

目录

第一章 智能制造大数据概述.....	2
1.1 工业制造过程概述.....	2
1.1.1 工业制造过程的基本概念.....	2
1.1.2 制造业发展阶段.....	4
1.1.3 制造业发展战略.....	6
1.1.4 制造业向智能制造转型升级.....	9
1.2 智能制造的基本概念.....	10
1.2.1 智能制造的定义.....	10
1.2.2 智能制造的发展历程.....	11
1.2.3 智能制造的发展方向.....	12
1.3 工业大数据的基本概念.....	13
1.3.1 工业大数据的定义.....	13
1.3.2 工业大数据的来源与类型.....	14
1.3.3 工业大数据的特征.....	16
1.3.4 工业大数据的采集.....	17
1.4 智能制造大数据分析需求与流程.....	20
1.4.1 智能制造大数据的典型分析需求.....	20
1.4.2 智能制造大数据的一般分析流程.....	21
1.5 本章小结.....	23
习题.....	24
参考文献.....	24

第一章 智能制造大数据概述

制造业是国民经济的主体，是立国之本、兴国之器、强国之基。进入新时代，加快建设制造强国、加快发展先进制造业，成为我国的国家战略。新一代人工智能与先进制造深度融合形成的智能制造技术，正在成为新一轮工业革命的核心驱动力。而大数据技术正在给传统制造业发展方式带来颠覆性、革命性的影响，对提升制造业智能化水平具有重要意义。本书以智能制造为载体，以数据处理技术为主线，介绍智能制造中的大数据分析技术与应用。

本章从工业制造过程的基本概念出发，结合制造业发展历程与世界制造业发展格局，分析推进制造业向智能制造转型升级的迫切需求。接着，从智能制造的科学问题出发，分析大数据技术对于智能制造的重要意义。最后，讲述工业大数据的基本概念，以及智能制造中的数据分析需求与分析流程。

1.1 工业制造过程概述

自十八世纪中叶开启工业文明以来，世界强国的兴衰史和中华民族的奋斗史一再证明，没有强大的制造业，就没有国家和民族的强盛。打造具有国际竞争力的制造业，是我国提升综合国力、保障国家安全、建设世界强国的必由之路。本节主要概述工业制造过程的基本概念、发展历程与发展战略。

1.1.1 工业制造过程的基本概念

从原始时代开始，人们就开始制造和使用工具。从石器时代到青铜器、铁器时代这段时期，生产力水平低下，制造技术主要依靠手工技艺，生产方式采用的是手工或半手工半机械化的手工业作坊式生产方式。随着自动化技术、信息技术、先进制造和管理技术的进步以及生产力的发展，人们对制造过程的定义和内涵的理解发生了较大的变化，逐渐形成了制造过程的基本概念。

（1）工业制造过程的定义

工业制造过程是指对制造资源（物料、能源、设备、工具、资金、技术、信息和人力等），按照市场要求，通过采购、加工、装配、检验、销售等过程，转化为可供人们使用的工业品与生活消费品的过程，通过改变原料的形状或特性来增加原料的价值。其中，由产品的加工、装配组成产品的基本制造过程属于狭义制造概念，典型工业制造过程可以一般化表示为图 1-1 中的过程。

工业制造过程的各个环节都会产生数据，能够反映设备状态、生产运行情况、成本消耗和资源调度等多方面信息。数据分析与信息处理是指对制造数据的采集、分析、处理、存储和分析等，涉及自上而下的生产指令和自下而上的反馈信息。数据分析与信息处理能够指导工厂进行市场分析、决策规划、产品研发、工艺规划、管理控制，对于保证产品质量、提升生产效率、增加企业效

益等都至关重要。

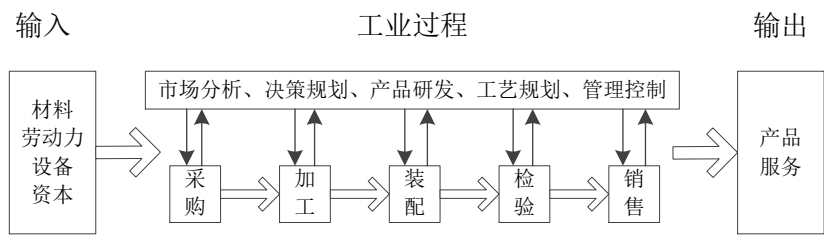


图 1-1 工业制造过程示意图

(2) 工业制造过程分类

根据生产的工艺流程不同，制造型生产过程可以分为离散工业和流程工业。其中，离散工业主要代表为机械装备制造，流程工业主要代表包括石化、冶金等原材料工业和火力发电、核能发电等电力能源工业。

离散工业与流程工业存在明显不同，具体如图 1-2 所示。离散工业的主要制造过程可以概括为制造装备总体设计、零件加工、装备组装，是由多个零件经过一系列不连续的工序加工、装配形成最终产品。其零件加工与组装是可拆分的物理过程，可以通过计算机集成制造技术实现数字化设计与生产。

流程工业是以原材料为主产品，原料进入生产线的不同装备，通过物理化学反应，以及进一步的形变、相变过程，在信息流与能源流的作用下，将原材料变化形成合格的产品。在这个过程中，工艺和产品较为固定，产品不能单件计量，产品加工过程不能分割，生产线的某一工序产品加工出现问题，会影响生产线的最终产品。

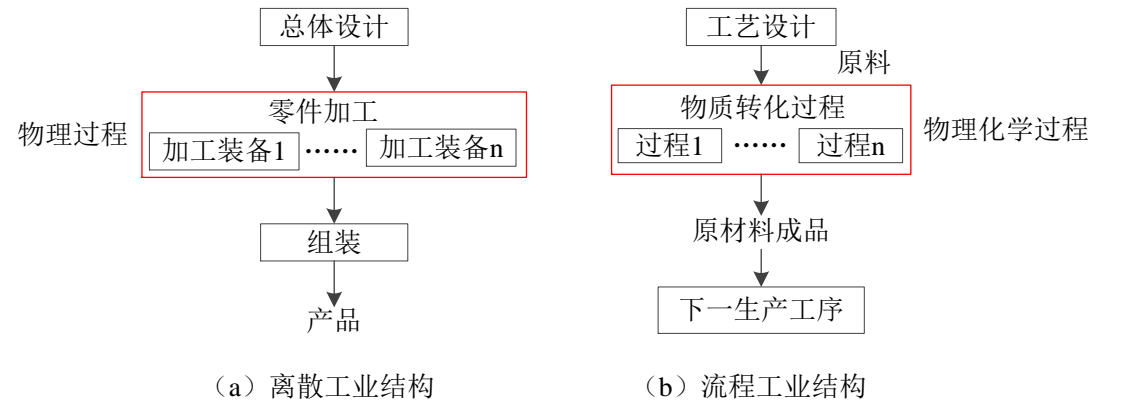


图 1-2 离散工业和流程工业的结构

下面以钢铁冶金过程为例，对上述离散工业和流程工业在实际工业过程中做具体介绍。

如图 1-3 所示，一个典型的钢铁冶金过程包括高炉炼铁、炼钢、连铸和轧钢四个部分。高炉炼铁是把铁矿石还原成生铁连续生产过程，将铁矿石、石灰石、炼焦煤按照一定比例投入到高炉中得到铁水。炼钢过程将生铁里的碳及其它杂质（如硅、锰）等氧化，产出比铁的物理、化学性能与力学性能更好的钢；转炉生产出来的钢水经过精炼炉精炼以后，需要将钢水铸造成不同类型、

不同规格的钢坯，连铸工段就是将精炼后的钢水连续铸造成钢坯的生产工序。轧钢过程是将连铸后的钢坯轧制成客户需要规格的钢材，主要包括冷轧和热轧，得到冷轧卷和热轧卷。这里的炼铁和炼钢过程都属于流程工业，而连铸和轧钢都具有典型的离散工业特征。

除此之外，为了保障钢铁厂的正常运行，需要装备制造将零部件组装成为大型制造机器，为钢铁厂提供制造装备，这属于离散工业；电厂将煤炭通过一系列处理转化为电能，为钢铁厂提供电力，此部分属于流程工业；还需要矿场为其提供各种矿石原料，此部分包含采矿、选矿、配矿等工序，既涉及具有离散和间歇特征的离散生产过程，也涉及连续作业的流程生产过程。

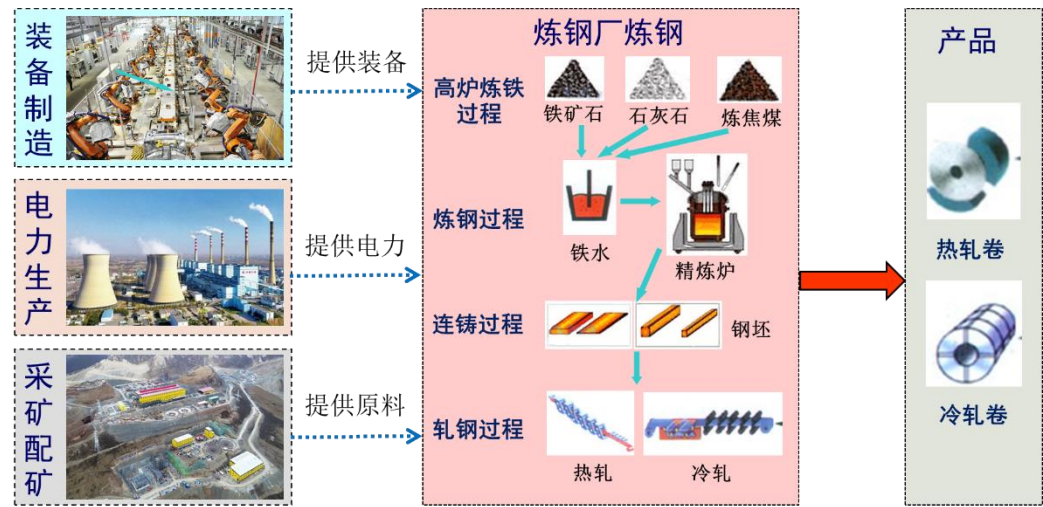


图 1-3 钢铁冶金过程示意图

1.1.2 制造业发展阶段

从世界经济发展史的角度看，世界的生产方式经历了四次重大的变革，即四次工业革命，分为蒸汽动力时代、电力时代、信息时代和智能互联时代。每个阶段的时间、代表性技术及产品、产业密集类型、劳动力手段、生产工具在表 1.1 中做简要概括，每一阶段重要的事件与技术特点具体介绍如下。

表 1.1 世界制造业发展史

时间	蒸汽动力时代	电力时代	信息时代				智能互联时代
	1760-1780	1860-1940	1940-1970	1970-1985	1985-1995	1995-2012	2012 至今
代表性技术及产品	蒸汽机	电动机	集成电路、PLC	个人计算机	网络和通信技术、互联网	现代集成制造、网络化制造	移动互联网、大数据、云计算、人工智能等新一代信息技术
产业密	资本密集	技术密集	技术	技术-信	信息密	信息-知	知识密集

集类型			密集	息密集	集	识密集	
劳动力手段	机械化	刚性自动化	刚-柔性自动化	柔性自动化	集成化	集成化、智能化	智能化、预测性
生产工具	机器设备	联动线	机电自动线	柔性生产线	制造系统	制造系统	制造平台

(1) 第一次工业革命——蒸汽动力时代

1776 年，英国机械师瓦特发明了第一批新型蒸汽机，并应用于实际生产。依靠蒸汽动力，纺织业、机器制造业取得了革命性的变化，汽车、火车、汽船等交通工具相继生产出来，煤炭、石油和钢铁等行业也得到迅速发展。蒸汽机的发明，将人类社会带入了蒸汽动力时代，蒸汽机成了大工业时代普遍应用的发动机，引发了第一次工业大革命。第一次工业革命标志着由机器占统治地位的近代工业化大生产时期的开始，为资本主义发展和机器大生产创造了广阔的前景。

(2) 第二次工业革命——电动力时代

1820 年，丹麦物理学家奥斯特发现了电磁效应，法国物理学家安培提出了电流相互作用定律；1821 年，英国化学家法拉第制作出了永磁电动机模型；1831 年，法拉第总结出电磁感应定律，发明了发电机；1864 年，英国物理学家麦克斯韦建立了电磁场理论。上述这些发明，发现为后来的电信、电器以及电动机、发电机等电力工业的应用奠定了理论基础。此后，电话、电报、电灯相继问世。电力机器与蒸汽机相比能耗低，且清洁、操作方便、占用空间小、性能优良、价格低廉，具有明显的竞争优势。1875—1894 年，法国试验应用电力，直接采用电动机驱动机床，推动了机床制造业的发展，结束了机床动力由人力、水力到蒸汽动力的时代。以电作为动力改变了机器的结构，开拓了机电制造技术的新局面，人类从此开始进入电动力时代。

(3) 第三次工业革命——信息时代

信息技术的发展主要以计算机出现为标志，信息经济时代大致可以分为数据处理时代、个人计算机时代、互联网时代、移动互联网时代和智能互联时代 5 个阶段。

第一阶段开始于第二次世界大战期间，军事工业发展的需要促使电子技术的研究与开发异常活跃。美国陆军支持宾夕法尼亚大学大学研制电子计算机，第一台电子计算机于 1945 年诞生；1958 年贝尔实验室成功研制出集成电路；1971 年出现了单片微处理机，之后在美国诞生了超大规模集成电路，电子计算机进入到超大规模集成电路为标志的第四代；1969 年，美国莫迪康公司（Modicom）开发了第一台可编程逻辑控制器（Programmable Logic Controller, PLC），PLC 把智能融入机器和过程自动化，广泛应用于工业、基础设施和建筑

业，标志着数据处理时代的到来。

第二阶段技术上的特征是源于 1981 年 IBM 的第一台桌上个人计算机的出现与发展，从而使信息处理技术发生了革命性变化。但处理的范围还局限于单机，而且所能处理的信息也很简单，主要是文字和二维图形。

第三阶段技术上的特征是网络和通信技术使得信息跨地域的迅速流通和共享得以实现。1995 年 10 月 Netscape 上市，成为互联网商业化的标志。在该阶段，信息主要通过门户网站、搜索引擎和社区进行单向传播。随着各国信息高速公路的建设，迎来了信息化进程中的网络时代。这个阶段的另一特征是信息处理等能力大大提高，能处理包括图形、图像、声音等多媒体信息。

第四阶段开始于 2004 年，以 Web 2.0 为标记；2009 年，智能手机的普及标志着移动互联网进入快速发展阶段。在该阶段，信息连接方式以 Wi-Fi、3G、4G 为主，信息传播方式主要通过社交网络平台、App、自媒体等实现双向互动。

第五阶段开始于 2012 年美国通用公司提出的工业互联网概念，随着 2012 年美国再工业化战略、2013 年德国“工业 4.0”制造战略、2015 年“中国制造 2025”战略的推出，信息技术的发展进入智能互联时代。该阶段的连接方式主要有移动 Wi-Fi 和 5G 等，并以物联网、机器人、智能硬件和人工智能为手段，以智能设备为中心出发点，实现整个网络世界之间的全方位智能互联、互动。

(4) 第四次工业革命——智能互联时代

随着移动互联网、大数据、人工智能、物联网、云计算等新一代信息技术的普及，带动了几乎所有领域发生了以智能化、服务化、网络化为特征的群体性技术革命。新一代信息技术与制造业深度融合，孕育了智能制造的新理念。在智能互联时代，人们工作与生活中的设备与设备、人与设备之间基于物联网和移动互联网实现互联，基于大数据与人工智能实现智能化，标志着以智能互联为特征的第四次工业革命的到来。

在智能互联时代，新一代信息技术与制造业深度融合，已经开始孕育一种制造业的全新理念——智能制造，推动移动互联网、大数据、人工智能、云计算、生物工程、新能源、新材料等领域取得新突破。同时，基于信息物理系统的智能装备、智能工厂等智能制造正在引领制造方式变革，大规模个性化定制、精准供应链管理、全生命周期管理等正在重塑产业价值链体系，可穿戴的智能产品、智能家电、智能汽车等智能终端产品不断拓展制造业新领域。

1.1.3 制造业发展战略

面对新一轮工业革命，世界各国高度重视制造业，积极抢占制造业制高点。我国确定并全力推进“制造强国战略”，加快建设制造强国、加快发展先进制造业。本节将具体解读我国制造业现状、全球制造业发展战略与《中国制造 2025》。

（1）我国制造业现状

我国已经是一个制造大国。自 2010 年以来，我国制造业增加值连续十余年位居世界第一。我国也是世界上工业体系最为健全的国家，在五百余种主要工业品中，超过四成的工业产品的产量位居世界第一。

我国已经具备了建设制造强国的基础和条件：首先，我国制造业拥有巨大市场，需求是最强大的发展动力；其次，我国制造业有着世界上门类最为齐全、独立完整的体系，具备强大的产业基础；第三，我国一直坚持信息化与工业化融合发展，在制造业数字化方面掌握了核心关键技术，具有强大的技术基础；第四，我国在制造业人才队伍建设方面已经形成了独特的人力资源优势；第五，我国制造业在自主创新方面已经取得了一些辉煌成就，“上天”、“入地”、“下海”、高铁、输电、发电、国防等各个领域都显示出我国制造业巨大的创新力量。

但是，与世界先进水平相比，我国制造业仍然还存在一定差距，在自主创新能力、资源利用效率、产业结构水平、信息化程度、质量效益等方面差距明显，转型升级和跨越发展的任务紧迫而艰巨。

目前，我国制造业存在的突出问题如下：①自主创新能力不强，核心技术对外依存度较高，产业发展需要的高端设备、关键零部件和元器件、关键材料等大多依赖进口。②产品质量问题突出，国家监督抽查产品质量不合格率偏高，制造业每年直接质量损失超过两千亿元。③资源利用效率低，单位国内生产总值（GDP）能耗约为世界平均水平的两倍，环保任务极重，绿色发展势在必行。④产业结构调整刻不容缓，技术密集型产业和生产性服务业弱，产业集聚和集群发展水平低，具有较强国际竞争力的大企业少。

整体而言，中国制造业整体竞争力还不强，从“制造大国”迈向“制造强国”，中国制造业任重而道远。

（2）世界各国都在高度重视并且全力振兴发展制造业

21 世纪以来，世界上主要国家都非常重视制造业发展战略。一些发达国家纷纷实施“再工业化”战略，强化制造业创新，重塑制造业竞争新优势；一些发展中国家也在加快谋划和布局，积极参与全球产业再分工，谋求新一轮竞争的有利位置。从全球产业发展大趋势来看，发达国家正利用在信息技术领域的领先优势，加快制造工业智能化的进程。

2012 年，美国提出了“先进制造业国家战略计划”，给出中小企业、劳动力、伙伴关系、联邦投资以及研发投资等相关的发展目标和具体实施建议；2019 年提出未来工业发展规划，将人工智能、先进制造技术、量子信息科学和 5G 技术列为推动美国繁荣和保护国家安全的 4 项关键技术。

2013 年，德国政府宣布启动“工业 4.0（Industry 4.0）”国家级战略规划，意图在新一轮工业革命中抢占先机，奠定德国工业在国际上的领先地位，工业

4.0 在国际上，引起极大关注。一般的理解，工业 1.0 对应蒸汽动力时代，工业 2.0 对应电力时代，工业 3.0 对应信息化时代，工业 4.0 则是利用信息化、智能化技术促进产业变革的时代，也就是对应智能互联时代。

2015 年，法国提出了“新工业法国”，确定以“未来工业”为核心，实现工业生产向数字化、智能化转型。

2015 年，韩国提出“制造业创新 3.0 计划”，旨在促进制造业与信息技术相融合，从而创造出新产业，提升韩国制造业的竞争力。

2016 年，日本政府发布“第五期科学技术基本计划”，首次提出“社会 5.0”概念，旨在利用数字化、智能化技术解决日本经济社会可持续发展的关键问题。

上述的战略计划并未冠以“智能制造”，但实际上都包含智能制造的内容。

(3) 中国制造 2025

我国为实现制造强国的战略目标，国务院于 2015 年发布了《中国制造 2025》战略规划。该战略立足国情和现实，以提质增效为中心，以加快新一代信息技术与制造业深度融合为主线，以智能制造为主攻方向，力争通过“三步走”实现制造强国的战略目标，三步走战略如图 1-4 所示。

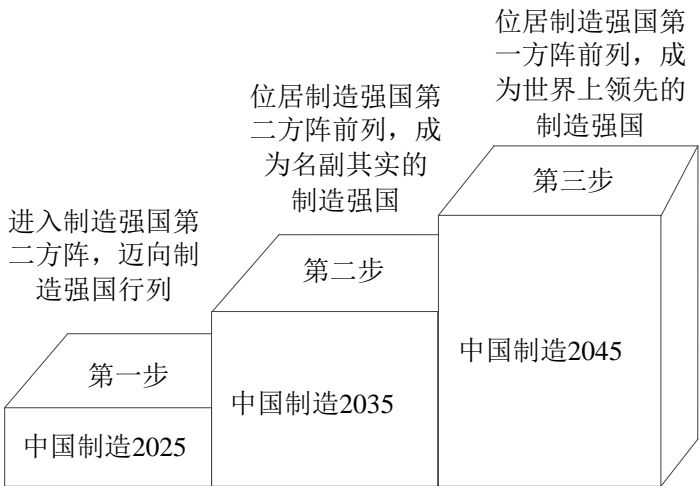


图 1-4 三步走战略

第一步：到 2025 年，制造业整体素质大幅提升，创新能力显著增强，劳动生产率明显提高，两化（工业化和信息化）融合迈上新台阶。重点行业单位工业增加值能耗、物耗及污染物排放达到世界先进水平。形成一批具有较强国际竞争力的跨国公司和产业集群，在全球产业分工和价值链中的地位明显提升。

第二步：到 2035 年，我国制造业整体达到世界制造强国阵营中等水平。创新能力大幅提升，重点领域发展取得重大突破，整体竞争力明显增强，优势行业形成全球创新引领能力，全面实现工业化。

第三步：新中国成立一百年时，制造业大国地位更加巩固，综合实力进入世界制造强国前列。制造业主要领域具有创新引领能力和明显竞争优势，建成全球领先的技术体系和产业体系。

中国制造 2025 可以概括为“一二三四五五十”的总体结构：

“一”：指从制造业大国向制造业强国转变，最终实现制造业强国这一个目标。

“二”：指通过信息化和工业化两化深度融合来引领和带动整个制造业的发展。

“三”：指要通过“三步走”战略，大体上每一步用十年左右的时间来实现我国从制造业大国向制造业强国转变的目标。

“四”：表示四项原则。其中，第一项原则是市场主导、政府引导；第二项原则是既立足当前，又着眼长远；第三项原则是全面推进、重点突破；第四项原则是自主发展和合作共赢。

“五五”：表示五条方针和五大工程。五条方针即创新驱动、质量为先、绿色发展、结构优化和人才为本；五大工程包括制造业创新中心建设工程、强化基础工程、智能制造工程、绿色制造工程和高端装备创新工程。

“十”：表示十大领域，包括新一代信息技术产业、高档数控机床和机器人、航空航天装备、海洋工程装备及高技术船舶、先进轨道交通装备、节能与新能源汽车、电力装备、农机装备、新材料、生物医药及高性能医疗器械十个重点领域。

《中国制造 2025》是我国实施制造强国战略第一个十年的行动纲领，旨在通过动员全社会力量参与发展先进制造业，对于制造业转型升级，推进我国的制造强国进程，抢占国际竞争制高点具有重要意义。

1.1.4 制造业向智能制造转型升级

全球制造业迎来了以数字化、网络化、智能化为发展方向的深刻变革，以新一代信息技术与先进制造技术深度融合为基本特征的智能制造，已成为这次新工业革命的核心驱动力。我国是制造大国，制造业正处于创新升级的重要历史关头。抓牢全球制造业变革带来的历史机遇，主动融入新一轮科技革命和产业变革，加快发展智能制造，是由大变强、全面创新升级的必由之路。

（1）智能制造是我国制造业实现创新发展的内在需求

当前，我国经济已由高速增长阶段迈入高质量发展阶段，社会发展的主要矛盾发生深刻变化，我国制造业面临的需求和环境也出现了深刻改变。整体来看，我国制造已经从总体供给不足过渡到有效供给不足，现有制造产品和服务已不能适应消费结构升级的需要；我国制造业的技术创新能力、资源利用效率、质量效益水平等与世界制造强国尚有差距。我国制造业面临着全面转型升级、全面创新发展、高质量发展的迫切需求。

我国制造业要实现转型升级、创新发展，根本出路在于进一步深化制造业供给侧结构性改革，依靠技术进步满足广大企业创新和转型升级的迫切需求，尤其要加快发展智能制造，推动制造业数字化网络化智能化的发展步伐，实施

制造企业数字化转型；提升制造业的整体技术水平和运营水平，优化制造产品品质和服务质量，提高制造业劳动生产率和经济效益，改变传统制造业的技术模式、生产方式、运营模式，使制造供给体系更好适应消费体系、市场环境的变化。

（2）智能制造为我国制造业创新发展提供了历史性机遇

全球新一轮科技革命和产业变革，为我国制造业的全面创新升级带来了新的历史机遇。我国在制造领域、现代信息技术领域都建立了规模庞大、门类齐全的产业体系，形成了结构完整、配套健全的供应链网络。作为工业制造与信息技术融合发展的交汇点，智能制造引领和推动新一轮工业革命，将进一步促进我国新一代信息技术、先进制造技术的深度融合，助推传统产业实施技术优化升级，支持新兴产业培育和发展，带动新技术、新产品、新装备发展，催生新的经济增长点，推动制造业迈入数字化网络化智能化阶段，促成我国制造业的历史性重大变革。

1.2 智能制造的基本概念

随着移动互联网、大数据、人工智能等新一代信息技术与制造业深度融合，智能制造新理念应运而生。智能制造可以在受到限制，没有经验知识、不能预测的环境下，根据不完全、不精确信息来完成拟人的制造任务，将从根本上提高制造业质量、效率和企业竞争力。本节从智能制造的定义、智能制造的发展历程和未来发展三个方向三个方面论述智能制造的基本概念。

1.2.1 智能制造的定义

工信部在 2016 年发布的《智能制造发展规划（2016-2020 年）》中对智能制造给出了明确的定义：智能制造是基于新一代信息通信技术与先进制造技术深度融合，贯穿于设计、生产、管理、服务等制造活动的各个环节，具有自感知、自学习、自决策、自执行、自适应等功能的新型生产方式。智能制造包括智能制造技术（Intelligent Manufacturing Technology, IMT）与智能制造系统（Intelligent Manufacturing System, IMS）。

（1）智能制造技术

智能制造技术是指利用计算机模拟制造专家的分析、判断、推理、构思和决策等智能活动，并将这些智能活动与智能机器有机融合，使其贯穿应用于制造企业的各个子系统（如经营决策、采购、产品设计、生产计划、制造、装配、质量保证和市场销售等）的先进制造技术。该技术能够实现整个制造企业经营运作的高度柔性化和集成化，取代或延伸制造环境中专家的部分脑力劳动，并对制造业专家的智能信息进行收集、存储、完善、共享、继承和发展，从而极大地提高生产效率。

(2) 智能制造系统

智能制造系统是由部分或全部具有一定自主性和合作性的智能制造单元组成的、在制造活动全过程中表现出相当智能行为的制造系统。其最主要的特征是在工作过程中对知识的获取、表达与使用。根据其知识来源，智能制造系统可分为两类：一是以专家系统为代表的非自主式制造系统，该类系统的知识由人类的制造知识总结归纳而来；二是自主式制造系统，该类系统可以在工作过程中不断自主学习完善与进化自有的知识，因而具有强大的适应性以及高度开放的创新能力。随着以神经网络与深度学习为代表的人工智能技术的发展，智能制造系统正逐步从非自主式智能制造系统向具有自学习、自进化与自组织的具有持续发展能力的自主式智能制造系统过渡发展。

智能制造系统是数字化、网络化、智能化信息技术与产品设计、生产、服务等产品全生命周期制造活动全面集成优化的工业生产系统。

1.2.2 智能制造的发展历程

智能制造作为制造业和信息技术深度融合的产物，其诞生和演变是和信息化发展相伴而生的。在长期实践演进过程中，智能制造形成了三种基本范式，即数字化制造——第一代智能制造，数字化、网络化制造——“互联网+制造”或第二代智能制造，数字化、网络化、智能化制造——新一代智能制造。

(1) 数字化制造—智能制造的第一范式

20 世纪中叶以后，随着制造业对技术进步的强烈需求，以及计算机、通信和数字控制等信息化技术的发明和广泛应用，制造系统进入了数字化制造时代，以数字化为标志的信息革命引领和推动了第三次工业革命。数字化制造是智能制造的第一种基本范式，也可以称为第一代智能制造。

与传统制造系统相比，数字化制造系统最本质的变化是在人和物理系统（Human-Physical Systems, HPS）之间增加了一个信息系统，从原来的“人-物理”二元系统发展成为“人-信息-物理”（Human-Cyber-Physical Systems, HCPS）三元系统。人的相当部分的感知、分析、决策和控制功能迁移给信息系统，信息系统可以代替人类完成部分脑力劳动。信息系统是由软件和硬件组成的系统，其主要作用是对输入的信息进行各种计算分析，并代替操作者去控制物理系统完成工作任务。

(2) 数字化、网络化制造—智能制造的第二范式

20 世纪末、21 世纪初，互联网技术快速发展并得到普及和广泛应用，“互联网+”不断推进制造业和互联网融合发展，制造技术与数字技术、网络技术的密切结合重塑制造业的价值链。数字化、网络化制造是智能制造的第二种基本范式，也可称为“互联网+制造”，或第二代智能制造。

“互联网+制造”是在数字化制造的基础上，用网络将人、流程、数据和事物连接起来，联通企业内部和企业间的“信息孤岛”，通过企业内、企业间的协同和各种社会资源的共享与集成，实现产业链的优化，快速、高质量、低成本地为市场提供所需的产品和服务。先进制造技术和数字化、网络化技术的融合，使得企业对市场变化具有更快的适应性，能够更好地收集用户对使用产品和对产品质量的评价信息，在制造柔性化、管理信息化方面达到了更高的水平。

(3) 数字化、网络化、智能化制造—新一代智能制造

当前，工业互联网、大数据及人工智能实现群体突破和融合应用，以新一代人工智能技术为主要特征的信息化开创了制造业数字化、网络化、智能化制造的新阶段。新一代智能制造的突破和广泛应用将重塑制造业的技术体系、生产模式、产业形态。

近年来，随着人工智能算法的重大突破、计算能力的极大提高，互联网引发了真正的大数据革命，在算法、算力、数据与其他各种先进技术互融互通的基础上，人工智能技术已经实现战略突破，进入了“新一代人工智能”时代。新一代人工智能具备了学习的能力，能够生成知识和更好地运用知识。

1.2.3 智能制造的发展方向

随着人工智能、5G、大数据、云计算、物联网等的进一步发展，在下一代智能制造中，机器与人的关系将由协作转向共融，借助云上“大脑”达到感知智能层级，数字工程师将处理某些专业领域的工作并与人进行交流，商业智能也会应用得更加广泛。

(1) 人机共融

随着人工智能的发展，人机不再是单纯的协作关系，可以是共融关系。即在同一自然空间内，充分利用人和机器人的差异性与互补性，通过人机个体间的融合，人机群体间的融合，人机融合后的共同演进，实现人机共融共生，人机紧密协调，自主完成感知与计算。实现人机共融后，机器人与人的感知过程、思维方式和决策方法将会紧密耦合。

人机共融是人与机器人关系的一种抽象概念，它主要有以下四个方面的内涵：① 人机智能融合，人与机器人在感知、思考、决策上有着不同层面的互补；② 人机协调，人与机器人能够顺畅交流，协调动作；③ 人机合作，人与机器人可以分工明确，高效地完成同一任务；④ 人机共进，人与机器人相处后，彼此间的认知更加深刻。

(2) 云端协同

云端协同技术借助于 5G 网络、云计算与人工智能技术，达到了感知智能层级，位于云端数据中心具有强大存储能力和运算能力的“大脑”，利用人工智能算法和其他先进的软件技术，通过 5G 通信网络来控制本地端，使云端能全

面感知环境、相互学习、共享知识，不仅能够降低成本，还会帮助本地端提高自学能力、自适应能力。云端协同技术将带来技术，社会、工业各个层面发生颠覆性的变化，包括新的价值链、新的技术、新的体系结构、新的体验和新的商业模式等。

（3）数字工程师

新一代智能制造系统进一步完善了信息系统的功能，使信息系统具备了认知和学习的能力，形成新一代“人-信息系统-物理系统”。信息系统能够代替人完成部分的认知和学习等脑力劳动，促使人和信息系统的关系发生了根本性的变化。未来的智能制造系统将会逐步摆脱对人的依赖，其信息系统具有更强的知识获取和知识发现的能力，能够代替人管理整个或者部分制造领域中的知识。我们将这种具有高度自主决策能力的智能化系统称为数字工程师。

数字工程师是人机协作时代的一个典型产物，是能够自我学习成长的具有灵敏情感反应的人类工作伙伴。数字工程师能在新一代智能制造的信息系统中发挥自身独特优势，具有强大的感知、计算分析与推理能力，同时具有学习提升、自主决策、产生知识的能力。

（4）商业智能

商业智能又称商业智慧或商务智能，利用现代的数据仓库技术、联机分析处理技术、人工智能技术和数据可视化展示技术进行数据分析和呈现，完成从数据到信息的转化，其目标是为决策提供支持。商业智能的核心是完成数据到信息的转化，为决策提供支撑。

从技术的角度，商业智能的执行过程是：企业决策人员以企业数据库为基础，通过利用联机分析处理和人工智能技术以及决策相关的专业知识，从数据中提取有价值的信息，然后根据信息作出决策。从应用的角度，商业智能可以协助用户对商业数据进行处理和分析。从数据的角度，商业智能将内部事务性数据、供应链上下游数据以及外部竞争数据通过抽取、转换和加载后转移到数据库中，然后通过聚集、切片、分类和分析等，将数据库中的数据转化为有价值的信息，为决策提供支撑。

1.3 工业大数据的基本概念

智能制造是工业大数据的载体和产生来源，其各环节的信息化、自动化系统所产生的数据构成了工业大数据的主体。另一方面，智能制造又是工业大数据形成数据产品的应用场景和目标。工业大数据描述了智能制造各生产阶段的真实情况，为人类理解、分析和优化制造过程提供了宝贵的数据资源。

1.3.1 工业大数据的定义

工业大数据是指在工业生产过程中整个产品全生命周期各个环节所产生的

各类数据的总称，涉及客户需求、销售、订单、计划、研发、设计、工艺、制造、采购、供应、库存、发货、交付、售后服务、运维、报废或回收再制造等。工业大数据以产品数据为核心，极大延展了传统工业数据的范围。

工业大数据具备双重属性：价值属性和产权属性。一方面，通过工业大数据分析等关键技术能够实现设计、工艺、生产、管理、服务等各个环节智能化水平的提升，满足用户定制化需求，提高生产效率并降低生产成本，为企业创造可量化的价值；另一方面，这些数据具有明确的权属关系和资产价值，企业能够决定数据的具体使用方式和边界，数据产权属性明显。工业大数据的价值属性实质上是基于工业大数据采集、存储、分析等关键技术，对工业生产、运维、服务过程中的数据实现价值提升或变现；工业大数据的产权属性则偏重于通过管理机制和管理方法帮助工业企业明晰数据资产目录与数据资源分布，确定所有权边界，为其价值的深入挖掘提供支撑。

1.3.2 工业大数据的来源与类型

工业大数据的来源广泛，主要包括生产经营相关业务数据、制造过程数据和企业外部数据三个方面，具体介绍如下：

第一类是生产经营相关业务数据。主要来自传统企业信息化范围，被收集存储在企业信息系统内部，包括传统工业设计和制造类软件、企业资源计划（Enterprise Resource Planning, ERP）、产品生命周期管理（Product Lifecycle Management, PLM）、供应链管理（Supply Chain Management, SCM）、客户关系管理（Customer Relationship Management, CRM）和环境管理系统（Environmental Management System, EMS）等信息系统中。通过这些企业信息系统积累大量的产品研发数据、生产性数据、经营性数据、客户信息数据、物流供应数据及环境数据。此类数据是工业领域传统的数据资产，在移动互联网等新技术应用环境下正在逐步扩大范围。

第二类是制造过程数据。主要指涵盖操作和运行情况、工况状态、环境参数等体现设备和产品运行状态的数据，通过制造执行系统（Manufacturing Execution System, MES）系统实时传递。狭义的工业大数据即指该类数据，即工业设备和产品快速产生的并且存在时间序列差异的大量数据。这类数据主要来源包括工业现场数据采集和工业产品数据采集。工业现场的数据采集针对现场工业控制系统和设备进行，通过生产现场的自动化与控制系统，如数据采集与监视控制系统（Supervisory Control And Data Acquisition, SCADA）、分布式控制系统（Distributed Control System, DCS）、现场总线控制系统（Fieldbus Control System, FCS）等，借助可编程控制器 PLC、传感器、采集器、射频识别等实现对地理位置集中的底层设备或分散的工业现场设备进行监视与数据采集。工业产品数据在产品或装备在客户端投入使用后，通过 4G、5G、窄带物

联网（Narrow Band Internet of Things, NB-IoT）等无线通信技术接入工业互联网，利用标识、传感器等获取产品信息、能耗、温度、工作电流、电压等实时指标数据。目前在智能装备大量应用的情况下，此类数据量增长最快

第三类是企业外部数据。包括工业企业产品售出之后的使用、运营情况的数据，同时还包括大量客户名单、供应商名单、外部的互联网等数据。

表 1-2 工业数据的主要来源

数据类型	主要来源
生产经营相关业务数据	来自企业信息系统，包括企业资源（ERP）、产品生命周期管理（PLM）、供应链管理（SCM）、客户关系管理（CRM）和能耗管理系统（EMS）等
制造过程数据	来自数据采集与监视控制系统（SCADA），分布式控制系统（DCS）、现场总线控制系统（FCS）等
企业外部数据	产品使用、运营数据；客户信息数据等

工业大数据的来源不同，产生方式各异，也导致数据类型多种多样。以下给出了基于不同划分标准的工业大数据类型。

从数据产生环节来看，工业大数据可以分为研发数据、生产数据、运维数据、管理数据和外部数据。研发数据主要包括研发设计数据、开发测试数据等；生产数据主要有控制信息、工况状态、工艺参数、系统日志等；运维数据包括监控报警事件、操作日志、性能指标数据等；管理数据主要有系统设备资产信息、客户与产品信息、产品供应链数据、业务统计数据等；外部数据是指与其他主体共享的数据等。

从数据形式看，制造业数据可以分为结构化数据、半结构化数据和非结构化数据。结构化数据是由二维表结构来逻辑表达和实现的数据，严格地遵循数据格式与长度规范，主要通过关系型数据库进行存储和管理，企业的 ERP、传感器采集的过程数据都属于结构化数据。例如环境温度、湿度数据；设备温度、压力数据；能耗数据等。半结构化数据并不符合关系型数据库或其他数据表的形式关联起来的数据模型结构，但包含相关标记，用来分隔语义元素以及对记录 and 字段进行分层。例如，接口数据；不同工人的个人信息就是半结构化数据。非结构化数据是数据结构不规则或不完整，没有预定义的数据模型，不方便用数据库二维逻辑表来表现的数据，包括所有格式的办公文档、文本、图片、XML/HTML、各类报表、图像和音频、视频信息等。

表 1-3 工业数据按数据类型分类表

结构化数据	半结构化数据	非结构化数据
① 环境数据，如环境温度、湿度等② 设备数据，如设备温度、压力、速度等③ 能耗数据、产品数据，如属性数据、指标数	① 接口数据，如 JSON 格式、XML 格式等；	① 工艺知识，如工艺机理、工程图纸等；② 信息文档，如 word 文档等；③ 生产监控信息，如监控

从数据处理的角度看，制造业数据可以分为原始数据与衍生数据。原始数据是指来自上游系统的，没有做过任何加工的数据；衍生数据是指通过对原始数据进行加工处理后产生的数据。衍生数据包括各种数据集市、汇总层、数据分析和挖掘结果等。虽然会从原始数据中产生大量衍生数据，但还是会保留一份未作任何修改的原始数据，一旦衍生数据发生问题，可以随时从原始数据重新计算。

1.3.3 工业大数据的特征

工业数据来源与数据类型多样，在应用大数据分析技术进行工业数据挖掘时，除了考虑传统大数据的典型形态特征外，还需要基于传统工业过程的本质特征，明确工业数据的新特征及其处理难点，归纳其应用特征，制定合理的工业大数据处理规划和技术路线。

(1) 工业大数据的形态特性

随着传感器的普及以及数据采集、存储技术的飞速发展，工业大数据同样呈现了大数据具备的“5V”特性，即规模性（Volume）、多样性（Variety）、高速性（Velocity）、价值密度低（Value）、真实性（Veracity）。

① 规模性：工业大数据体量很大，大量机器设备的高频数据和工业互联网数据持续涌入，大型工业企业的数据集将达到 PB 级甚至 EB 级。

② 多样性：指数据类型多样和来源广泛。工业数据分布广泛，数据来源于机器设备、工业产品、管理系统、物联网等各个环节，并且结构复杂，既有结构化和半结构化的传感数据，也有非结构化数据。

③ 高速性：指生产过程中对数据的获取和处理实时性要求高，部分传感器信号采集频率高、数据产生速度快，同时生产现场可能要求数据处理分析时间达到毫秒级，从而为智能制造生产优化提供及时的决策依据。

④ 价值密度低：工业大数据更强调用户价值驱动和数据本身的可用性，而在制造业的海量数据中，有用的数据所占比重极低，导致整个制造业数据的价值密度低，想要从海量数据有用的信息也就更加困难。

⑤ 真实性：是指数据的真实性、完整性和可靠性，更加关注数据质量，以及数据处理、分析技术和方法的可靠性。相对于分析结果的高可靠性要求，工业大数据的真实性比较低。工业应用中因为技术、成本等原因，很多关键的量没有被测量或者没有被精确测量，同时某些数据具有固有的不可预测性，这些情况导致往往数据真实性不高，是数据分析和利用最大的障碍。

工业大数据作为对工业相关要素的数字化描述，除了具备传统的大数据“5V”共性特点以外，还具有多模态、强关联、高通量等形态特征：

① 多模态：工业大数据是工业系统在赛博空间的映像，必须反映工业系

统的系统化特征及其各方面要素，所以记录的数据必须完整，往往需要复杂的结构、不同格式、不同维度的信息进行表征。同时，当生产过程工况发生改变时，工业大数据也会呈现出多模态的特征。

② 强关联：强关联反映的是工业的系统性及其复杂动态关系，本质是指物理对象之间和过程的语义关联，包括产品部件之间的关联关系，生产过程的数据关联，产品生命周期设计、制造、服务等不同环节数据之间的关联。

③ 高通量：高通量即工业传感器要求瞬时写入超大规模数据。机器设备所产生的时序数据涉及海量的设备与测点，数据采集频度高（产生速度快）、数据总吞吐量大、持续不断，呈现出高通量的特征。

（2）工业大数据的应用特性

工业大数据的应用特征是工业对象本身的特性或需求所决定的，工业大数据的应用特征可归纳为跨尺度、协同性、多因素、强机理等几个方面。

① 跨尺度：跨尺度是由工业的复杂系统性所决定。工业过程向大规模、动态性、集成化发展，多单元、多产品生产动态运行，采集到的过程数据涉及到多个空间层面，不同空间层面的数据采集规则不尽相同，需要将这些不同空间尺度的信息集成到一起；另外，跨尺度不仅体现在空间尺度，还体现在时间尺度上，不同业务常常需要将毫秒级、分钟级、小时级等不同时间尺度的信息集成起来。

② 协同性：协同性是由于工业系统强调子系统间的动态协同而引起的。在工业系统中，当聚焦到具体的某台设备、某个部门、某个管理的指标变化时，需要兼顾工艺流程、生产调度等信息的变化，因而需要进行信息集成，从而促成信息的自动流动，加强信息感知能力，减小面临的不确定性。

③ 多因素：多因素特征则是由工业过程本身的物化反应特性决定。工业过程中包含的物化反应通常十分复杂，在认识工业过程时需要全面、历史地了解其全貌。因此，往往要求尽可能完整地收集与工业对象相关的各类数据，从多个方面对系统进行描述。

④ 强机理：工业过程具有的复杂机理，任何生产单元及其子系统发生的数据变化、各类现象特征，都可能通过物质流、能量流、信息流在不同系统层级间传播并不断演化，因而表现出较强的机理特性。通过机理分析来处理数据间定性的问题，进而运用数据来确定各类定量关系，进而可以获得准确、可靠的数据分析模型。

1.3.4 工业大数据的采集

工业数据采集是利用泛在感知技术对多源设备、异构系统、运营环境、人等要素信息进行实时高效采集，通过各类通信手段接入不同设备、系统和产品，采集大范围、深层次的工业数据，以及异构数据的协议转换与边缘处理。

(1) 工业数据采集范围

工业数据采集广义范围包括工业现场设备的数据采集、工厂外智能产品/装备的数据采集，以及 ERP、MES 等应用系统的数据采集，具体如下：

① 工业现场设备的数据采集：主要通过现场总线、工业以太网、工业光纤网络等工业通信网络实现对工厂内设备的接入和数据采集。可分为三类：对传感器、变送器、采集器等专用采集设备的数据采集；对 PLC、远程终端单元（Remote Terminal Unit, RTU）、嵌入式系统、进程间通信（Inter-Process Communication, IPC）等通用控制设备的数据采集；对机器人、数控机床、搬运机器人等专用智能设备/装备的数据采集。主要用于工业现场生产过程的可视化和持续优化，实现智能化控制与决策。

② 工厂外智能产品/装备的数据采集：主要通过工业物联网实现对工厂外智能产品/装备的远程接入和数据采集，采集智能产品/装备运行时的关键指标数据，用于实现智能产品/装备的远程监控、健康状态监测和远程维护等应用。

③ 对 ERP、MES 等应用系统的数据采集：主要由工业互联网平台通过接口和系统集成方式实现对 MES、ERP 等应用系统的数据采集。

(2) 工业数据采集体系框架

工业数据采集体系架构包括设备接入、协议转换、边缘数据处理三层，向下接入设备或智能产品，向上与工业互联网平台/工业应用系统对接，如图 1-5 所示：



图 1-5 工业数据采集体系框架

① 设备接入：通过工业以太网、工业光纤网络、工业总线、3G/4G、NB-IoT 等各类有线和无线通信技术，接入各种工业现场设备、智能产品/装备，采集工业数据。

② 协议转换：一方面运用协议解析与转换、中间件等技术兼容 Modbus、

CAN、Profinet 等各类工业通信协议，实现数据格式转换和统一。另一方面利用 HTTP、MQTT 等方式将采集到的数据传输到云端数据应用分析系统或数据汇聚平台。

③ 边缘数据处理：基于高性能计算、实时操作系统、边缘分析算法等技术支撑，在靠近设备或数据源头的网络边缘侧进行数据预处理、存储以及智能分析应用，提升操作响应灵敏度、消除网络堵塞，并与云端数据分析形成协同。

(3) 工业数据采集协议—OPC

目前在工业数据采集领域，多种工业协议标准并存，各种工业协议标准不统一、互不兼容，导致协议解析、数据格式转换和数据互联互通困难。协议转换技术就是将不同的工业通信协议通过协议解析、数据转换和地址空间重映射等技术手段转换成统一协议，实现设备数据采集的信息交互以及和信息系统的互联互通。可以支持常用的工业协议，通过协议转换降低了设备组网的难度，实现了访问的统一性。

OPC 全称是 Object Linking and Embedding (OLE) for Process Control，它的出现为基于 Windows 的应用程序和现场过程控制应用建立了桥梁，是一项应用于自动化行业及其他行业的数据安全交换可互操作性标准。它独立于平台，并确保来自多个厂商的设备之间信息的无缝传输，OPC 基金会负责该标准的开发和维护。

OPC 数据访问 (Open Platform Communication Data Access, OPC DA)：OPC 数据访问是指提供从数据源读取和写入特定数据的手段，OPC 数据访问 (DA) 定义了包括数据值，更新时间与数据品质信息的相关标准。

OPC 数据访问对象由分层结构构成，OPC 客户端通过创建一个 OPC Server 对象来建立一个到服务器的连接，服务器对象提供方法通过浏览地址空间分层寻找项目及其属性，如数据类型和访问权限，为了访问数据，客户端根据相同的设置（如更新时间）将 OPC 项目分组到一个 OPC Group 对象。

OPC 采用客户端/服务器 (C/S) 方式进行信息交换。OPC A&E 定义了报警与事件的信息，OPC HDA 定义了查询、分析历史数据和含有时标的数据的方法，OPC 服务器封装了过程信息来源，使信息可以通过它的接口访问，OPC 客户端连接到 OPC 服务器后，可以访问和使用它所提供的数据。工业过程中常见的数据采集框架如图 1-6 所示。

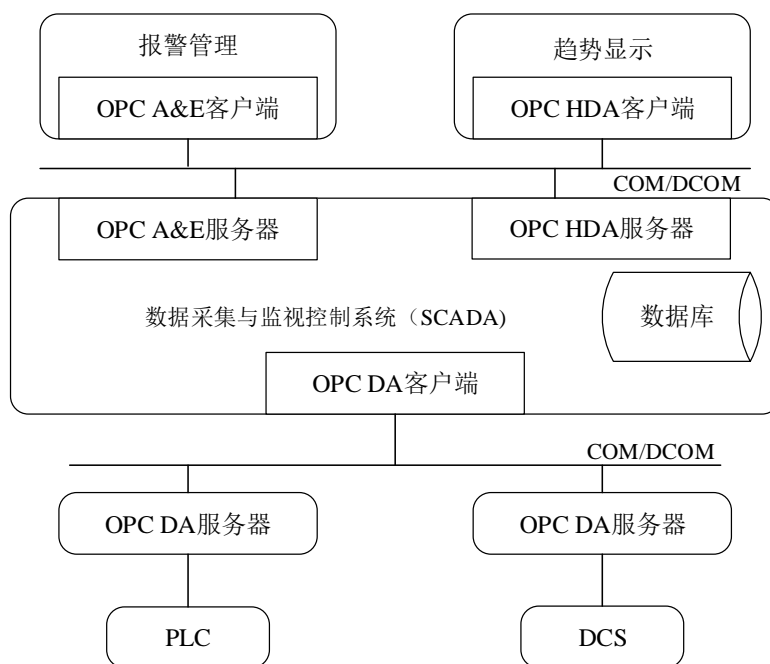


图 1-6 OPC 采集框架图

工业数据采集实现了对生产现场各种工业数据的实时采集和整理，为企业的 MES、ERP 等信息系统提供大量工业数据，是智能制造大数据分析的基础和前提。通过对采集的工业大数据深入挖掘与分析，实现生产过程优化和智能化决策。

1.4 智能制造大数据分析需求与流程

工业大数据分析其本质是通过促进数据的自动流动去解决和控制业务问题，减少决策过程中所带来的不确定性，并尽量克服人工决策的缺点，是智能制造的核心。本节主要讨论一些典型的智能制造大数据分析需求，并对大数据分析的基本流程与关键技术进行简单介绍。

1.4.1 智能制造大数据的典型分析需求

当前制造业中存在着数据量大、数据利用率低的矛盾，许多场景中的数据规模甚至已经超出传统数据分析方法所能承受的极限。通过大数据分析技术发掘蕴含在数据中的宝贵知识和财富，在实际工业生产中发挥着重要作用。利用大数据推动智能制造主要包括智能化设计、计划调度、质量监控、生产过程优化和运行维护等。

（1）智能化设计

大数据驱动的产品智能设计，根据前端互联网用户评价等数据快速准确地分析和预测市场需求，通过后端制造、运维等数据动态关联产品结构、功能设计方案，并基于历史设计方案学习提高设计方案评价及智能决策能力，从而形成前后端横向集成的主动设计模式。

(2) 计划调度

大数据驱动的生产计划调度，依据车间制造过程数据，通过数据分析挖掘车间实时状态参数与加工时间、等待时间、运输时间的复杂演变规律及映射关系，可实现车间调度中产品完工时间的精准预测，并基于预测结果实现复杂动态环境下的车间实时调度。

(3) 质量监控

大数据驱动的质量监控，依据产品制造过程数据及质检数据实现产品追溯，通过关联分析识别影响质量的主要因素，如原料性能参数、设备状态参数、工艺参数、车间环境参数等，并建立质量影响因素与质量性能的映射模型以有效预测产品质量，进一步利用智能优化算法自适应地实时调整影响产品质量的控制参数，实现产品质量自适应控制与优化。

(4) 生产过程优化

通过分析产品质量、成本、能耗、效率、成材率等关键指标与工艺、设备参数之间的关系，优化产品设计和工艺。以实际的生产数据为基础，建立生产过程的仿真模型，优化生产流程。

(5) 运行维护

在实际生产线上，任何一个小的故障不及时解决，都可能会影响到整个生产系统的正常运行，带来巨大的经济损失。因此，快速、准确地进行故障检测、诊断和预警，一直是制造业企业迫切需求。然而，生产线是一个复杂的系统，整个生产线有数十道甚至上百道工序，每个工序有多级操作，每步操作又对应不同的机器和设备。对于这样的复杂系统，用传统方法去定量分析建模、找到异常状况的影响因素是十分困难的。

大数据驱动的运行维护，通过实时监测的制造过程数据及设备性能参数等时间序列数据，揭示系统故障特征时序变化规律及征兆性表征，主动提前发现系统运行过程中的潜在异常，并通过历史异常数据聚类分析对潜在异常进行诊断，精准定位异常源，结合历史诊断数据做出运维决策进行预防性维护，从而在重大异常发生之前消除隐患。

1.4.2 智能制造大数据的一般分析流程

在工业生产过程中，大数据分析的基本流程主要包括数据采集、数据存储、数据预处理、数据分析等，是通过数据分析挖掘出蕴含的知识以指导生产的过程，如图 1-7 所示。

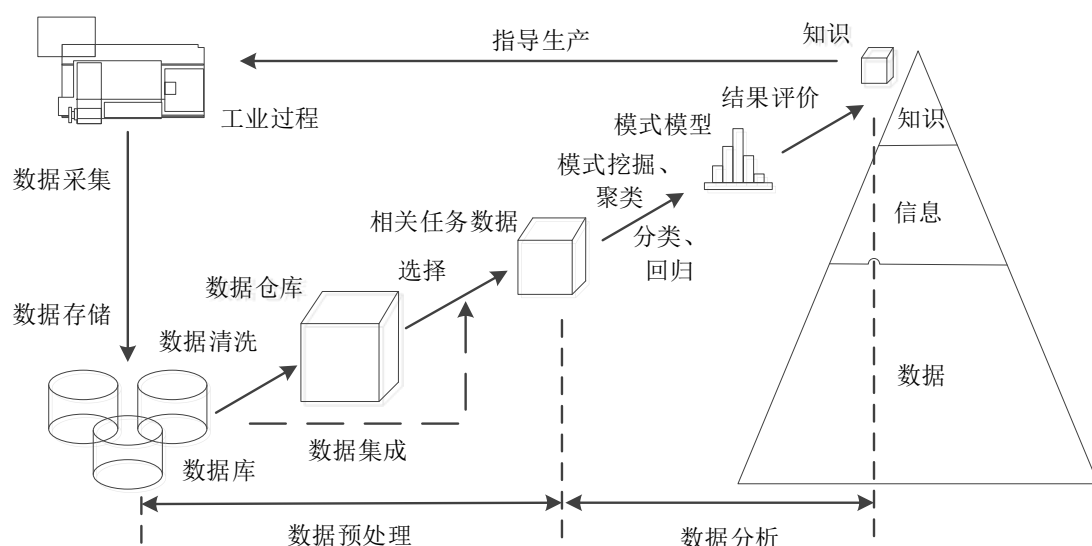


图 1-7 工业大数据分析流程

（1）数据采集与存储

数据采集是利用数据采集装置，从系统外部采集数据并输入到系统内部。数据的采集是获得有效数据的重要途径，同时也是工业大数据分析和应用的基础。数据采集与治理的目标是从企业内部和外部等数据源获取各种类型的数据，并围绕数据的使用，建立数据标准规范和管理机制流程，保证数据质量，提高数据管控水平。在智能制造中，数据分析往往需要更精细化的数据，因此对数据采集能力有着较高的要求。

数据存储是指将数据以某种格式记录在计算机内部或外部存储介质上进行保存，其存储对象包括数据流在加工过程中产生的临时文件或加工过程中需要查找的信息。在数据存储中，数据流反映了系统中流动的数据，表现出动态数据的特征；数据存储反映系统中静止的数据，表现出静态数据的特征。

制造业数据具有体量大、关联复杂、时效要求高等特点，对数据存储技术提出了很高的要求。数据存储管理系统可以分为单机式存储和分布式存储两类。单机式数据存储较为传统，一般采用关系数据库与本地文件系统结合的存储方式，无法为大规模数据提供高效存储和快速计算的支持。分布式数据存储工作节点多，能够提供大量的存储空间，同时能够与互联网技术结合，数据请求及处理速度较快。

（2）数据预处理

数据预处理是指对所收集数据进行分类或分组前所做的审核、筛选、排序等必要的处理。现实世界中数据大体上都是不完整、不一致的脏数据，无法直接进行数据挖掘，或挖掘结果差强人意。为了提高数据挖掘的质量产生了数据预处理技术。

数据预处理有多种方法，包括数据清理、数据集成、数据变换、数据归约

等。这些数据处理技术在数据挖掘之前使用，能提高数据挖掘结果的质量，降低实际挖掘所需要的时间。数据清洗通过填补缺失值、光滑噪声数据，识别或删除离群点，并解决不一致性问题，来“清洗”数据。数据集成将数据由多个数据源合并成一个一致性的数据存储。数据归纳是使用数据编码方案，以便得到原始数据的简化或“压缩”表示。数据变换是将数据规范化、数据离散化和概念分层化。

（3）数据分析

数据分析是指用适当的分析方法对收集来的大量数据加以汇总和理解并消化，以求最大化地开发数据的功能，发挥数据的作用。数据分析是为了提取有用信息和形成结论而对数据加以详细研究和概括总结的过程，是智能制造重要环节之一。与其他领域数据分析不同，制造业数据分析需要融合生产过程的机理模型，以“数据驱动+机理驱动”的双驱动模式进行数据分析，从而建立高精度、高可靠性的模型来真正解决实际工业问题。

工业数据分析常见的方法包括频繁模式挖掘、聚类分析、分类分析和回归分析。频繁模式挖掘是挖掘频繁出现在数据集中的模式（如项集、子序列或子结构），将在第 4 章中具体讨论；聚类分析是根据样本之间的相似度将整个样本集合聚集成若干个类的过程，为第 5 章内容；分类分析根据已经掌握的每类若干样本的数据信息，总结出分类的规律性，建立判别公式和判别规则，是第 6 章内容；回归分析确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法，为第 7 章内容。

1.5 本章小结

本章为智能制造大数据概述，主要介绍工业制造过程基本概念与发展战略；智能制造的定义、内涵，工业大数据的定义、来源、类型、特点及采集方式；智能制造中的数据分析需求与分析流程。

工业制造过程：介绍了工业制造过程的概念、意义与四个发展阶段；介绍了我国制造业现状和全球制造业发展格局；介绍了智能制造对于制造业全面转型升级的重要意义。

智能制造：介绍了智能制造的基本概念、四个方面的内涵、三个发展阶段和四个未来方向。

工业大数据：介绍了工业大数据的基本概念、来源、类型、特征和工业数据采集的定义、范围、框架、系统等内容。

智能制造大数据分析需求与分析流程：介绍了工业大数据四个方面的分析需求与分析流程概述。

习题

- 1、简述工业过程的基本概念及其分类。
- 2、简述制造业的发展历程。
- 3、简述智能制造的定义与内涵。
- 4、简述智能制造的发展阶段。
- 5、工业大数据的来源有哪些？包括哪些典型类型？分别具有什么特点？
- 6、智能制造大数据分析通常有哪些典型需求？请结合实际应用进行陈述。
- 7、简述大数据分析的基本流程。

参考文献

- [1] 周济. 智能制造——“中国制造 2025”的主攻方向[J]. 中国机械工程, 2015, 26(17): 2273-2284.
- [2] 周济, 李培根. 智能制造导论[M]. 北京: 高等教育出版社, 2021.2.
- [3] 李培根. 智能制造概论[M]. 北京: 清华大学出版社, 2021.10.
- [4] 李杰, 倪军, 王正安. 从大数据到智能制造[M]. 上海: 上海交通大学出版社, 2016. 05.
- [5] 柴天佑. 大数据与制造流程自动化发展战略研究[M], 科学出版社, 2019.
- [6] 刘敏, 严隽薇. 智能制造: 理念系统与建模方法[M]. 北京: 清华大学出版社, 2019.04.
- [7] 钱锋, 桂卫华. 人工智能助力制造业优化升级[J]. 中国科学基金, 2018, 32(03): 257-261. DOI:10.16262/j.cnki.1000-8217.2018.03.004.
- [8] 柴天佑, 丁进良. 流程工业智能优化制造[J]. 中国工程科学, 2018, 20(04): 51-58.
- [9] D. De Silva, S. Sierla, D. Alahakoon, E. Osipov, X. Yu and V. Vyatkin. Toward Intelligent Industrial Informatics: A Review of Current Developments and Future Directions of Artificial Intelligence in Industrial Applications[J]. *IEEE Industrial Electronics Magazine*, 2020, 14(02), 57-72.
- [10] Feng Qian, Weimin Zhong, Wenli Du. Fundamental Theories and Key Technologies for Smart and Optimal Manufacturing in the Process Industry. *Engineering*, 2017, 154-160. <http://dx.doi.org/10.1016/J.ENG.2017.02.011>
- [11] 钟志华, 臧冀原, 延建林, 苗仲桢, 杨晓迎, 古依莎娜. 智能制造推动我国制造业全面创新升级[J]. 中国工程科学, 2020, 22(06): 136-142.
- [12] Ji Zhou, Peigen Li, Yanhong Zhou, Baicun Wang, Jiyuan Zang, Liu Meng. Toward New-Generation intelligent manufacturing [J]. *Engineering*, 2018, 4(01) :28-47.

- [13] Baicun Wang, Fei Tao, Xudong Fang, Chao Liu, Yufei Liu, Theodor Freiheit. Smart manufacturing and intelligent manufacturing: A comparative review[J]. *Engineering*, 2021, 7(06): 738-757.
- [14] 张映锋, 张党, 任杉. 智能制造及其关键技术研究现状与趋势综述[J]. 机械科学与技术, 2019, 38(03): 329-338.
- [15] P. Patel, M. I. Ali, A. Sheth. From raw data to smart manufacturing: AI and semantic web of things for industry 4.0[J]. *IEEE Intelligent Systems*, 2018, 33(04): 79-86.
- [16] G. D'Emilia, A. Gaspari, E. Natale. Mechatronics applications of measurements for smart manufacturing in an industry 4.0 scenario[J]. *IEEE Instrumentation & Measurement Magazine*, 2019, 22(02): 35-43.
- [17] 张曙. 工业 4.0 和智能制造[J]. 机械设计与制造工程, 2014, 43(08) :1-5.
- [18] 中国信息通信研究院. 工业大数据白皮书（2019 版）[R]. 2019-09.
- [19] 王晨, 郭朝晖, 王建民. 工业大数据及其技术挑战[J]. 电信网技术, 2017(08): 1-4.
- [20] Hafiz Mughees Ahmad, Afshin Rahimi. Deep learning methods for object detection in smart manufacturing: A survey[J]. *Journal of Manufacturing Systems*, 2022, 64: 181-196.
- [21] 何文韬, 邵诚. 工业大数据分析技术的发展及其面临的挑战[J]. 信息与控制, 2018, 47(04): 398-410.