros），如果你已经拥有一些关于它的数据或知识点，你会立刻想象出这些被称为“海中独角兽”的生物，它们以长长的螺旋牙闻名。<sup>4</sup>

这个例子展示了数据可视化对我们来说是多么的直观。我们拥有强大的能力来深入理解并综合或呈现各种主题的视觉支持。图1-19 分解了分析、发现、展示、感知和交互，这些共同构成了数据可视化背后的科学。

---

<sup>4</sup> 我最近对独角鲸有点兴趣。我从未见过它们，但现在我知道我要去加拿大和格陵兰附近的北极水域，才有机会看到这些美丽的生物。我了解到它们是社交动物，经常成群结队地出现，并且拥有令人印象深刻的潜水能力。我的意思是，如果我在水下憋气1.5分钟我就很开心了；而独角鲸可以潜到最多1500米（4921英尺）的深度，并且能长时间潜水，通常超过25分钟。真是令人惊叹的力量！

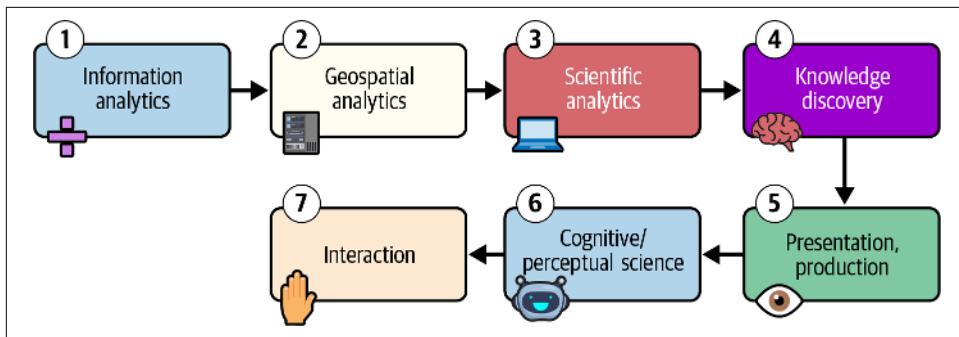


图 1-19. 分析数据可视化的主要组成部分

比我们的本能更强大的是处理代表真实物体的3D几何体和属性的能力，利用我们的现代计算系统。实际上，这构成了我们创建更好视觉效果的基本工具，可以更精确地实现我们设定的任何目标。

说到相关性，你听说过Jacques吗？不是比利时歌手Jacques Brel，也不是法国前总统Jacques Chirac，不是海军军官Jacques Cousteau，而是法国制图师Jacques Bertin（1918–2010）？

谈论数据可视化时，很难不提及Bertin的视觉变量。它们由Bertin提出，因为这些变量以视觉方式清晰且有效地编码信息，因此至今仍是以图形方式表示数据的基本属性。图1-20展示了这些变量，包括方向、大小、形状、颜色、纹理、明度和位置。

通过战略性地选择和组合视觉变量，我们可以创建有效的可视化，使观众更容易理解复杂的信息（对于制图和二维应用，这些可视化最常以地图的形式出现）。现在，是时候将这些模块与应用层联系起来了。



你将在第5章和第7章中找到一些关于你的3D可视化目标的灵感。第5章中，我们将创建一个多模态3D查看器，以处理各种3D模态。第7章中，我们将开发一些分析工具，如阈值滑块，用于实时与我们的数据集交互，并为3D体验带来沉浸式元素。

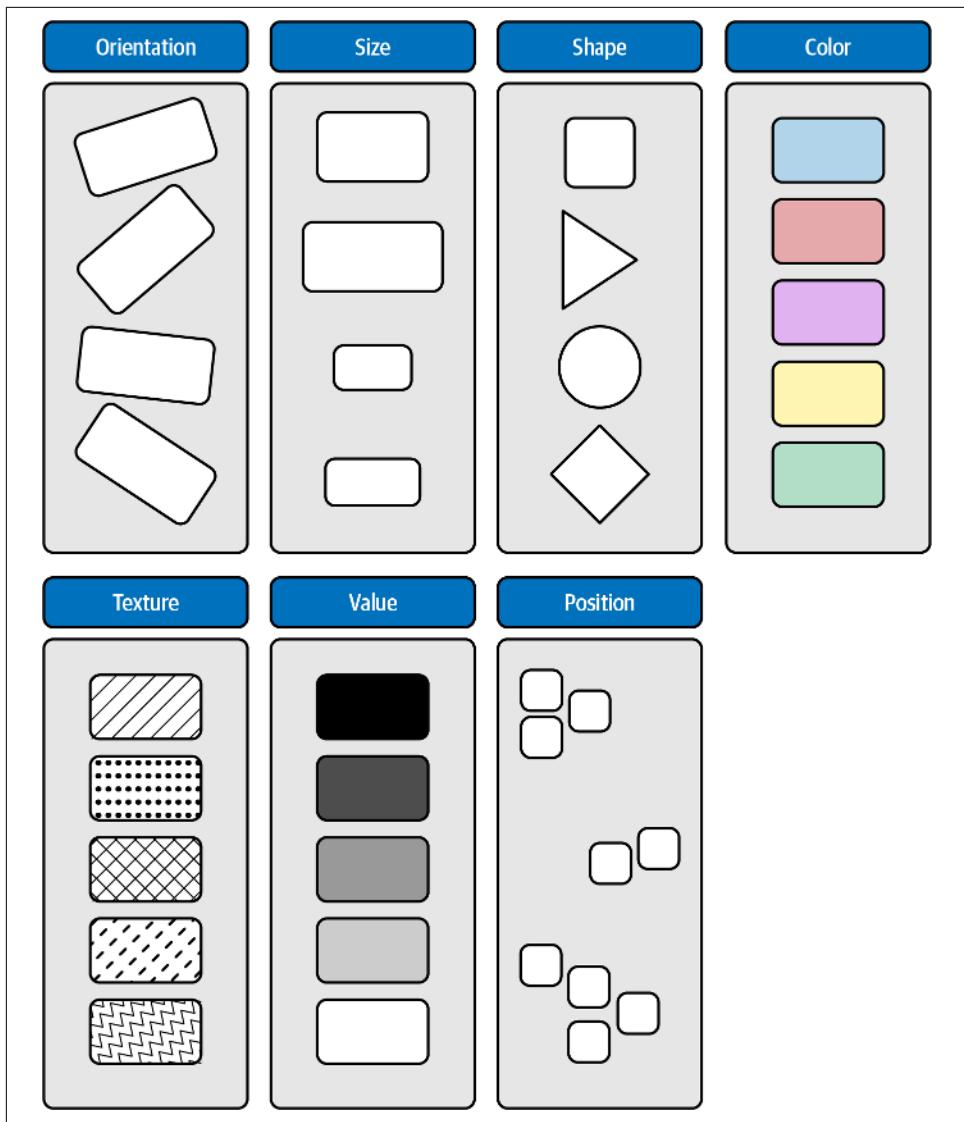


图1-20. Bertin的视觉变量

## 应用程序（软件）开发

这是一个可选步骤，重点是对我们开发的模型和算法进行细致的开发和部署。虽然本书并不打算为您提供完整的应用开发解决方案，但让我分享一下如图1-21所示的关键阶段的简单视图。

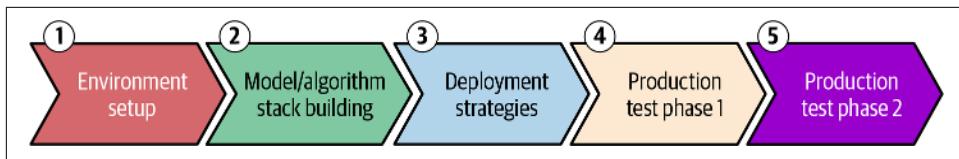


图 1-21. 创建交付生产就绪解决方案的应用模块的主要步骤

如您所见，到那时，您应该已经完成了阶段1——环境搭建，以及阶段2——模型/算法堆栈。这些内容已通过之前的步骤完成。然而，在本次生产周期中，我们通常会重构代码，确保其优化到尽可能顺畅运行。此外，还会进行兼容性检查，以实现与目标部署环境的无缝集成，无论是基于云的基础设施、边缘计算设备，还是现有系统。这意味着，尤其对于模型，我们必须考虑依赖关系，并构建一个能够无缝处理三维数据集的数据处理流程。这里的重点是优化——确保效率和处理能力。



想象一下，你正在开发一个用于建筑设计的3D计算机辅助设计（CAD）应用程序。对于这个示例，环境设置涉及安装和配置诸如OpenCASCADE用于几何建模以及PyTorch用于深度学习集成的软件库，并确保与建筑师和设计师使用的目标操作系统兼容。

这通常是我们决定与现实场景相一致的部署策略的地方。其核心思想是定义将软件更改从开发环境部署到生产环境的方法。这是你的软件架构定义的一部分。如果你打算将你的3D数据科学旅程做成一个真正的产品，我推荐你阅读 Mark Richards 和 Neal Ford（O'Reilly 出版）的《软件架构基础》。



如果您打算将计算机代码解决方案转化为有形且可销售的产品，那么您可以选择多种产品化方法。您可以创建独立应用程序、软件即服务（SaaS）、嵌入式软件、软件组件，甚至提供专业服务。创造力不仅止步于数据背后的科学。然而，产品的选择自然会影响软件的架构。

如今，您可以受益于云部署，它提供了可扩展性和可访问性，可能帮助您尽量减少停机时间，降低风险，并确保实施成功且高效。让我们考虑另一个例子。在 the

地理空间分析领域，利用亚马逊网络服务（Amazon Web Services）、微软 Azure（Microsoft Azure）或谷歌云平台（Google Cloud Platform）等云服务，确保了全球用户在分析和可视化大规模三维空间数据集时的**理论解耦而模块化**，仍有一些重要阶段：严格的生产测试对于验证部署的3D数据科学应用的稳健性和可扩展性至关重要。这包括性能测试以评估响应时间和资源利用率、压力测试以评估峰值负载处理能力，以及安全测试以增强对漏洞的防护。这是第一阶段，主要涉及内部测试。

想象一下部署一个用于肿瘤检测的3D医学成像应用程序。在生产测试阶段，性能测试评估该应用程序高效处理高分辨率3D医学图像的能力。压力测试评估系统如何处理来自多位医疗专业人员的同时请求，而安全测试则确保患者数据的机密性。

然而，仍然缺少一个重要部分，这部分由我称之为生产测试第二阶段来处理。它通常包括一个试点阶段，启动对有限用户群的推广并收集初步反馈和见解。随后进行分阶段部署，逐步扩大到更大用户群。

这个特定的过程是一个“独立的世界”。它在很大程度上扩展了本书的范围，但当你准备将你的想法转化为产品时，它也非常有趣。因此，我可以推荐一系列七本书，逐步引导你完成实现优秀软件产品的过程：

- 《实用程序员》 David Thomas 和 Andrew Hunt 著（AddisonWesley Longman）
- Modern Software Engineering by  
David Farley(Pearson Education)
- Code Complete: 由 Steve McConnell 著作的《软件构建实用手册》（Microsoft Press）
- Google的软件工程 由 Titus Winters、Tom Mansreck 和 Hyrum Wright (O'Reilly) 撰写
- 《软件架构基础》作者 Mark Richards 和 Neal Ford (O'Reilly)
- 软件架构：艰难之处 作者：Neal Ford、Mark Richards、Pramod Sadalage 和 Zhamak Dehghani (O'Reilly)
- 《设计机器学习系统》作者：Chip Huyen (O'Reilly)

在这个阶段，你应该对3D数据科学工作流程背后的内容有了更清晰的认识。理解这些过程为数据采集、预处理、配准、语义提取、数据集优化、分析、可视化以及针对特定应用的定制奠定了基础，从而推动革命性进展的实现。

但理解基础知识并不止于此。在3D数据科学的世界里，自动化至关重要。虽然实现这些工作流程的自动化并非易事，但达到一定程度的自动化能够带来巨大的益处。让我们接着来探讨这一点。

## 自动化的理由

简化我们团队在生产中的任务和工作流程可以大大提升项目执行效率。事实上，有效的工作流程管理使我们能够构建、自动化并管理工作流程，从而清晰地了解整个过程。它有助于消除混乱和无序，简化流程。其中一个尤其关键的方面是：自动化。自动化任务带来了诸多益处，如图1-22所示。

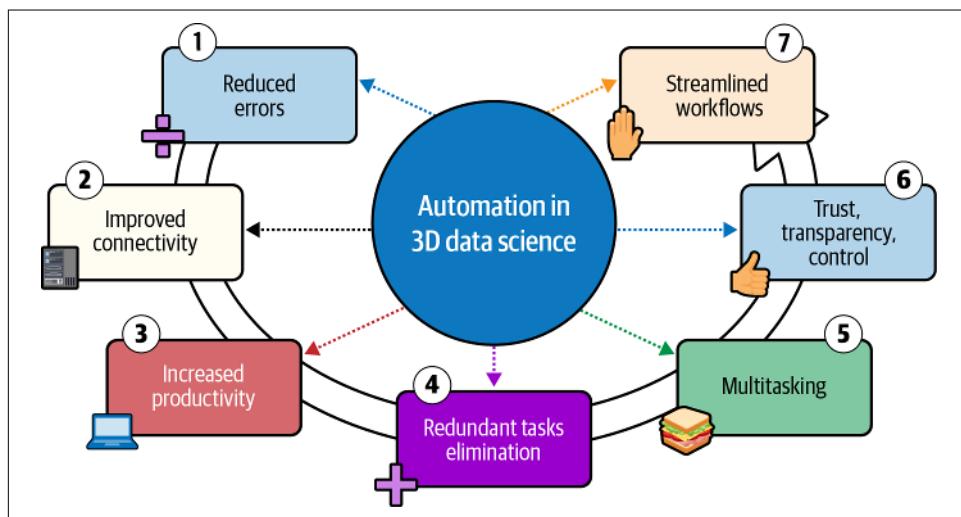


图 1-22. 3D 数据科学中自动化的好处

与3D数据科学相关，我们可以将思考从以人为中心的流程扩展到自动化工作流程。这将我们的目标定位于开发自动化和空间AI，以加快推理过程的速度。这在开发包含语义化任务的工作流程时尤为关键，例如需要识别物体的场景。

当我们从行业的角度来看时，我们看到机器人研究在提供自主3D记录系统方面取得了显著进展。这些系统能够在无人干预的情况下捕捉环境的3D点云。然而，仅仅收集没有上下文的数据集是不足以构成有效的

决策。我们需要专家的知识来提取必要的信息，并为决策创建可行的数据支持。

自动化全自主认知决策系统的过程是一个极具吸引力的前景。然而，这带来了重大的挑战，主要涉及3D数据的知识提取、整合和表示。因此，3D数据集的结构设计必须使计算机能够将其作为通过推理和基于智能体系统进行信息提取的基础。这为理解3D数据科学如何依赖于3D数据专业知识铺平了道路，而这又需要熟练的知识管理。实现这一点是自动化工作流程的最终阶段（图1-23）。

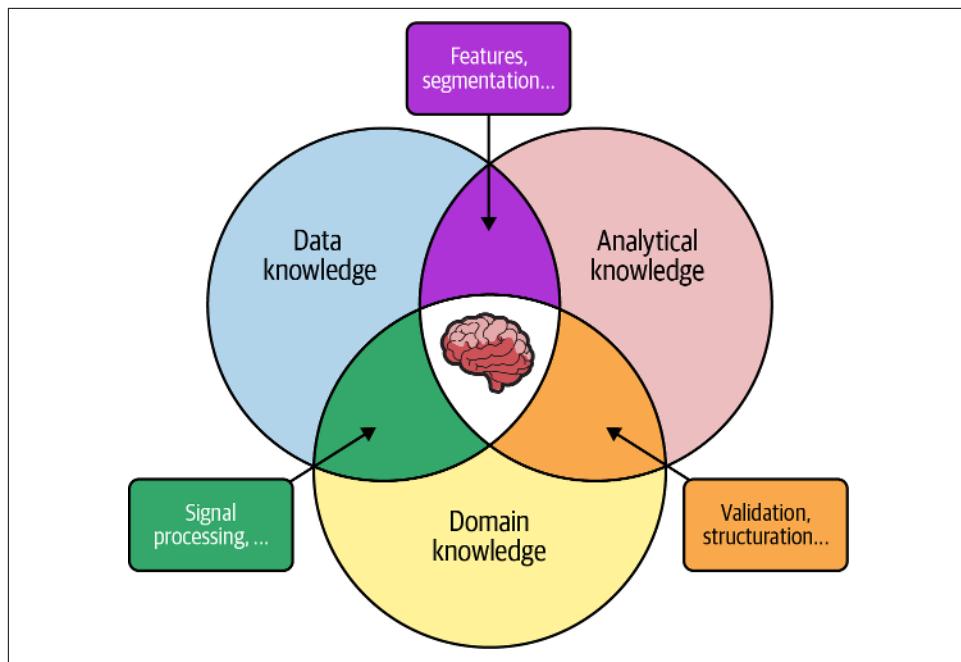


图1-23. 用于实现100%自动化和基于3D数据科学推理的知识组织

设想如果我们试图通过理解采集过程的特性及影响该过程的不同因素，来理解数据特性的起源。这样的理解使得能够预测数据特性。这个以数据为中心的过程需要对数据、场景和数据处理领域及这些领域之间的影响有明确的认识。因此，我们确定了三个主要领域：

## 数据知识

数据域由有关数据特征的知识和有关获取过程的知识（例如，获取方法、获取技术、获取工具）组成。

## 领域知识

数字化场景的领域由关于物体的知识、它们在场景中的分布、场景上下文以及外部因素组成。

## 分析性知识

最后，我们需要了解工作流程中使用的算法（例如，使用条件、算法的目的、配置、优选使用条件）。

如果我们在设计系统时考虑到这些因素，就可以确保实现全自动化不会导致我们模块的恶意逆向工程。

现在我们已经排列好了各种工作流程模块以及自动化的案例，我们是不是在一条直截了当的道路上？实际上，有相当多的挑战你应该了解，以更好地掌握前方的道路。

## 3D数据科学中的工作流程挑战

我们对现实世界的应用感兴趣，因此理解我们将要面对的挑战非常重要。我们应该注意什么？3D数据科学有哪些局限性？让我们以一种综合的方式来探索我为你整理的内容。

基本上，每当我们有一个现实世界的数据科学工作流程时，我们总是会有三个主要阶段：数据准备、构建模型以解决特定任务，以及推断和分析。考虑到这些步骤，我总结了我们将面临的八个主要挑战（图1-24）。

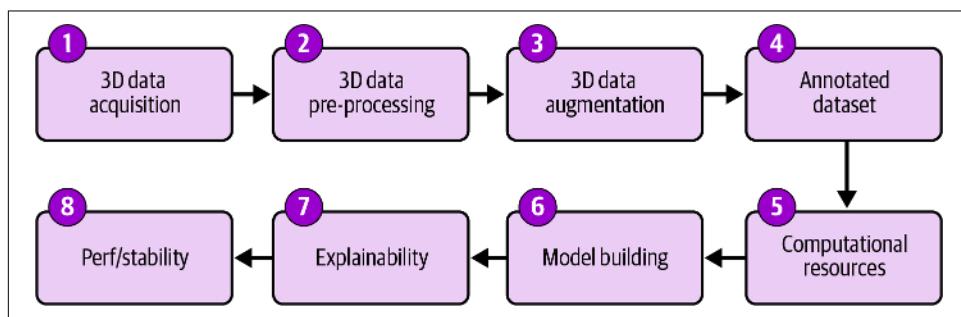


图1-24。三维数据科学中的主要工作流程挑战

这些挑战没有特定的顺序，但我按照图1-12中展示的主要工作流程，以一种逻辑方式将它们连接起来。

## 3D 数据采集的挑战

理解数据的来源非常重要，尤其是在你尝试构建架构时，这样你才能确定它将针对你的应用进行调优。请记住，数据科学的这一应用非常专业且在特定的实际应用中表现卓越。在这种情况下，你有不同的传感器，例如LiDAR传感器、主动传感器和被动传感器。你还有不同的平台，可能是移动的或静态的。你也有其他的采集方法，这使得数据收集的过程不那么直接，因此在使用数据之前，最好明确数据是如何收集的。

## 3D数据预处理

在将数据集输入到架构之前，几乎总是需要对数据进行预处理以确保其能够正常工作，如前所示。预处理可以简单到仅仅是平移、旋转数据，将其归一化到特定范围，或者进行平铺和采样，以确保始终有相同数量的数据点通过网络。这些都是针对3D深度学习应用预处理3D数据的一般示例。这个阶段带来了一些挑战，即我们应该进行哪些预处理以及如何进行。

## 3D数据增强

通常，我们不会拥有足够数量的带注释数据集，这意味着为了确保涵盖应用中的全部变异范围，我们需要进行一些数据增强。这可能包括噪声注入、数据合成，或其他对数据执行旋转、平移或缩放的方法，所有这些都是为了扩展数据集的可能配置，例如。

## 缺乏带注释的数据集

3D 数据科学中最大的挑战之一是对标注数据集的需求。事实上，我们经常依赖于从标注数据中学习的模型。要以监督方式训练这样的模型，您需要带有标签的数据。处理 3D 数据时，这一点尤为关键。可用的既包含数据又包含语义（标签）的标注 3D 场景非常少，且针对具体应用找到这样的数据更加困难。这使得使用监督学习技术准确训练模型变得具有挑战性。

让我们举个例子。我曾经从事文化遗产项目。在这类工作中，拥有一个你在现场收集的数据集意味着你通常找不到该特定对象的另一个样本。你需要提供另一层关于你想要分类的数据的包容性类的抽象。所有应用都需要找到足够多的每个类别的对象，以充分标注这些分割工作流程。通常这归结为拥有大量数据，而这并不

真的有可能。这是一个大问题。当尝试提供良好注释的多模态数据集时，这种复杂性会进一步加剧。有一句著名的话：“垃圾进，垃圾出。”当你想在生产环境中推送某些东西时，你需要确保用于训练的数据具有顶级的标签质量。通常，这意味着需要大量的人为监督和有限的无监督机制，以尝试促进标签过程，因为这可能非常耗时。这是一个巨大的挑战。

## 计算资源

这也是一个巨大的挑战。我们正在处理庞大的数据集，这将促使我们需要进入深度学习领域，而不是传统的机器学习。然而，如果你的数据集非常小，机器学习通常能够表现良好，前提是你只是获得了最佳结果，而不是交集的并集。这是第一个挑战。然后，还有另一个挑战：你手头的计算资源。

我们处理极其复杂的任务。我们需要以最高效的方式执行许多连续操作。通常，不建议在中央处理单元（CPU）上训练深度学习模型以用于生产环境的应用（尽管在CPU上进行推理是可能的）。对于训练来说，我们需要谨慎确保我们的计算在GPU（图形处理单元）上执行，因为GPU更加注重性能，或者在TPU（张量处理单元）上执行。这意味着我们通常需要转向云计算或并行计算。如果你使用带有CUDA的GPU，我们稍后会涉及这一点。在这种情况下，通常你的所有并行计算都会被处理好，使得利用所有可用的CUDA核心更加高效。

## 模型构建

模型构建是一个重大挑战，因为首先，我们需要确定将选择哪种范式以及使用哪种机制。然后，我们必须选择一个具体任务，接着选择一个具体方法来实现我们想要的功能。然后，我们需要确保定义训练、推理以及如何以合适的比例进行泛化，以获得稳健的性能。这还意味着需要监控性能，确保我们获得尽可能高的性能。这是最大的挑战。

## 可解释性

可解释性和可说明性是什么意思？模型有多个层次，模型会进行大量复杂计算以找到执行特定任务的最优参数，这意味着通常很难准确理解模型工作的规则或逻辑。这是你在处理人工神经网络时经常会发现的情况。很有趣

在深度学习、可解释性和可解释性人工智能方面正在进行相关工作，这些内容将在第8章和第10章中介绍。

## 性能与稳定性

最后一个挑战是性能和稳定性。你希望你的模型在特定任务或多个任务上拥有最佳性能，同时你也希望具备稳定性，这意味着无论是在训练期间、推理期间，还是两者兼有时，你都希望模型能智能且稳健地表现。对于训练来说，这意味着你希望每个epoch中的变异数量是可控的。对于推理而言，你的目标是拥有一个能够很好地泛化的模型。确保你以一种真正代表你实际应用需求的方式准备数据。

请记住，我们不是从学术角度或研究角度来看待问题，在这些角度下，构建一个没有直接应用的理论是可以接受的。我们希望解决现实世界中的实用问题。这就是为什么在解决手头的问题时，我们需要确保所有这些挑战都得到解决。当你思考3D数据科学时，牢记这八点非常重要。它将指导你贯穿整本书及更远的学习，确保每当你构建某样东西时，都能正确应对所有这些挑战。

表1-1总结了3D数据科学中的工作流程挑战，应有助于指导您未来的开发工作。

表 1-1. 3D 数据科学中的工作流程挑战

3D 数据科学中的工作流程挑战		
3D 数据 收购 挑战	理解……的起源和特征 3D数据尤为重要，尤其是考虑像传感器类型（LiDAR、 平台（移动，静态），和收购 方法论。	仔细考虑传感器特性、平台 数据期间的局限性和获取方法 收集。传感器、平台和选择 方法论应与应用保持一致 需求。
3D 数据 预处理	这一步确保数据质量并且 与后续处理的兼容性 步骤。关键操作包括翻译， 旋转、归一化、平铺和采样。	根据情况应用适当的预处理技术 数据特征和后续的要求 处理步骤。技术如噪声消除， 离群值移除和数据归一化增强数据 质量和模型性能。
3D 数据 增强	这一步有助于解决……的挑战 有限的带注释数据集。它 像噪声注入、合成数据这样的技术 生成、旋转、平移和平移到 增加数据的多样性并提升模型 泛化。	利用各种数据增强技术进行扩展 训练数据集并提高模型的鲁棒性。 增强技术，如旋转、缩放，以及 添加噪声，创建现有数据的变体， 减少过拟合风险并提升模型 泛化。

挑战与策略		
缺乏注释的数据集	获得足够的带标签的3D数据监督学习是一个重大挑战。这对于专业领域尤其如此应用中带注释的数据集是稀缺。	探索半监督或其他替代方法当标记数据有限时的无监督学习。利用迁移学习技术，使用在相关数据集上预训练的模型用于引导训练过程。投入数据标注工作以为特定领域创建高质量的标注应用程序。
计算的资源	3D深度学习通常需要大量计算资源。训练大型模型在庞大数据集上需要强大的硬件如GPUs或TPUs，需要云计算解决方案以满足高效处理。	优化代码和算法以实现高效计算，考虑并行处理和GPU利用率。探索提供可扩展性的云计算平台按需提供计算资源，使得大型3D模型的训练和部署。
模型构建	选择合适的模型架构，训练方法论和表现评估指标是至关重要的。平衡模型复杂性与泛化能力之间的权衡同时确保稳健的训练和推断过程是一个关键挑战。	根据仔细选择模型架构特定的3D数据分析任务和特性。尝试不同的训练策略、超参数调优和正则化实现模型最佳性能的技术。持续监控训练进度并进行验证模型在独立数据集上的表现以防止防止过拟合并确保泛化能力。
可解释性	理解3D深度背后的推理学习模型的预测可能具有挑战性。解释复杂模型的决策和提供对预测的清晰解释是建立信任和确保的关键因素负责任的人工智能开发。	调查解释性技术，例如特征可视化或注意力图以提供见解模型决策制定。开发方法用于生成可供人类理解的模型解释预测，提升透明度和问责制。
性能和稳定性	在实现强大模型性能的同时在训练过程中保持稳定性以及推理至关重要。管理训练变异性并确保一致性、可靠性在现实世界应用中的模型预测是一个持续的挑战。	实施具有适当性的稳健培训程序正则化技术和学习率调度以减轻训练不稳定性。彻底评估模型在各种数据集上的表现以确保稳定性和在现实世界场景中的可靠预测。

## 工业中的3D数据科学

在结束本章之前，我想为您提供一个关于3D数据科学实际应用的可选视角（图1-1中的第4阶段）。您可以在3D数据科学资源中心找到一个更详细的视图，该视图描绘了大多数您在处理3D数据科学时可能遇到的应用所涵盖的主要领域。

每个领域中都有许多任务，由3D工程、3D分析、3D重建和3D可视化中的角色支持。然后，对于每个这些任务和角色，您可以围绕各种知识支柱定义一个工作流程。以这种方式构建，一个清晰的

路径出现，使我们能够了解如何进入这些不同的类别（图1-25）。

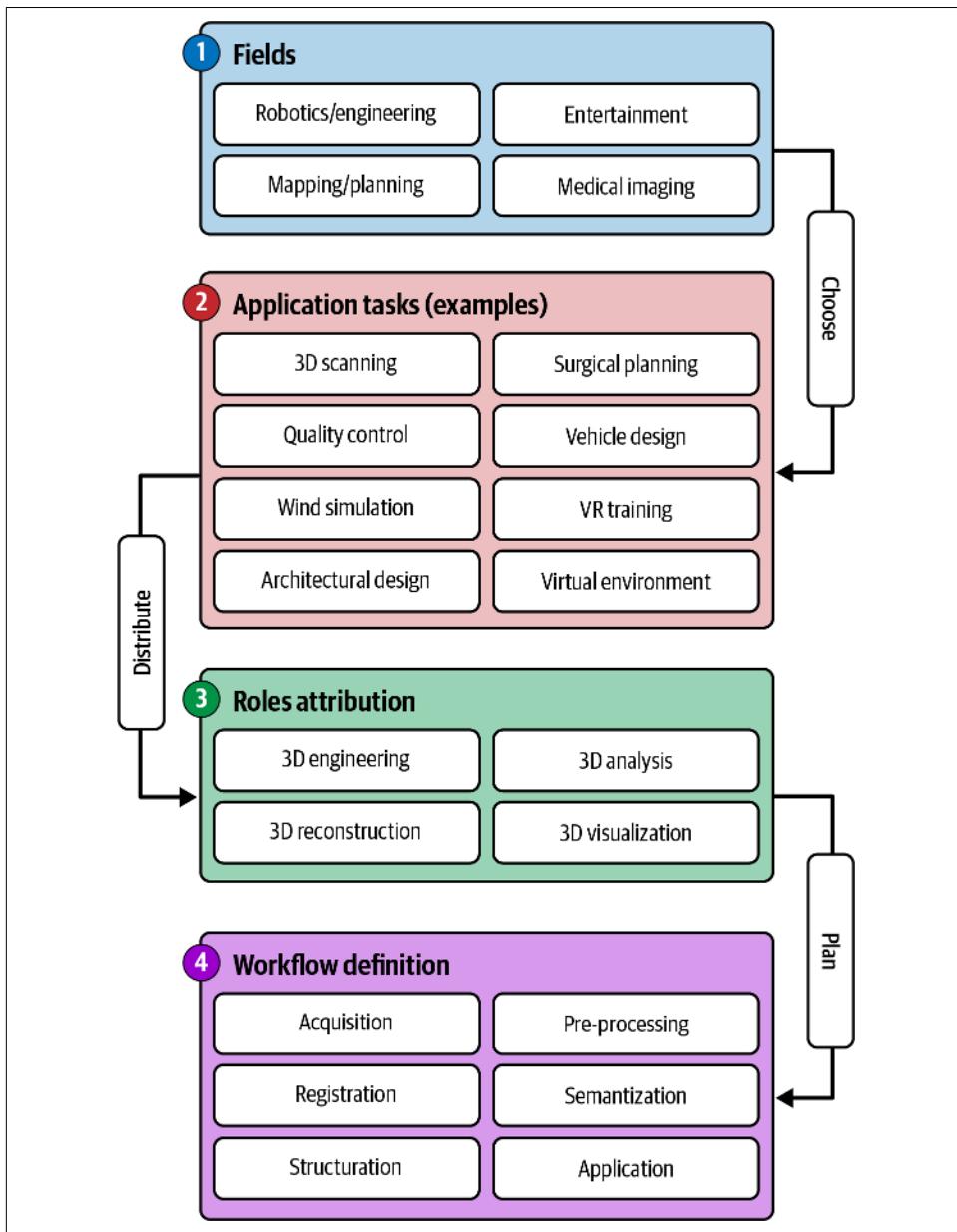


图 1-25. 行业中的3D数据科学

识别领域、应用和角色可以让您了解在定义工作流程之前所需的流程和目标定位。这样的工作流程可以通过对当前行业参与者的仔细观察和研究策划得出。



除了学术界之外，3D数据科学领域还有一批充满活力的行业领袖，如NVIDIA、Microsoft、Google、Amazon和Facebook。它们通过复杂的解决方案推动了创新。这些以技术实力著称的公司，对推进3D数据科学应用做出了重要贡献。

## 摘要

到此为止，您已经迈出了进入3D数据科学的第一步。我们研究了维度在科学中的作用，以及空间人工智能如何以几何体作为基础，模拟我们大脑处理3D数据时的分析部分。在继续下一章之前，有九个关键学习点需要理解：

- 3D数据科学分析三维数据，以提取知识和见解。它弥合了我们三维世界与数据科学之间的鸿沟。
- 空间人工智能是人工智能的一个专门领域，专注于理解和与三维空间中的世界进行交互。它是最佳三维数据科学工作流程的成果。
- 维度是某物的可测量方面或特征。在3D数据科学中，维度通常指空间范围。
- 3D数据科学提供了许多机会，例如低入门成本、高多功能性以及一个面向未来的职业路径。
- 几何学、拓扑学和语义学对于从三维数据中提取有意义的信息至关重要。几何学包括对空间属性的研究。拓扑学关注空间的内在属性。语义学指的是与元素或结构相关的意义、解释或重要性。
- 点云数据是3D数据科学中的一个核心数据模态。点云是表示记录环境的离散空间集合。
- 3D数据科学模块化工作流程包括数据采集、预处理、配准、3D数据分类、结构化、3D建模、3D数据分析、3D数据可视化和应用开发。

- 自动化3D数据科学工作流程对于提高效率和可扩展性至关重要。
- 3D数据科学中的工作流程挑战包括3D数据采集、3D数据预处理、3D数据增强、缺乏带注释的数据集、计算资源、模型构建、可解释性、性能和稳定性。

本章中确立的工作流程定义构成了本书中我们努力的基础。接下来，我们将探讨帮助你开启3D数据科学之旅的基本资源和工具。

## 动手项目

随着你阅读本书，你会发现每章都包含一个练习，你可以通过完成这个小型实践项目来自我评估应用所学知识的能力。对于本章，请熟悉点云数据，这是3D数据科学的核心组成部分。你可以通过以下六个步骤完成：

1. OpenTopography 提供各种 3D 点云数据集。访问网站并选择一个您感兴趣的数据集。考虑位置、大小和点密度等因素。
  2. 下载所选数据集的一小部分。
  3. 可视化点云：有许多工具可用于点云可视化，我们将在下一章中介绍。作为第一步，您可以使用OpenTopography生成的查看器来探索您的点云，以避免在此阶段安装软件。
  4. 观察点云并尝试识别不同的特征或物体。你能辨别出哪些形状或模式？你能对点云中捕捉到的环境或物体做出任何推断吗？
  5. 当你处理小规模点云子集时，思考一下五个V（体量、速度、价值、真实性和多样性）在处理大规模点云数据时可能带来的挑战。
  6. 考虑可能需要的预处理步骤，以清理和准备该点云数据，以便进行进一步分析或模型训练。
- 本练习侧重于使用外部工具和资源对点云数据进行动手探索。完成该练习大约需要30到45分钟。