

于信息技术与社会发展的深度融合。智慧城市的五个核心能力要素密切关联且相互影响，但不可互为替代，均是开展新一阶段智慧城市整体、局部乃至具体项目建设、运行需要关注的核心能力要素。对核心能力要素解释为：

- 数据治理：围绕数据这一新的生产要素进行能力构建，包括数据责权利管控、全生命周期管理及其开发利用等。
- 数字孪生：围绕现实世界与信息世界的互动融合进行能力构建，包括社会孪生、城市孪生和设备孪生等，将推动城市空间摆脱物理约束，进入数字空间。
- 边际决策：基于决策算法和信息应用等进行能力构建，强化执行端的决策能力，从而达到快速反应、高效决策的效果，满足对社会发展的敏捷需求。
- 多元融合：强调社会关系和社会活动的动态性及其融合的高效性等，实现服务可编排和快速集成，从而满足各项社会发展的创新需求。
- 态势感知：围绕对社会状态的本质反映及模拟预测等进行能力构建，洞察可变因素与不可见因素对社会发展的影响，从而提升生活质量。

根据智慧城市参考基本原理，智慧城市建设与发展内容主要面向城市治理、惠民服务、生态宜居、产业发展等相关城市场景构建服务能力，为政府、企业、市民等提供服务。服务场景的构建，不仅需要技术与平台、基础设施等方面的共性技术支撑能力构建和数据要素支撑，也需要安全保障体系、建设运营体系的建设，同时还离不开产业环境、信息环境等环境的支撑。

## 2) 成熟度等级

智慧城市的建设与发展过程是一个逐渐迈向更高成熟状态的长期渐进过程，是基础设施、信息资源、网络安全、体制机制、惠民服务、城市治理等各方面能力的持续提升过程。结合成熟度理论和方法，可以构建智慧城市成熟度，实现从动态评价的角度来对城市智慧化的发展阶段进行衡量。依托科学发展规律，可将智慧城市发展成熟度划分为规划级、管理级、协同级、优化级、引领级 5 个等级，如图 1-8 所示。

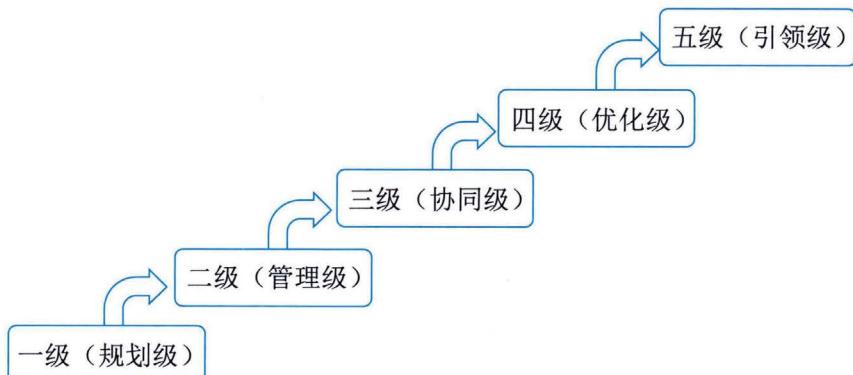


图 1-8 智慧城市成熟度等级参考模型

- 一级（规划级）：应围绕智慧城市的发展进行策划，明确相关职责分工和工作机制等，初步开展数据采集和应用，确保相关活动有序开展。
- 二级（管理级）：应明确智慧城市发展战略、原则、目标和实施计划等，推进城市基础

设施智能化改造，多领域实现信息系统单项应用，对智慧城市全生命周期实施管理。

- 三级（协同级）：应管控智慧城市各项发展目标，实施多业务、多层次、跨领域应用系统的集成，持续推进信息资源的共享与交换，推动惠民服务、城市治理、生态宜居、产业发展等的融合创新，实现跨领域的协同改进。
- 四级（优化级）：应聚焦智慧城市与城市经济社会发展深度融合，基于数据与知识模型实施城市经济、社会精准化治理，推动数据要素的价值挖掘和开发利用，推进城市竞争力持续提升。
- 五级（引领级）：应构建智慧城市敏捷发展能力，实现城市物理空间、社会空间、信息空间的融合演进和共生共治，引领城市集群治理联动，形成高质量发展共同体。

### 3. 数字乡村

数字乡村是伴随网络化、信息化和数字化在农业农村经济社会发展中的应用，以及农民现代信息技能的提高而内生的农业农村现代化发展和转型进程，既是乡村振兴的战略方向，也是建设数字中国的重要内容。2019年中共中央办公厅、国务院办公厅印发《数字乡村发展战略纲要》指出：立足新时代国情农情，要将数字乡村作为数字中国建设的重要方面，加快信息化发展，整体带动和提升农业农村现代化发展。进一步解放和发展数字化生产力，注重构建以知识更新、技术创新、数据驱动为一体的乡村经济发展政策体系，注重建立层级更高、结构更优、可持续性更好的乡村现代化经济体系，注重建立灵敏高效的现代乡村社会治理体系，开启城乡融合发展和现代化建设新局面。

《数字乡村发展战略纲要》明确了到2035年，数字乡村建设取得长足进展。城乡“数字鸿沟”大幅缩小，农民数字化素养显著提升。农业农村现代化基本实现，城乡基本公共服务均等化基本实现，乡村治理体系和治理能力现代化基本实现，生态宜居的美丽乡村基本实现。到21世纪中叶，全面建成数字乡村，助力乡村全面振兴，全面实现农业强、农村美、农民富。

### 4. 数字生活

数字生活是依托互联网和一系列数字科技技术应用为基础的一种生活方式，可以方便快捷地带给人们更好的生活体验和工作便利。数字生活主要体现在如下方面。

(1) 生活工具数字化。数字化生活时代，信息技术和产品成为极其重要的生活工具，人们将像享受空气、阳光、水一样享受数字化生活工具带来的舒适和便捷。根据摩尔定律和梅特卡夫定律，随着技术的不断创新与广泛扩散，其应用成本将显著下降，而其价值则显著增加。

(2) 生活方式数字化。在数字社会中，借助于数字化技术，每个人的工作、学习、消费、交往、娱乐等各种活动方式都将具有典型的数字化特征，数字家庭成为未来家庭的发展趋势。体现在工作更加弹性化和自主化；终身学习与随时随地学习成为可能；网络购物跻身主流消费方式；人际交往的范围与空间无限扩大等。

(3) 生活内容数字化。数字生活时代，人们工作、学习、消费和娱乐的内容具有典型的数字化特征。生活内容数字化体现在作品内容以创造、处理和分配信息为主；学习内容个性化；信息成为重要消费内容；娱乐内容数字化等。

## 1.4.4 数字生态

随着新一代信息技术创新和迭代速度的明显加快，其在提高社会生产力、优化资源配置等方面的作用日益凸显。营造良好数字生态，有利于充分激发数字技术的创新活力、要素潜能、发展空间，引领和驱动经济结构调整、产业发展升级、消费需求增长、治理格局优化，为加快建设数字经济、数字社会、数字政府提供良好环境和有力支撑。特别要看到，世界主要国家均把信息化作为国家战略重点和优先发展方向，通过优化数字生态加快推动数字化转型发展。《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》指出“坚持放管并重，促进发展与规范管理相统一，构建数字规则体系，营造开放、健康、安全的数字生态”。

### 1. 数据要素市场

随着数字经济的快速发展，数据作为数字经济的关键要素，对我国经济高质量发展的重要作用日益凸显。数据作为生产要素参与生产，需要进行市场化配置，形成生产要素价格及其体系。数据要素价格体系的建立，又是建立在数据所有制基础上的。因此谁掌握数据资产，在一定程度上就可以影响体系建立。数据作为新型生产要素，具有劳动工具和劳动对象的双重属性。首先数据作为劳动对象，通过采集、加工、存储、流通、分析环节，具备了价值和使用价值；其次，数据作为劳动工具，通过融合应用能够提升生产效能，促进生产力发展。

数据要素市场就是将尚未完全由市场配置的数据要素转向由市场配置的动态过程，其目的是形成以市场为根本调配机制，实现数据流动的价值或者数据在流动中产生价值。数据要素市场化配置是一种结果，而不是手段。数据要素市场化配置是建立在明确的数据产权、交易机制、定价机制、分配机制、监管机制、法律范围等保障制度的基础上。未来数据要素市场的发展，需要不断动态调整以上保障制度，最终形成数据要素的市场化配置。

### 2. 数字营商环境

良好的营商环境是一个国家或地区经济软实力和综合竞争力的重要体现。市场化、法治化、国际化、便利化的营商环境是一个国家、一个地区经济社会高质量发展的重要因素。随着数字经济蓬勃发展，与传统的营商环境相对应，数字经济时代的新型营商环境成为广泛关注的议题。2022年1月发布的《“十四五”数字经济发展规划》提出，要更加优化数字营商环境，加速弥合数字鸿沟。

国家工业信息安全发展研究中心2021年12月提出的全球数字营商环境评价指标体系。该评价体系包含5个一级指标：①数字支撑体系，包含普遍接入、智慧物流设施、电子支付设施；②数据开发利用与安全，包含公共数据开放、数据安全；③数字市场准入，包含数字经济业态市场准入、政务服务便利度；④数字市场规则，包含平台企业责任、商户权利与责任、数字消费者保护；⑤数字创新环境，包含数字创新生态、数字素养与技能、知识产权保护。

### 3. 网络安全保护

随着《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》《关键信息基础设施安全保护条例》等法律法规的颁布，以及网络安全等级保护2.0标准体系的发布，使我国网络安全法律法规和制度标准更加健全。但在百年变局和世纪疫情交织叠加，国际环境日趋复杂，网络霸权主义对世界和平与发展构成威胁，全球产业链供应链

遭受冲击，网络空间安全面临的形势持续复杂多变。网络空间对抗趋势更加突出，大规模针对性网络攻击行为增加，安全漏洞、数据泄露、网络诈骗等风险增加。针对这些问题需要坚持总体国家安全观和正确的网络安全观，贯彻新发展理念，构建网络安全新格局，全面加强网络安全保障体系和能力建设。

强大的网络安全产业实力是保障网络空间安全的根本和基石。习近平总书记多次就网络安全产业作出重要指示，强调“要坚持网络安全教育、技术、产业融合发展，形成人才培养、技术创新、产业发展的良性生态”，为网络安全事业高质量发展指明方向，并提供根本遵循。

## 1.5 数字化转型与元宇宙

随着众多信息通信新技术的迅速发展与普及应用，信息空间成长为第三空间，并与物理空间和社会空间共同构成人类社会的三元空间。新一轮科技革命与产业革命交互演进，面向组织的战略发展、业务模式、生产管理、运行管理等全方位的数字化转型，已成为数字经济时代广大组织的必选题。以云计算、大数据、人工智能等为代表的新一代信息技术发展迅猛，成为驱动组织数字化转型的关键要素。组织需要通过深化应用数字技术，打造敏捷、韧性、创新的数字化能力，重构传统业务流程和价值链，推动实现全要素、全链条、全层级的数字化转型。随着各领域数字化转型的发展和持续深化，元宇宙这一新概念也随之流行。元宇宙本质上是对现实世界的虚拟化、数字化过程，需要对内容生产、经济系统、用户体验以及实体世界内容等进行大量改造。

### 1.5.1 数字化转型

数字化转型（Digital Transformation）是建立在数字化转换（Digitization）、数字化升级（Digitalization）基础上，进一步触及组织核心业务，以新建一种业务模式为目标的高层次转型。数字化转型是开发数字化技术及支持能力以新建一个富有活力的数字化商业模式，只有组织对其业务进行系统性、彻底的（或重大和完全的）重新定义，而不仅仅是IT，而是对组织活动、流程、业务模式和员工能力的方方面面进行重新定义的时候，成功才会得以实现。

#### 1. 驱动因素

从全球视角来看，当前国际社会主要矛盾聚焦在发达国家企图垄断市场、资源和技术与发展中国家的发展愿望之间的矛盾。发达国家生产力没有飞跃式发展（第四次科技革命姗姗来迟），世界范围内市场、资源开发程度越来越充分，众多发展中国家想进一步改善人民生活，进一步参与到世界市场和资源的竞争中。纵观历史，无论是国际竞争关系、产业转型升级和新经济发展，还是当前我国社会主要矛盾变化带来的新特征、新要求，都有其发展规律和演进范式，即“生产力飞跃、生产要素变化、信息传播效率突破和社会‘智慧主体’规模扩容的叠加，将会促使人类社会生产关系的创新变革，最终引发经济与民生的深层发展”。这个范式驱动完成了原始经济到农业经济，再到工业经济的转型过程，同样会驱动工业经济向数字经济的转型。

##### 1) 生产力飞升：第四次科技革命

科学技术是第一生产力。近代人类发展过程中，已经完成了三次科技革命，正在经历第

四次科技革命，每次科技革命都对应一个科学范式，其深刻影响着世界格局的变化，是人类社会发展的根本动力，也是国际社会主要矛盾的发源地。

第一科学范式为经验范式。它偏重于经验事实的描述和明确具体的实用性的科学的研究范式。在研究方法上以归纳为主，带有较多盲目性的观测和实验。第二科学范式为理论范式。它主要指偏重理论总结和理性概括，强调较高普遍的理论认识而非直接实用意义的科学的研究范式。第三科学范式为模拟范式。它是一个与数据模型构建、定量分析方法以及利用计算机来分析和解决科学问题的研究范式。第四科学范式为数据密集型研究范式。它针对数据密集型科学，是由传统的假设驱动向基于科学数据进行探索的科学方法转变而生成的科学的研究范式。其研究方法是基于计算机生产实践产生的数据，按照驱动理论获得猜想与假设，完成数据自动化的计算和原理探索，即由计算机实施第一、第二、第三科学范式。第四范式通过新型信息技术的数据洞察，从大数据中自动化挖掘实践经验、理论原理并自行开展模拟仿真，完成基于数据的自决策和自优化，极大地繁荣应用科学技术。

### 2) 生产要素变化：数据要素的诞生

数据是与土地、劳动力、资本和技术并列的主要生产要素，表明数据将会是未来社会数字化、智能化发展的重要基础。数据是一项重要的经济资源，其对经济社会的全面持续发展、经济组织转型和参与个体生活质量非常重要且不可或缺。数据记载信息，信息融合知识，知识孕育智慧，过去人们已经持续了几十年的信息化建设，人们把智慧解构成知识，把知识分解为信息，把信息拆解为数据。随着人工智能、区块链和大数据等技术的出现，过去分散在各个环节的数据，重新归集为显性信息、知识和智慧，数据的经济价值越发凸显，因此数据对我国高质量发展的作用，与土地、设备、原材料、资本、劳动、技术同等重要，具备了单列为生产要素的现实条件。

### 3) 信息传播效率突破：社会互联网新格局

随着科学技术的发展，各种网络服务随之而来，互联网社交网络就是其中之一。人们的日常生活逐渐从现实社交网络转移到互联网虚拟社交网络中。互联网社交网络下，人们可以跟不在身边的朋友进行面对面的交流，还可以寻找有共同爱好的陌生人。从而形成在线社区，构成了庞大的社交网络平台，为用户提供便捷交流的渠道。

社交网络信息传输具有永生性、无限性、即时性以及方向性的特征。永生性指尽管在传播过程中可以控制信息，但它并不会被破坏或者消灭。比如：到一条信息且尚未传播该消息，但该消息实实在在地存在，信息的载体还可以继续传播。无限性是指信息可以像病毒一样无限地传播下去。即时性是社交网络信息传播的速度从通信器向接收者传播信息的时间大大缩短，甚至可以忽略。方向性意味着信息传播具有目的性，某些信息的传播仅是为了传递给特定的人。

随着互联网的发展，在互联网上传播信息已成为信息扩散的主要渠道。互联网的特性是信息可以跨越时间和地理障碍在网络上迅速传播。

### 4) 社会“智慧主体”规模：快速复制与“智能+”

过去，我们认为的“智慧主体”都是自然人，复制一个“智慧主体”的难度很大，需要教育、培育、培养等众多的手段方法。同时，其周期也较为漫长，培育一个自然人的“智慧主体”，往往需要超过20年的时间。漫长的复杂交互环境，以及自然环境因素等限制，都制约了

社会“智慧主体”规模的扩大与繁荣，从而使互联网的节点容量出现瓶颈，随着社会的进一步演进，这种瓶颈会阻碍人类社会的高质量建设，会影响人类社会的进一步发展和演进。

现在，社会的“智慧主体”已经不单纯是自然人，他可以是一个互联网账号、一台自动驾驶的汽车、一部智能手机，或者是工厂中的一套智能机器人。这些新兴“智慧主体”，具有不同于自然人的全量社会化活动模式，如消费选择等，但其在数据生产、数据开发利用、劳动力贡献和决策能力等方面，具备了自然人很多关键特征，在不知不觉中已经让这些新主体参与到了人们社会活动的方方面面，乃至与自然人享有同等的社会空间，如未来某一时刻无人驾驶的汽车主体与自然人道路参与主体享有同等的道路权。

新兴的“智慧主体”具备较强的可复制性、自我学习能力、更加广泛的连接能力和更加标准的交互手段等。新兴“智慧主体”规模和种类的快速扩张，会引发人类社会的深层次变革，改变自然人主体的劳动方式，劳动密集型的社会劳动逐步消退，智力密集型的社会劳动持续强化，自然人“智慧主体”甚至会全面退出生产制造过程领域，让自然人的竞争力聚焦在新兴“智慧主体”不会具备的领域。这个领域是以“服务”为典型代表，因为该领域会面对更加复杂的交互过程、更多的风险融合应对和情感因素管控等。

## 2. 基本原理

随着经济与社会的持续发展，同领域相关参与者因为数量的持续增多和发展水平趋于一致等，再加上我国处在中高速发展阶段，这些因素共同导致了经济与社会的竞争越来越充分、越来越激烈。随着我国社会主要矛盾从人民日益增长的物质文化同落后的社会生产之间的矛盾，转变为人民日益增长的美好生活需要和不平衡、不充分的发展之间的矛盾，以及信息时代带来的信息高效、充分且大规模传播，信息对象过程加速，乃至出现信息淹没等情况，这进一步加剧了经济与社会参与者的竞争，这表现在产品和服务的生命周期迭代越来越快，组织运行决策越来越高效，组织的转型升级周期越来越短，组织的业务发展越来越敏捷等。

传统发展视角下，组织为提升自身的竞争力，往往通过优化组织结构体系（如组织结构扁平化），提升工艺技术与装备（如应用新技术或自动化装备），降低业务成本（如人员容量、材料成本、加工成本等）等方式展开，这种优化与提升从某种程度上实现了对组织竞争力和竞争优势的保持和增强。这种发展模式下，组织通过治理和管理体系强化组织的协同性和创新力，并降低组织风险；通过减少客户个性选择驱动业务规模化发展，优化产品生产和服务交付成本。

数字经济时代，经济与社会竞争的进一步加剧，传统发展视角下的竞争力与竞争优势的保持和增强等方法，越来越难以支撑组织的发展需求，主要体现在：

- 决策瓶颈：以组织架构构建的治理与管理体系决策效率容易遇到瓶颈，并且组织规模越大、行政层级越多、决策效率效能越容易达到瓶颈。
- 变革制约：组织变革是一项系统工程，这不仅仅包括新组织、新工艺、新产品、新营销等的策划、规划和设计等，其部署落实也是一组复杂的工作，变革的效能常常受组织文化、人员技能、技术现状等方面制约，太多的变革一致性无法解决。
- 知识资产流失：组织研发或沉淀的各类经验，如使用传统的知识体系（如用文档资料管理），容易随着人员流动而流失，这是因为传统知识方法需要相关人员全部掌握。

- 需求响应延迟：组织为了有效地控制成本，最常用的方法是固化管理、工艺等，通过“简单可复制”的模式，达到一致性和成本最优化，这会导致组织对客户或服务对象的个性化需求延迟满足乃至放弃满足。

组织的数字化转型就是基于组织既有的治理与管理体系、工艺路径和产品技术、服务活动定义等，打造更加高效的决策效率、更灵活的工艺调度、更多元的产品与服务技术应用和更丰富的业务模式等。数字化转型需要组织结合信息技术的开发利用，对组织完成深层次变革，可参考模型如图 1-9 所示。



图 1-9 数字组织运行参考框架

### 1) 能力因子定义和数字化“封装”

实施数字化转型，组织需要把各项能力和活动进行清晰的结构化并定义，形成细化的可灵活调度和编排的能力因子，这些能力因子是有层次或可组合的，如能力域、能力子域、能力项、能力分项、能力子项等，对于数字化转型不同成熟度的组织来说，主要体现在能力因子定义颗粒度、学科性和有效性等方面。

能力因子的定义可驱动组织的管理精细化，更重要的是能够实现对这些能力因子的数字化“封装”，这种封装不只是对业务流程、工艺过程和技术内容的“包装”，而是需要向具体活动的人员、技术（含内部控制等）、资源、数据、流程（过程、动作）的模块化“封装”，打造基于数据的标准化输入与输出，形成类似信息化系统中的对象、类、模块等的组件。在工业类组织中体现为数字装备、数字化管理单元、数字产品等，目的是实现“智能+”。

### 2) 基于“互联网+”的调度和决策

实施数字化转型，组织需要在已有治理与管理体系、工艺体系、服务体系、产品体系的基础上，通过使用“互联网+”的模式，将组织沉淀的各类知识经验进行数字化提炼，形成数字算法、模型和框架等，满足信息系统能够理解和使用的方式，让调度和决策脱离“自然人”，从

而提高调度和决策效率及其科学性。这部分工作是数字化转型中一项持续性工作，其科技含量比较高，也是组织数字化转型中的难点，主要体现在：

- 业务融合：将知识经验形成数字化调用模式，需要业务和信息技术的充分融合，需要实施这些工作的业务人员具备一定的数字技能，或者信息技术人员能够深入理解业务。
- 持续坚持：通过数字模式开展决策与调度活动，开始时的效果、效率、效能并不一定理想，这就需要组织能够持续坚持，通过持续改进活动，提升数据模式的价值。
- 文化冲突：调度与决策的科学化、敏捷化，依赖组织的知识沉淀，这就需要组织解决文化冲突，引导成员适应数字化带来的各种变化，积极贡献知识经验，消除自我成长顾虑以及驾驭数字的“恐惧”等。
- 效果判别：通常情况下，治理和管理关注判断与决策的正确性，执行操作关注过程的精确性，而使用数字模式实施决策和调度时，其精确性被凸显出来，对决策和调度的数据及应用过程提出了更高的要求，需要组织投入更多的智力资源。

### 3) 转型控制

数字化转型往往不是指一个结果的表达，而是一个持续的过程，组织需要能够有效管控转型过程，无论是服务组织还是工业组织，都不能一蹴而就地完成转型升级。组织需要充分借鉴信息化与工业化、信息化与领域现代化等深度融合的最佳实践，结合自身的实际情况，持续建设、优化和改进数字化转型过程。

## 3. 智慧转移

数字化转型基本原理揭示了个体智慧（知识、技能和经验等）由“自然人”个体，转移到组织智慧（计算机、信息系统等掌握的）的必要性和重要性。这种“智慧转移”也称“智慧移植”，需要经历系列的过程才能完成，每个组织开展此类活动的模式与方法存在差异，也可以参考图 1-10 所示的“智慧转移”模型。

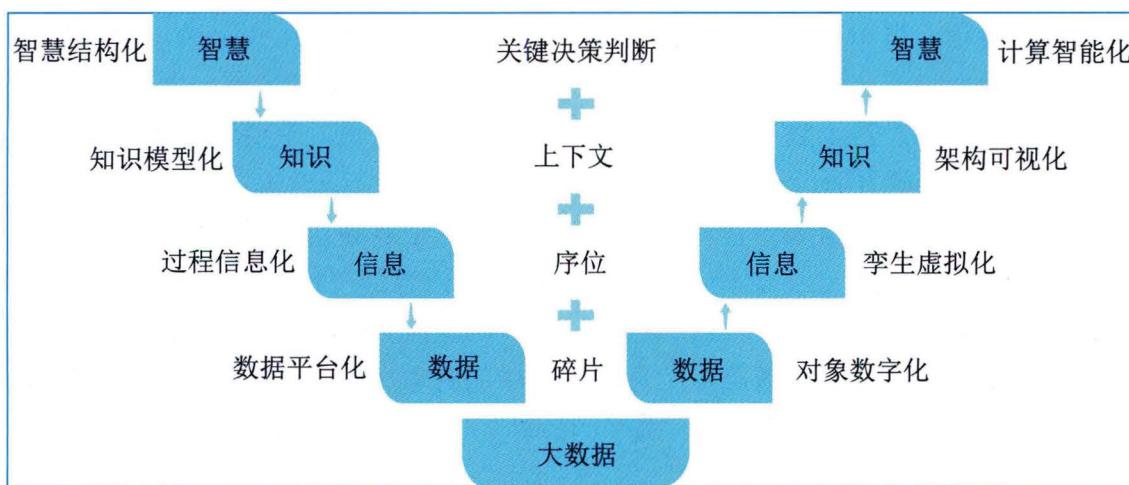


图 1-10 智慧转移的 S8D 模型

DIKW 模型很好地诠释了数据（Data）、信息（Information）、知识（Knowledge）和智慧

(Wisdom) 之间的关系，并揭示了他们的转化过程与方法。S8D 模型就是基于 DIKW 模型，构筑了“智慧—数据”“数据—智慧”两大过程的 8 个转化活动。

### 1) “智慧—数据” 过程

该过程通常指信息系统规划、建设、运行过程，也就是传统讲的“信息化过程”。该过程：①通过智慧结构化明确了业务体系层面的内容；②通过知识模型化定义业务活动的逻辑关系；③通过过程信息化（管理和工艺流程化）明确各执行操作系列要求；④通过数据平台化实现了数据的采集、存储和共享等。

### 2) “数据—智慧” 过程

该过程通常指数据的开发利用和资源管理的过程，即人们常说的“智慧化过程”，重点解决基于各类组织组成对象（人员、流程、业务、工艺、装备等）“数字关系”的“脑力替代”。该过程在大数据“筑底”后，多元化数据能够被开发利用：①通过对象数字化实现对各类对象的数字化表达；②通过孪生虚拟化完成物理对象到信息空间的映射；③通过架构可视化实现业务知识模型与经验沉淀的复用和创新；④通过计算智能化实现多元条件下的调度和决策。

数据是筑底构建可计算智慧的关键，通过“智慧—数据”过程将人类智慧形成了数据表达，并通过数据流动，提高了组织业务与工作效能，实现“体力劳动替代”。接下来通过“数据—智慧”工程，在大数据的基础上，逆变数据为计算智能，完成了智慧载体由自然人到计算机和信息系统的转移，其价值不仅仅可实现智慧的在线时效（ $7\times 24$  无休），更可以实现“智慧挤压”（多方法多维度综合判断）和更高级别的“智慧萃取”（新智慧的生成），进一步实现智慧的可复制。这一过程也是第四科学范式的基本框架，是第四次科技革命的触发逻辑。

## 4. 持续迭代

组织数字化转型需要在能力因子不断细化的基础上，针对能力因子的数字化转型实施迭代，可类比为整体数字化转型与局部数字转型的关系。组织每个能力因子数字化“封装”的持续迭代主要包含四项活动，即：信息物理世界（也称数字孪生，CPS）建设、决策能力边际化（Power to Edge, PtoE）部署、科学社会物理赛博机制构筑（Cyber-Physical-Social Systems, CPSS）、数字框架与信息调制（Digital Frame and Information Modulation, DFIM），如图 1-11 所示。针对能力因子的持续迭代可以从任何一项活动开始实施四项活动，形成持续迭代闭环。

### 1) 信息物理世界建设

针对能力因子中的各类对象，实施数字孪生建设，并在此基础上加入该因子与其他因子之间的配置关系。组织可以通过该项活动实现能力因子相关数据挖掘与数据开发利用，从而发现新的技术和逻辑，提升各项工作效率。

### 2) 决策能力边际化部署

决策能力边际化是指处置执行层面的装置和人员能够基于决策算法模型等，敏捷取得更高的决策能力（权），达到敏捷响应的效果。组织可以通过该项活动实现决策权融合与调制，达到装备智能化和提高决策效能的价值效果。

### 3) 科学物理赛博机制构筑

科学物理赛博机制构筑是在CPS的基础上，汇聚组织内能力因子的环境因素力量（或组织维度的外部社会力量），建设高密度数据框架，参照社会运行原理，封装、解构和重构各能力因子协同关系。组织可以通过该项活动实现对各能力因子的灵活组合机制，形成能够面对各类需求的动态调度能力。

### 4) 数据框架与信息调制设计

组织能力因子的数字密度越高，对其可控性就越高，对应的安全可靠性也越高。组织通过优化能力因子的数字框架模型，并提升数据采集获取的精准度和及时性，能够有效地提升组织对能力因子的应急与动员能力，从而具备更加可靠的已知风险管控能力和未知风险的应对能力。

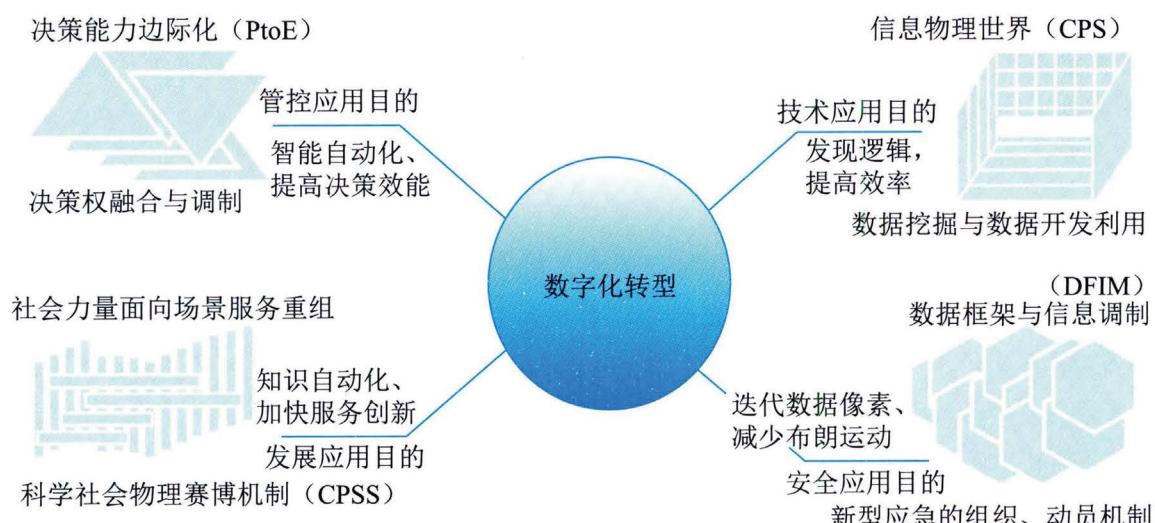


图 1-11 组织能力数字化转型及持续迭代参考模型 (CPSD 模型)

## 1.5.2 元宇宙

元宇宙 (Metaverse) 是一个新兴概念，是一大批技术的集成。北京大学陈刚教授对元宇宙的定义是：元宇宙是利用科技手段进行链接与创造的，与现实世界映射与交互的虚拟世界，具备新型社会体系的数字生活空间。清华大学沈阳教授对元宇宙的定义是：元宇宙是整合多种新技术而产生的新型虚实相融的互联网应用和社会形态，它基于扩展现实技术提供沉浸式体验，以及数字孪生技术生成现实世界的镜像，通过区块链技术搭建经济体系，将虚拟世界与现实世界在经济系统、社交系统、身份系统上密切融合，并且允许每个用户进行内容生产和编辑。中国社会科学院学者左鹏飞从时空性、真实性、独立性、连接性四个方面去交叉定义元宇宙。他指出：从时空性来看，元宇宙是一个空间维度上虚拟而时间维度上真实的数字世界；从真实性来看，元宇宙中既有现实世界的数字化复制物，也有虚拟世界的创造物；从独立性来看，元宇宙是一个与外部真实世界既紧密相连，又高度独立的平行空间；从连接性来看，元宇宙是一个把网络、硬件终端和用户囊括进来的一个永续的、广覆盖的虚拟现实系统。

## 1. 主要特征

随着虚拟现实、人工智能、数字孪生、云计算等关键技术逐步迭代发展，用户对更沉浸的虚拟世界有了更深入、更丰富的需求。元宇宙的流行是互联网发展到了一定的高度，也可以认为是互联网发展的另一阶段。元宇宙的主要特征包括：

- 沉浸式体验：元宇宙的发展主要基于人们对互联网体验的需求，这种体验就是即时信息基础上的沉浸式体验。
- 虚拟身份：人们已经拥有大量的互联网账号，未来人们在元宇宙中，随着账号内涵和外延的进一步丰富，将会发展成为一个或若干个数字身份，这种身份就是数字世界的一个或一组角色。
- 虚拟经济：虚拟身份的存在就促使元宇宙具备了开展虚拟社会活动的能力，而这些活动需要一定的经济模式展开，即虚拟经济。
- 虚拟社会治理：元宇宙中的经济与社会活动也需要一定的法律法规和规则的约束，就像现实世界一样，元宇宙也需要社区化的社会治理。

总之，元宇宙作为现实世界的孪生空间和虚拟世界，其物理属性淡化，但社会属性将会被强化，我们在现实社会中的大量特征和活动，都逐渐会在元宇宙中体现出来。

## 2. 发展演进

元宇宙作为多技术的集成融合和现实世界虚拟化，其发展一方面受到各类技术创新、发展和演进的影响，另一方面受经济与社会发展进程的约束。从互联网发展的基本规律和数字化转型进程来看，元宇宙首先会在社交、娱乐和文化领域发展，形成虚拟“数字人”，逐步再向虚拟身份方向演进，形成“数字人生”，此时的元宇宙偏向个体用户需求。但随着元宇宙中虚拟经济的发展和现实中组织数字化转型的深入，元宇宙向“数字组织”领域延伸，从而影响现实世界的经济与社会发展整体数字化转型升级，形成“数字生态”。之后伴随相关法律法规、标准规范的生成，网信事业的发展以及网络文明的进一步完善，元宇宙的虚拟世界形态持续迭代，形成“数字社会治理”，实现物理空间、社会空间和信息空间三元空间的协同发展新格局。

## 1.6 本章练习

### 1. 选择题

(1) 下列说法正确的是\_\_\_\_\_。

- A. 信息只存在家庭中      B. 信息只存在图书馆中  
C. 信息只存在校园里      D. 信息无处不在

参考答案: D

(2) 信息的基础是\_\_\_\_\_。

- A. 数据      B. 知识      C. 事实      D. 概念

参考答案: A

(3) 工业互联网的体系不包括\_\_\_\_\_。

- A. 网络
- B. 平台
- C. 技术
- D. 安全

参考答案: C

(4) 支撑科学研究、技术开发、产品研制的具有公益属性的基础设施属于\_\_\_\_\_。

- A. 信息基础设施
- B. 融合基础设施
- C. 创新基础设施
- D. 网络基础设施

参考答案: C

(5) GB/T 39116《智能制造能力成熟度模型》规定了企业智能制造能力在不同阶段应达到的水平。若企业应对装备、系统等开展集成，实现跨业务活动间的数据共享，则该企业属于\_\_\_\_\_水平。

- A. 一级（规划级）
- B. 二级（规范级）
- C. 三级（集成级）
- D. 四级（优化级）

参考答案: C

(6) 《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》中从数字经济、数字政府、数字社会、\_\_\_\_\_四个维度出发勾勒了建设数字中国的宏伟蓝图。

- A. 数字生态
- B. 数字技术
- C. 数字服务
- D. 数字人才

参考答案: A

(7) \_\_\_\_\_不属于智慧城市核心能力要素。

- A. 数据治理、边际决策、多元融合
- B. 数据治理、数字孪生、边际决策
- C. 数据管理、数字孪生、态势感知
- D. 数字孪生、多元融合、态势感知

参考答案: C

(8) 智慧城市发展过程中，能够明确智慧城市发展战略、原则、目标和实施计划等，推进城市基础设施的智能化改造，多领域实现信息系统单项应用，对智慧城市全生命周期实施管理，则该智慧城市成熟度处于\_\_\_\_\_水平。

- A. 规划级
- B. 管理级
- C. 协同级
- D. 优化级

参考答案: B

## 2. 思考题

(1) 请列出大数据的特点，并简要叙述大数据有哪些重要应用领域。

参考答案: 略

(2) 请简述产业数字化和数字化的关系及其意义。

参考答案: 略

(3) 请简述数字中国的时代特征及数字中国建设的重大意义。

参考答案: 略

## 第2章 信息技术发展

信息技术是在信息科学的基本原理和方法下，获取信息、处理信息、传输信息和使用信息的应用技术总称。从信息技术的发展过程来看，信息技术在传感器技术、通信技术和计算机技术的基础上，融合创新和持续发展，孕育和产生了物联网、云计算、大数据、区块链、人工智能和虚拟现实等新一代信息技术，成为支撑当今经济活动和社会生活的基石，代表着当今先进生产力的发展方向。

从宏观上讲，信息技术与信息化、信息系统密不可分。信息技术是实现信息化的手段，是信息系统建设的基础。信息化的需求驱动信息技术高速发展，信息系统的广泛应用促进了信息技术的迭代创新。近年来，随着新一代信息技术的发展，信息及其相关的数据成为重要生产要素和战略资源、使得人们能更高效地进行资源优化配置，持续推动传统产业不断升级、社会劳动生产率的不断提升，从而带动全球信息化发展和数字化转型，新一代信息技术已成为世界各国投资和重点发展的战略性产业。

### 2.1 信息技术及其发展

信息技术是以微电子学为基础的计算机技术和电信技术的结合而形成的，对声音的、图像的、文字的、数字的和各种传感信号的信息进行获取、加工、处理、存储、传播和使用的技术。按表现形态的不同，信息技术可分为硬技术（物化技术）与软技术（非物质化技术）。前者指各种信息设备及其功能，如传感器、服务器、智能手机、通信卫星、笔记本电脑。后者指有关信息获取与处理的各种知识、方法与技能，如语言文字技术、数据统计分析技术、规划决策技术、计算机软件技术等。

#### 2.1.1 计算机软硬件

计算机硬件（Computer Hardware）是指计算机系统中由电子、机械和光电元件等组成的各种物理装置的总称。这些物理装置按系统结构的要求构成一个有机整体，为计算机软件运行提供物质基础。计算机软件（Computer Software）是指计算机系统中的程序及其文档，程序是计算任务的处理对象和处理规则的描述；文档是为了便于了解程序所需的阐明性资料。程序必须安装入机器内部才能工作，文档一般是给人看的，不一定安装入机器。

硬件和软件互相依存。硬件是软件赖以工作的物质基础，软件的正常工作是硬件发挥作用的重要途径。计算机系统必须要配备完善的软件系统才能正常工作，从而充分发挥其硬件的各种功能。硬件和软件协同发展，计算机软件随硬件技术的迅速发展而发展，而软件的不断发展与完善又促进硬件的更新，两者密切交织发展，缺一不可。随着计算机技术的发展，在许多情况下，计算机的某些功能既可以由硬件实现，也可以由软件来实现。因此硬件与软件在一定意

义上来说没有绝对严格的界线。

## 2.1.2 计算机网络

在计算机领域中，网络就是用物理链路将各个孤立的工作站或主机相连在一起，组成数据链路，从而达到资源共享和通信的目的。凡将地理位置不同，并具有独立功能的多个计算机系统通过通信设备和线路连接起来，且以功能完善的网络软件（网络协议、信息交换方式及网络操作系统等）实现网络资源共享的系统，均可称为计算机网络。从网络的作用范围可将网络类别划分为个人局域网（Personal Area Network, PAN）、局域网（Local Area Network, LAN）、城域网（Metropolitan Area Network, MAN）、广域网（Wide Area Network, WAN）、公用网（Public Network）、专用网（Private Network）。

### 1. 网络标准协议

网络协议是为计算机网络中进行数据交换而建立的规则、标准或约定的集合。网络协议由三个要素组成，分别是语义、语法和时序。语义是解释控制信息每个部分的含义，它规定了需要发出何种控制信息，完成的动作以及做出什么样的响应；语法是用户数据与控制信息的结构与格式，以及数据出现的顺序；时序是对事件发生顺序的详细说明。人们形象地将这三个要素描述为：语义表示要做什么，语法表示要怎么做，时序表示做的顺序。

#### 1) OSI

国际标准化组织（ISO）和国际电报电话咨询委员会（CCITT）联合制定的开放系统互连参考模型（Open System Interconnect, OSI），其目的是为异种计算机互连提供一个共同的基础和标准框架，并为保持相关标准的一致性和兼容性提供共同的参考。OSI 采用了分层的结构化技术，从下到上共分物理层、数据链路层、网络层、传输层、会话层、表示层和应用层。

广域网协议是在 OSI 参考模型的最下面三层操作，定义了在不同的广域网介质上的通信。广域网协议主要包括：PPP 点对点协议、ISDN 综合业务数字网、xDSL（DSL 数字用户线路的统称：HDSL、SDSL、MVL、ADSL）、DDN 数字专线、x.25、FR 帧中继、ATM 异步传输模式。

#### 2) IEEE 802 协议族

IEEE 802 规范定义了网卡如何访问传输介质（如光缆、双绞线、无线等），以及如何在传输介质上传输数据的方法，还定义了传输信息的网络设备之间连接的建立、维护和拆除的途径。遵循 IEEE 802 标准的产品包括网卡、桥接器、路由器以及其他一些用来建立局域网络的组件。IEEE 802 规范包括：802.1（802 协议概论）、802.2（逻辑链路控制层 LLC 协议）、802.3（以太网的 CSMA/CD 载波监听多路访问 / 冲突检测协议）、802.4（令牌总线 Token Bus 协议）、802.5（令牌环 Token Ring 协议）、802.6（城域网 MAN 协议）、802.7（FDDI 宽带技术协议）、802.8（光纤技术协议）、802.9（局域网上的语音 / 数据集成规范）、802.10（局域网安全互操作标准）、802.11（无线局域网 WLAN 标准协议）。

#### 3) TCP/IP

Internet 是一个包括成千上万相互协作的组织和网络的集合体。TCP/IP 是 Internet 的核心。

TCP/IP 在一定程度上参考了 OSI，它将 OSI 的七层简化为四层：①应用层、表示层、会话层三个层次提供的服务相差不是很大，所以在 TCP/IP 中，它们被合并为应用层一个层次。②由于传输层和网络层在网络协议中的地位十分重要，所以在 TCP/IP 中它们被作为独立的两个层次。③因为数据链路层和物理层的内容相差不多，所以在 TCP/IP 中它们被归并到网络接口层一个层次里。

在应用层中，定义了很多面向应用的协议，应用程序通过本层协议利用网络完成数据交互的任务。这些协议主要有 FTP（File Transfer Protocol，文件传输协议）、TFTP（Trivial File Transfer Protocol，简单文件传输协议）、HTTP（Hypertext Transfer Protocol，超文本传输协议）、SMTP（Simple Mail Transfer Protocol，简单邮件传输协议）、DHCP（Dynamic Host Configuration Protocol，动态主机配置协议）、Telnet（远程登录协议）、DNS（Domain Name System，域名系统）、SNMP（Simple Network Management Protocol，简单网络管理协议）等。

传输层主要有两个传输协议，分别是 TCP 和 UDP（User Datagram Protocol，用户数据报协议），这些协议负责提供流量控制、错误校验和排序服务。

网络层中的协议主要有 IP、ICMP（Internet Control Message Protocol，网际控制报文协议）、IGMP（Internet Group Management Protocol，网际组管理协议）、ARP（Address Resolution Protocol，地址解析协议）和 RARP（Reverse Address Resolution Protocol，反向地址解析协议）等，这些协议处理信息的路由和主机地址解析。

由于网络接口层兼并了物理层和数据链路层，所以网络接口层既是传输数据的物理媒介，也可以为网络层提供一条准确无误的线路。

## 2. 软件定义网络

软件定义网络（Software Defined Network，SDN）是一种新型网络创新架构，是网络虚拟化的一种实现方式，它可通过软件编程的形式定义和控制网络，其通过将网络设备的控制面与数据面分离开来，从而实现了网络流量的灵活控制，使网络变得更加智能，为核心网络及应用的创新提供了良好的平台。

利用分层的思想，SDN 将数据与控制相分离。在控制层，包括具有逻辑中心化和可编程的控制器，可掌握全局网络信息，方便运营商和科研人员管理配置网络和部署新协议等。在数据层，包括哑交换机（与传统的二层交换机不同，专指用于转发数据的设备），仅提供简单的数据转发功能，可以快速处理匹配的数据包，适应流量日益增长的需求。两层之间采用开放的统一接口（如 OpenFlow 等）进行交互。控制器通过标准接口向交换机下发统一标准规则，交换机仅需按照这些规则执行相应的动作即可。SDN 打破了传统网络设备的封闭性。此外，南北向和东西向的开放接口及可编程性，也使得网络管理变得更加简单、动态和灵活。

SDN 的整体架构由下到上（由南到北）分为数据平面、控制平面和应用平面，具体如图 2-1 所示。其中，数据平面由交换机等网络通用硬件组成，各个网络设备之间通过不同规则形成的 SDN 数据通路连接；控制平面包含了逻辑上为中心的 SDN 控制器，它掌握着全局网络信息，负责各种转发规则的控制；应用平面包含着各种基于 SDN 的网络应用，用户无须关心底层细节就可以编程、部署新应用。

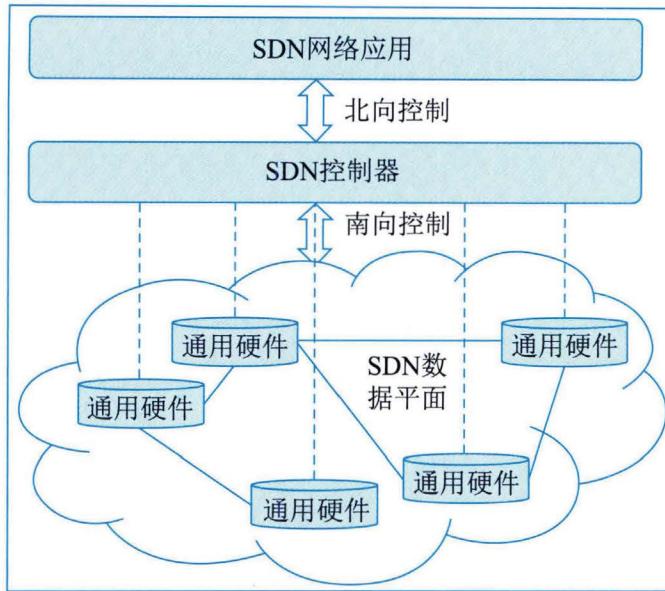


图 2-1 SDN 体系架构图

控制平面与数据平面之间通过 SDN 控制数据平面接口（Control-Data-Plane Interface, CDPI）进行通信，它具有统一的通信标准，主要负责将控制器中的转发规则下发至转发设备，最主要应用的是 OpenFlow 协议。控制平面与应用平面之间通过 SDN 北向接口（NorthBound Interface, NBI）进行通信，而 NBI 并非统一标准，它允许用户根据自身需求定制开发各种网络管理应用。

SDN 中的接口具有开放性，以控制器为逻辑中心，南向接口负责与数据平面进行通信，北向接口负责与应用平面进行通信，东西向接口负责多控制器之间的通信。最主流的南向接口 CDPI 采用的是 OpenFlow 协议。OpenFlow 最基本的特点是基于流（Flow）的概念来匹配转发规则，每一个交换机都维护一个流表（Flow Table），依据流表中的转发规则进行转发，而流表的建立、维护和下发都是由控制器完成的。针对北向接口，应用程序通过北向接口编程来调用所需的各种网络资源，实现对网络的快速配置和部署。东西向接口使控制器具有可扩展性，为负载均衡和性能提升提供了技术保障。

### 3. 第五代移动通信技术

第五代移动通信技术（5th Generation Mobile Communication Technology, 5G）是具有高速率、低时延和大连接特点的新一代移动通信技术。

国际电信联盟（ITU）定义了 5G 的八大指标，与 4G 的对比如表 2-1 所示。

表 2-1 4G 与 5G 主要指标对标

指标名称	流量密度 / (Tb/s · km <sup>2</sup> )	连接数密度 / (万 · km <sup>-2</sup> )	时延 / ms	移动性 / (km · h <sup>-1</sup> )	能效 / 倍	用户体验速率 / b · s <sup>-1</sup>	频道效率 / 倍	峰值速率 / Gb · s <sup>-1</sup>
4G	0.1	10	空口 10	350	1	10M	1	10
5G	10	100	空口 1	500	100	0.1 ~ 1G	3	20

5G 国际技术标准重点满足灵活多样的物联网需要。在正交频分多址（Orthogonal Frequency

Division Multiple Access, OFDMA) 和多入多出 (Multiple Input Multiple Output, MIMO) 基础技术上, 5G 为支持三大应用场景, 采用了灵活的全新系统设计。在频段方面, 与 4G 支持中低频不同, 考虑到中低频资源有限, 5G 同时支持中低频和高频频段, 其中中低频满足覆盖和容量需求, 高频满足在热点区域提升容量的需求, 5G 针对中低频和高频设计了统一的技术方案, 并支持百 MHz 的基础带宽。为了支持高速率传输和更优覆盖, 5G 采用 LDPC (一种具有稀疏校验矩阵的分组纠错码)、Polar (一种基于信道极化理论的线性分组码) 新型信道编码方案、性能更强的大规模天线技术等。为了支持低时延、高可靠, 5G 采用短帧、快速反馈、多层 / 多站数据重传等技术。

国际电信联盟 (ITU) 定义了 5G 的三大类应用场景, 即增强移动宽带 (eMBB)、超高可靠低时延通信 (uRLLC) 和海量机器类通信 (mMTC)。增强移动宽带主要面向移动互联网流量爆炸式增长, 为移动互联网用户提供更加极致的应用体验; 超高可靠低时延通信主要面向工业控制、远程医疗、自动驾驶等对时延和可靠性具有极高要求的垂直行业应用需求; 海量机器类通信主要面向智慧城市、智能家居、环境监测等以传感和数据采集为目标的应用需求。

### 2.1.3 存储和数据库

#### 1. 存储技术

存储分类根据服务器类型分为: 封闭系统的存储和开放系统的存储。封闭系统主要指大型机等服务器。开放系统指基于包括麒麟、欧拉、UNIX、Linux 等操作系统的服务器。开放系统的存储分为: 内置存储和外挂存储。外挂存储根据连接的方式分为直连式存储 (Direct-Attached Storage, DAS) 和网络化存储 (Fabric-Attached Storage, FAS)。网络化存储根据传输协议又分为网络接入存储 (Network-Attached Storage, NAS) 和存储区域网络 (Storage Area Network, SAN)。DAS、NAS、SAN 等存储模式之间的技术与应用对比如表 2-2 所示。

表 2-2 常用存储模式的技术与应用对比

存储系统架构	DAS	NAS	SAN
安装难易度	不一定	简单	困难
数据传输协议	SCSI/FC/ATA	TCP/IP	FC
传输对象	数据块	文件	数据块
使用标准文件共享协议	否	是 (NFS/CIFS...)	否
异种操作系统文件共享	否	是	需要转换设备
集中式管理	不一定	是	需要管理工具
管理难易度	不一定	以网络为基础, 容易	不一定, 但通常很难
提高服务器效率	否	是	是
灾难忍受度	低	高	高, 专有方案
适合对象	中小组织服务器 捆绑磁盘 (JBOD)	中小组织 SOHO 族 组织部门	大型组织 数据中心

(续表)

存储系统架构	DAS	NAS	SAN
应用环境	局域网 文档共享程度低 独立操作平台 服务器数量少	局域网 文档共享程度高 异质格式存储需求高	光纤通道存储域网 网络环境复杂 文档共享程度高 异质操作系统平台 服务器数量多
业务模式	一般服务器	Web 服务器 多媒体资料存储 文件资料共享	大型资料库 数据库等
档案格式复杂度	低	中	高
容量扩充能力	低	中	高

存储虚拟化（Storage Virtualization）是“云存储”的核心技术之一，它把来自一个或多个网络的存储资源整合起来，向用户提供一个抽象的逻辑视图，用户可以通过这个视图中的统一逻辑接口来访问被整合的存储资源。用户在访问数据时并不知道真实的物理位置。它带给人们直接的好处是提高了存储利用率，降低了存储成本，简化了大型、复杂、异构的存储环境的管理工作。

存储虚拟化使存储设备能够转换为逻辑数据存储。虚拟机作为一组文件存储在数据存储的目录中。数据存储是类似于文件系统的逻辑容器。它隐藏了每个存储设备的特性，形成一个统一的模型，为虚拟机提供磁盘。存储虚拟化技术帮助系统管理虚拟基础架构存储资源，提高资源利用率和灵活性，提高应用正常运行时间。

绿色存储（Green Storage）技术是指从节能环保的角度出发，用来设计生产能效更佳的存储产品，降低数据存储设备的功耗，提高存储设备每瓦性能的技术。绿色存储是一个系统设计方案，贯穿于整个存储设计过程，包含存储系统的外部环境、存储架构、存储产品、存储技术、文件系统和软件配置等多方面因素。

绿色存储技术的核心是设计运行温度更低的处理器和更有效率的系统，生产更低能耗的存储系统或组件，降低产品所产生的电子碳化合物，其最终目的是提高所有网络存储设备的能源效率，用最少的存储容量来满足业务需求，从而消耗最低的能源。以绿色理念为指导的存储系统最终是存储容量、性能、能耗三者的平衡。

绿色存储技术涉及所有存储分享技术，包括磁盘和磁带系统、服务器连接、存储设备、网络架构及其他存储网络架构、文件服务和存储应用软件、重复数据删除、自动精简配置和基于磁带的备份技术等可以提高存储利用率、降低建设成本和运行成本的存储技术，其目的是提高所有网络存储技术的能源效率。

## 2. 数据结构模型

数据结构模型是数据库系统的核心。数据结构模型描述了在数据库中结构化和操纵数据的方法，模型的结构部分规定了数据如何被描述（例如树、表等）。模型的操纵部分规定了数据的添加、删除、显示、维护、打印、查找、选择、排序和更新等操作。

常见的数据结构模型有三种：层次模型、网状模型和关系模型，层次模型和网状模型又统称为格式化数据模型。

### 1) 层次模型

层次模型是数据库系统最早使用的一种模型，它用“树”结构表示实体集之间的关联，其中实体集（用矩形框表示）为结点，而树中各结点之间的连线表示它们之间的关联。在层次模型中，每个结点表示一个记录类型，记录类型之间的联系用结点之间的连线（有向边）表示，这种联系是父子之间的一对多的联系。这就使得层次数据库系统只能处理一对多的实体联系。每个记录类型可包含若干个字段，这里记录类型描述的是实体，字段描述实体的属性。每个记录类型及其字段都必须命名。各个记录类型、同一记录类型中各个字段不能同名。每个记录类型可以定义一个排序字段，也称码字段，如果定义该排序字段的值是唯一的，则它能唯一地标识一个记录值。

一个层次模型在理论上可以包含任意有限个记录类型和字段，但任何实际的系统都会因为存储容量或实现复杂度而限制层次模型中包含的记录类型个数和字段个数。在层次模型中，同一双亲的子女结点称为兄弟结点，没有子女结点的结点称为叶结点。层次模型的一个基本的特点是任何一个给定的记录值只能按其层次路径查看，没有一个子女记录值能够脱离双亲记录值而独立存在。

### 2) 网状模型

网状数据库系统采用网状模型作为数据的组织方式。网状模型用网状结构表示实体类型及其实体之间的联系。网状模型是一种可以灵活地描述事物及其之间关系的数据库模型。

现实世界中事物之间的联系更多的是非层次关系的，一个事物和另外的几个都有联系，用层次模型表示这种关系很不直观，网状模型克服了这一弊病，可以清晰地表示这种非层次关系。用有向图结构表示实体类型及实体间联系的数据结构模型称为网状模型。网状模型取消了层次模型的不能表示非树状结构的限制，两个或两个以上的结点都可以有多个双亲结点，则此有向树变成了有向图，该有向图描述了网状模型。

网状模型中以记录为数据的存储单位。记录包含若干数据项。网状数据库的数据项可以是多值的和复合的数据。每个记录有一个唯一标识它的内部标识符，称为码（DatabaseKey，DBK），它在一个记录存入数据库时由数据库管理系统（Database Management System，DBMS）自动赋予。DBK 可以看作记录的逻辑地址，可作记录的替身或用于寻找记录。网状数据库是导航式（Navigation）数据库，用户在操作数据库时不但说明要做什么，还要说明怎么做。例如在查找语句中不但要说明查找的对象，而且要规定存取路径。

### 3) 关系模型

关系模型是在关系结构的数据库中用二维表格的形式表示实体以及实体之间的联系的模型。关系模型是以集合论中的关系概念为基础发展起来的。关系模型中无论是实体还是实体间的联系均由单一的结构类型关系来表示。

关系模型的基本假定是所有数据都表示为数学上的关系，就是说  $n$  个集合的笛卡儿积的一个子集，有关这种数据的推理通过二值的谓词逻辑来进行，这意味着对每个命题都只有两种可

能的值：要么是真，要么是假。数据通过关系演算和关系代数的一种方式来操作。关系模型是采用二维表格结构表达实体类型及实体间联系的数据模型。

关系模型允许设计者通过数据库规范化的提炼，去建立一个信息的一致性的模型。访问计划和其他实现与操作细节由 DBMS 引擎来处理，而不应该反映在逻辑模型中。这与 SQL DBMS 普遍的实践是对立的，在它们那里性能调整经常需要改变逻辑模型。

基本的关系建造块是域或者叫数据类型。元组是属性的有序多重集（Multiset），属性是域和值的有序对。关系变量（Relvar）是域和名字的有序对（序偶）的集合，它充当关系的表头（Header）。关系是元组的集合。尽管这些关系概念是在数学上的定义的，它们可以宽松地映射到传统数据库概念上。表是关系公认的可视表示；元组类似于行的概念。

关系模型的基本原理是信息原理，即所有信息都表示为关系中的数据值。所以，关系变量在设计时是相互无关联的；反而，设计者在多个关系变量中使用相同的域，如果一个属性依赖于另一个属性，则通过参照完整性来强制这种依赖性。

### 3. 常用数据库类型

数据库根据存储方式可以分为关系型数据库（SQL）和非关系型数据库（Not Only SQL, NoSQL）。

#### 1) 关系型数据库

网状数据库（以网状数据模型为基础建立的数据库）和层次数据库（采用层次模型作为数据组织方式的数据库）已经很好地解决了数据的集中和共享问题，但是在数据独立性和抽象级别上仍有很大欠缺。用户在对这两种数据库进行存取时，仍然需要明确数据的存储结构，指出存取路径。而后来出现的关系数据库较好地解决了这些问题。关系数据库系统采用关系模型作为数据的组织方式。关系数据库是在一个给定的应用领域中，所有实体及实体之间联系的集合。关系型数据库支持事务的 ACID 原则，即原子性（Atomicity）、一致性（Consistency）、隔离性（Isolation）、持久性（Durability），这四种原则保证在事务过程当中数据的正确性。

#### 2) 非关系型数据库

非关系型数据库是分布式的、非关系型的、不保证遵循 ACID 原则的数据存储系统。NoSQL 数据存储不需要固定的表结构，通常也不存在连接操作。在大数据存取上具备关系型数据库无法比拟的性能优势。

常见的非关系数据库分为：

- 键值数据库：类似传统语言中使用的哈希表。可以通过key来添加、查询或者删除数据，因为使用key主键访问，会获得很高的性能及扩展性。Key/Value模型对于信息系统来说，其优势在于简单、易部署、高并发。
- 列存储（Column-oriented）数据库：将数据存储在列族中，一个列族存储经常被一起查询，比如人们经常会查询某个人的姓名和年龄，而不是薪资。这种情况下姓名和年龄会被放到一个列族中，薪资会被放到另一个列族中。这种数据库通常用来应对分布式存储海量数据。
- 面向文档（Document-Oriented）数据库：文档型数据库可以看作是键值数据库的升级

版，允许之间嵌套键值，而且文档型数据库比键值数据库的查询效率更高。面向文档数据库会将数据以文档形式存储。

- 图形数据库：允许人们将数据以图的方式存储。实体会作为顶点，而实体之间的关系则会作为边。比如有三个实体：Steve Jobs、Apple和Next，则会有两个Founded by的边将Apple和Next连接到Steve Jobs。

### 3) 不同存储方式数据库的优缺点

关系型数据库和非关系型数据库的优缺点，如表 2-3 所示。

表 2-3 常用存储数据库类型优缺点

数据库类型	特点类型	描述
关系型数据库	优点	<ul style="list-style-type: none"> <li>● 容易理解：二维表结构是非常贴近逻辑世界的一个概念，关系模型相对网状、层次等其他模型来说更容易理解</li> <li>● 使用方便：通用的SQL语言使得操作关系型数据库非常方便</li> <li>● 易于维护：丰富的完整性（实体完整性、参照完整性和用户定义的完整性）大大减低了数据冗余和数据不一致的概率</li> </ul>
	缺点	<ul style="list-style-type: none"> <li>● 数据读写必须经过SQL解析，大量数据、高并发下读写性能不足（对于传统关系型数据库来说，硬盘I/O是一个很大的瓶颈）</li> <li>● 具有固定的表结构，因此扩展困难</li> <li>● 多表的关联查询导致性能欠佳</li> </ul>
非关系型数据库	优点	<ul style="list-style-type: none"> <li>● 高并发：大数据下读写能力较强（基于键值对的，可以想象成表中的主键和值的对应关系，且不需要经过SQL层的解析，所以性能非常高）</li> <li>● 基本支持分布式：易于扩展，可伸缩（因为基于键值对，数据之间没有耦合性，所以非常容易水平扩展）</li> <li>● 简单：弱结构化存储</li> </ul>
	缺点	<ul style="list-style-type: none"> <li>● 事务支持较弱</li> <li>● 通用性差</li> <li>● 无完整约束，复杂业务场景支持较差</li> </ul>

## 4. 数据仓库

传统的数据库系统中缺乏决策分析所需的大量历史数据信息，因为传统的数据库一般只保留当前或近期的数据信息。为了满足中高层管理人员预测、决策分析的需要，在传统数据库的基础上产生了能够满足预测、决策分析需要的数据环境——数据仓库。数据仓库相关的基础概念包括：

- 清洗/转换/加载（Extract/Transformation/Load, ETL）：用户从数据源抽取出所需的数据，经过数据清洗、转换，最终按照预先定义好的数据仓库模型，将数据加载到数据仓库中去。
- 元数据：关于数据的数据，指在数据仓库建设过程中所产生的有关数据源定义、目标定义、转换规则等关键数据。同时元数据还包含关于数据含义的商业信息。典型的元数据包括：数据仓库表的结构、数据仓库表的属性、数据仓库的源数据（记录系

统）、从记录系统到数据仓库的映射、数据模型的规格说明、抽取日志和访问数据的公用例行程序等。

- 粒度：数据仓库的数据单位中保存数据的细化或综合程度的级别。细化程度越高，粒度级就越小；相反，细化程度越低，粒度级就越大。
- 分割：结构相同的数据被分成多个数据物理单元。任何给定的数据单元属于且仅属于一个分割。
- 数据集市：小型的，面向部门或工作组级的数据仓库。
- 操作数据存储（Operation Data Store, ODS）：能支持组织日常的全局应用的数据集合，是不同于DB的一种新的数据环境，是DW扩展后得到的一个混合形式。它具有四个基本特点：面向主题的、集成的、可变的、当前或接近当前的。
- 数据模型：逻辑数据结构，包括由数据库管理系统为有效进行数据库处理提供的操作和约束；用于表示数据的系统。
- 人工关系：在决策支持系统环境中用于表示参照完整性的一种设计技术。

数据仓库是一个面向主题的、集成的、非易失的且随时间变化的数据集合，用于支持管理决策。常见的数据仓库的体系结构如图 2-2 所示。

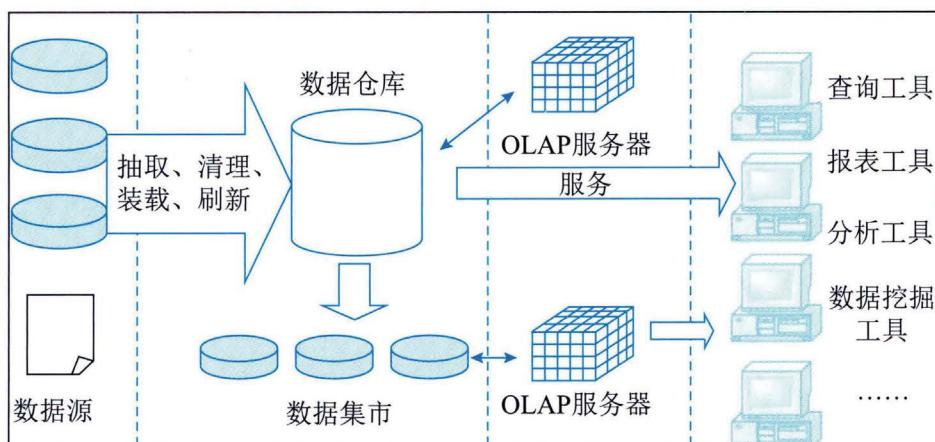


图 2-2 数据仓库体系结构

(1) 数据源。它是数据仓库系统的基础，是整个系统的数据源泉。通常包括组织内部信息和外部信息。内部信息包括存放于关系型数据库管理系统中的各种业务处理数据和各类文档数据。外部信息包括各类法律法规、市场信息和竞争对手的信息等。

(2) 数据的存储与管理。它是整个数据仓库系统的核心。数据仓库的真正关键是数据的存储和管理。数据仓库的组织管理方式决定了它有别于传统数据库，同时也决定了其对外部数据的表现形式。要决定采用什么产品和技术来建立数据仓库的核心，则需要从数据仓库的技术特点着手分析。针对现有各业务系统的数据，进行抽取、清理并有效集成，按照主题进行组织。数据仓库按照数据的覆盖范围可以分为组织级数据仓库和部门级数据仓库（通常称为数据集市）。

(3) 联机分析处理（On-Line Analytic Processing, OLAP）服务器。OLAP 对分析需要的数据进行有效集成，按多维模型予以组织，以便进行多角度、多层次的分析，并发现趋势。其具

体实现可以分为：基于关系数据库的 OLAP（Relational OLAP, ROLAP）、基于多维数据组织的 OLAP（Multidimensional OLAP, MOLAP）和基于混合数据组织的 OLAP（Hybrid OLAP, HOLAP）。ROLAP 基本数据和聚合数据均存放在 RDBMS 之中；MOLAP 基本数据和聚合数据均存放于多维数据库中；HOLAP 基本数据存放于关系数据库管理系统（Relational Database Management System, RDBMS）之中，聚合数据存放于多维数据库中。

（4）前端工具。前端工具主要包括各种查询工具、报表工具、分析工具、数据挖掘工具以及各种基于数据仓库或数据集市的应用开发工具。其中数据分析工具主要针对 OLAP 服务器，报表工具、数据挖掘工具主要针对数据仓库。

## 2.1.4 信息安全

常见的信息安全问题主要表现为：计算机病毒泛滥、恶意软件的入侵、黑客攻击、利用计算机犯罪、网络有害信息泛滥、个人隐私泄露等。随着物联网、云计算、人工智能、大数据等新一代信息技术的广泛应用，信息安全也面临着新的问题和挑战。

### 1. 信息安全基础

信息安全强调信息（数据）本身的安全属性，主要包括以下内容。

- 保密性（Confidentiality）：信息不被未授权者知晓的属性。
- 完整性（Integrity）：信息是正确的、真实的、未被篡改的、完整无缺的属性。
- 可用性（Availability）：信息可以随时正常使用的属性。

信息必须依赖其存储、传输、处理及应用的载体（媒介）而存在，因此针对信息系统，安全可以划分为四个层次：设备安全、数据安全、内容安全、行为安全。

信息系统一般由计算机系统、网络系统、操作系统、数据库系统和应用系统组成。与此对应，信息系统安全主要包括计算机设备安全、网络安全、操作系统安全、数据库系统安全和应用系统安全等。

网络安全技术主要包括：防火墙、入侵检测与防护、VPN、安全扫描、网络蜜罐技术、用户和实体行为分析技术等。

### 2. 加密解密

为了保证信息的安全性，就需要采用信息加密技术对信息进行伪装，使得信息非法窃取者无法理解信息的真实含义；需要采用加密算法提取信息的特征码（校验码）或特征矢量，并与有关信息封装在一起，信息的合法拥有者可以利用特征码对信息的完整性进行校验；需要采用加密算法对信息使用者的身份进行认证、识别和确认，以对信息的使用进行控制。

发信者将明文数据加密成密文，然后将密文数据送入网络传输或存入计算机文件，而且只给合法收信者分配密钥。合法收信者接收到密文后，实行与加密变换相逆的变换，去掉密文的伪装并恢复出明文，这一过程称为解密（Decryption）。解密在解密密钥的控制下进行。用于解密的一组数学变换称为解密算法。

加密技术包括两个元素：算法和密钥。密钥加密技术的密码体制分为对称密钥体制和非对

称密钥体制两种。相应地，对数据加密的技术分为两类，即对称加密（私人密钥加密）和非对称加密（公开密钥加密）。对称加密以数据加密标准（Data Encryption Standard, DES）算法为典型代表，非对称加密通常以RSA（Rivest Shamir Adleman）算法为代表。对称加密的加密密钥和解密密钥相同，而非对称加密的加密密钥和解密密钥不同，加密密钥可以公开而解密密钥需要保密。

### 3. 安全行为分析技术

传统安全产品、技术、方案基本上都是基于已知特征进行规则匹配来进行分析和检测。基于特征、规则和人工分析，以“特征”为核心的检测分析存在安全可见性盲区，有滞后效应、无力检测未知攻击、容易被绕过，以及难以适应攻防对抗的网络现实和快速变化的组织环境、外部威胁等问题。另一方面，虽然大多数的攻击可能来自组织以外，但最严重的损害往往是由内部人员造成的，只有管理好内部威胁，才能保证信息和网络安全。

用户和实体行为分析（User and Entity Behavior Analytics, UEBA）提供了用户画像及基于各种分析方法的异常检测，结合基本分析方法（利用签名的规则、模式匹配、简单统计、阈值等）和高级分析方法（监督和无监督的机器学习等），用打包分析来评估用户和其他实体（主机、应用程序、网络、数据库等），发现与用户或实体标准画像或行为异常的活动所相关的潜在事件。UEBA以用户和实体为对象，利用大数据，结合规则以及机器学习模型，并通过定义此类基线，对用户和实体行为进行分析和异常检测，尽可能快速地感知内部用户和实体的可疑或非法行为。

UEBA是一个完整的系统，涉及算法、工程等检测部分，以及用户与实体风险评分排序、调查等用户交换和反馈。从架构上来看，UEBA系统通常包括数据获取层、算法分析层和场景应用层。

### 4. 网络安全态势感知

网络安全态势感知（Network Security Situation Awareness）是在大规模网络环境中，对能够引起网络态势发生变化的安全要素进行获取、理解、显示，并据此预测未来的网络安全发展趋势。安全态势感知不仅是一种安全技术，也是一种新兴的安全概念。它是一种基于环境的、动态的、整体的洞悉安全风险的能力。安全态势感知的前提是安全大数据，其在安全大数据的基础上进行数据整合、特征提取等，然后应用一系列态势评估算法生成网络的整体态势状况，应用态势预测算法预测态势的发展状况，并使用数据可视化技术，将态势状况和预测情况展示给安全人员，方便安全人员直观便捷地了解网络当前状态及预期的风险。

网络安全态势感知的关键技术主要包括：海量多元异构数据的汇聚融合技术、面向多类型的网络安全威胁评估技术、网络安全态势评估与决策支撑技术、网络安全态势可视化等。

## 2.1.5 信息技术的发展

作为信息技术的基础，计算机软硬件、网络、存储和数据库、信息安全等都在不断的发展创新，引领着当前信息技术发展的潮流。

在计算机软硬件方面，计算机硬件技术将向超高速、超小型、平行处理、智能化的方向发展，计算机硬件设备的体积越来越小、速度越来越高、容量越来越大、功耗越来越低、可靠性越来越高。计算机软件越来越丰富，功能越来越强大，“软件定义一切”概念成为当前发展的主流。

在网络技术方面，计算机网络与通信技术之间的联系日益密切，甚至是已经融为一体。作为国家最重要的基础设施之一，5G成为当前的主流，面向物联网、低时延场景的窄带物联网（Narrow Band Internet of Things, NB-IoT）和增强型机器类型通信（enhanced Machine-Type Communication, eMTC）、工业物联网（Industrial Internet of Things, IIoT）和低延时高可靠通信（Ultra Reliable Low Latency Communication, URLLC）等技术，将进一步得到充分发展。

在存储和数据库方面，随着数据量的不断爆炸式增长，数据存储结构也越来越灵活多样，日益变革的新业务需求驱使数据库及应用系统的存在形式愈发丰富，这些变化均对各类数据库的架构和存储模式提出了挑战，推动数据库技术不断向着模型拓展、架构解耦的方向演进。

在信息安全方面，传统计算机安全理念将过渡到以可信计算理念为核心的计算机安全，由网络普及应用引发的技术与应用模式的变革，正在进一步推动信息安全网络化关键技术的创新；同时信息安全标准的研究与制定，信息安全产品和服务的集成和融合，正引领着当前信息安全技术朝着标准化和集成化的方向发展。

## 2.2 新一代信息技术及应用

信息技术在智能化、系统化、微型化、云端化的基础上不断融合创新，促进了物联网、云计算、大数据、区块链、人工智能、虚拟现实等新一代信息技术的诞生。新一代信息技术与信息资源充分开发利用形成的新模式、新业态等，是信息化发展的主要趋势，也是信息系统集成领域未来的重要业务范畴。

### 2.2.1 物联网

物联网（The Internet of Things）是指通过信息传感设备，按约定的协议将任何物品与互联网相连接，进行信息交换和通信，以实现智能化识别、定位、跟踪、监控和管理的网络。物联网主要解决物品与物品（Thing to Thing, T2T）、人与物品（Human to Thing, H2T）、人与人（Human to Human, H2H）之间的互连。另外，许多学者在讨论物联网时经常会引入 M2M 的概念：可以解释为人与人（Man to Man）、人与机器（Man to Machine）或机器与机器（Machine to Machine）。

#### 1. 技术基础

物联网架构可分为三层：感知层、网络层和应用层。感知层由各种传感器构成，包括温度传感器，二维码标签、RFID 标签和读写器，摄像头，GPS 等感知终端。感知层是物联网识别物体、采集信息的来源。网络层由各种网络，包括互联网、广电网、网络管理系统和云计算平台等组成，是整个物联网的中枢，负责传递和处理感知层获取的信息。应用层是物联网和用户的接口，它与行业需求结合以实现物联网的智能应用。

物联网的产业链包括传感器和芯片、设备、网络运营及服务、软件与应用开发和系统集成。物联网技术在智能电网、智慧物流、智能家居、智能交通、智慧农业、环境保护、医疗健康、城市管理（智慧城市）、金融服务与保险业、公共安全等方面有非常关键和重要的应用。

## 2. 关键技术

物联网关键技术主要涉及传感器技术、传感网和应用系统框架等。

### 1) 传感器技术

传感器是一种检测装置，它能“感受”到被测量的信息，并能将检测到的信息按一定规律变换成为电信号或其他所需形式的信息输出，以满足信息的传输、处理、存储、显示、记录和控制等要求。它是实现自动检测和自动控制的首要环节，也是物联网获取物理世界信息的基本手段。

射频识别技术（Radio Frequency Identification, RFID）是物联网中使用的一种传感器技术，在物联网发展中备受关注。RFID 可通过无线电信号识别特定目标并读写相关数据，而无须识别系统与特定目标之间建立机械或光学接触。RFID 是一种简单的无线系统，由一个询问器（或阅读器）和很多应答器（或标签）组成。标签由耦合元件及芯片组成，每个标签具有扩展词条唯一的电子编码，附着在物体上标识目标对象，它通过天线将射频信息传递给阅读器，阅读器就是读取信息的设备。RFID 技术让物品能够“开口说话”。这就赋予了物联网一个特性——可跟踪性，即可以随时掌握物品的准确位置及其周边环境。

### 2) 传感网

微机电系统（Micro-Electro-Mechanical Systems, MEMS）是由微传感器、微执行器、信号处理和控制电路、通信接口和电源等部件组成的一体化的微型器件系统。其目标是把信息的获取、处理和执行集成在一起，组成具有多功能的微型系统，集成于大尺寸系统中，从而大幅地提高系统的自动化、智能化和可靠性水平。MEMS 赋予了普通物体新的“生命”，它们有了属于自己的数据传输通路、存储功能、操作系统和专门的应用程序，从而形成一个庞大的传感网，使物联网能够通过物品来实现对人的监控与保护。未来，衣服可以通过传感网“告诉”洗衣机放多少水和洗衣粉最经济；文件夹会“检查”人们忘带了什么重要文件；食品蔬菜的标签会向顾客的手机介绍“自己”是否真正“绿色安全”。

### 3) 应用系统框架

物联网应用系统框架是一种以机器终端智能交互为核心的、网络化的应用与服务。它将使对象实现智能化的控制，涉及 5 个重要的技术部分：机器、传感器硬件、通信网络、中间件和应用。该框架基于云计算平台和智能网络，可以依据传感器网络获取的数据进行决策，改变对象的行为控制和反馈。以智能停车场为例，当车辆驶入或离开天线通信区时，天线以微波通信的方式与电子识别卡进行双向数据交换，从电子车卡上读取车辆的相关信息，从司机卡上读取司机的相关信息，自动识别电子车卡和司机卡，并判断该车卡是否有效和司机卡的合法性，核对车道控制电脑并显示与该电子车卡和司机卡一一对应的车牌号码及驾驶员等资料信息。车道控制电脑自动将通过时间、车辆和驾驶员的有关信息存入数据库中，车道控制电脑根据读到的

数据判断是正常卡、未授权卡、无卡还是非法卡，据此做出相应的回应和提示。另外，家中的老人通过佩戴嵌入智能传感器的手表，在外地的子女可以随时通过手机查询父母的血压、心跳是否稳定。智能化的住宅在主人上班时，传感器自动关闭水电气和门窗，定时向主人的手机发送消息，汇报安全情况。

### 3. 应用和发展

物联网的应用领域涉及人们工作与生活的方方面面。在工业、农业、环境、交通、物流、安保等基础设施领域的应用，有效地推动了这些方面的智能化发展，使得有限的资源能更加合理地使用分配，从而提高了行业效率、效益；在家居、医疗健康、教育、金融与服务业、旅游业等与生活息息相关领域的应用，通过与社会科学和社会治理的充分融合创新，实现了服务范围、服务方式和服务质量等方面的巨大变革和进步。

## 2.2.2 云计算

云计算（Cloud Computing）是分布式计算的一种，指的是通过网络“云”将巨大的数据计算处理程序分解成无数个小程序，然后通过由多部服务器组成的系统进行处理和分析这些小程序得到结果并返回给用户。在云计算早期，就是简单的分布式计算，进行任务分发并对计算结果进行合并。当前的云计算已经不单单是一种分布式计算，而是分布式计算、效用计算、负载均衡、并行计算、网络存储、热备份冗余和虚拟化等计算机技术混合演进并跃升的结果。

### 1. 技术基础

云计算是一种基于互联网的计算方式，通过这种方式将网络上配置为共享的软件资源、计算资源、存储资源和信息资源，按需求提供给网上的终端设备和终端用户。云计算也可以理解为向用户屏蔽底层差异的分布式处理架构。在云计算环境中，用户与实际服务提供的计算资源相分离，云端集合了大量计算设备和资源。

当使用云计算服务时，用户不需要安排专门的维护人员，云计算服务的提供商会为数据和服务器的安全做出相对较高水平的保护。由于云计算将数据存储在云端（分布式的云计算设备中承担计算和存储功能的部分），业务逻辑和相关计算都在云端完成，因此，终端只需要一个能够满足基础应用的普通设备即可。

云计算实现了“快速、按需、弹性”的服务，用户可以随时通过宽带网络接入“云”并获得服务，按照实际需求获取或释放资源，根据需求对资源进行动态扩展。

按照云计算服务提供的资源层次，可以分为基础设施即服务（Infrastructure as a Service, IaaS）、平台即服务（Platform as a Service, PaaS）和软件即服务（Software as a Service, SaaS）三种服务类型。

IaaS 向用户提供计算能力、存储空间等基础设施方面的服务。这种服务模式需要较大的基础设施投入和长期运营管理经验，其单纯出租资源的盈利能力有限。

PaaS 向用户提供虚拟的操作系统、数据库管理系统、Web 应用等平台化的服务。PaaS 服务的重点不在于直接的经济效益，而更注重构建和形成紧密的产业生态。

SaaS 向用户提供应用软件（如 CRM、办公软件等）、组件、工作流等虚拟化软件的服务，

SaaS一般采用Web技术和SOA架构，通过Internet向用户提供多租户、可定制的应用能力，大大缩短了软件产业的渠道链条，减少了软件升级、定制和运行维护的复杂程度，并使软件提供商从软件产品的生产者转变为应用服务的运营者。

## 2. 关键技术

云计算的关键技术主要涉及虚拟化技术、云存储技术、多租户和访问控制管理、云安全技术等。

### 1) 虚拟化技术

虚拟化是一个广义术语，在计算机领域通常是指计算元件在虚拟的基础上而不是真实的基础上运行。虚拟化技术可以扩大硬件的容量，简化软件的重新配置过程。CPU的虚拟化技术可以单CPU模拟多CPU并行，允许一个平台同时运行多个操作系统，并且应用程序都可以在相互独立的空间内运行而互不影响，从而显著提高计算机的工作效率。

虚拟化技术与多任务以及超线程技术是完全不同的。多任务是指在一个操作系统中多个程序同时并行运行，而在虚拟化技术中，则可以同时运行多个操作系统，而且每一个操作系统中都有多个程序运行，每一个操作系统都运行在一个虚拟的CPU或者虚拟主机上。超线程技术只是单CPU模拟双CPU来平衡程序运行性能，这两个模拟出来的CPU是不能分离的，只能协同工作。

容器（Container）技术是一种全新意义上的虚拟化技术，属于操作系统虚拟化的范畴，也就是由操作系统提供虚拟化的支持。目前最受欢迎的容器环境是Docker。容器技术将单个操作系统的资源划分到孤立的组中，以便更好地在孤立的组之间平衡有冲突的资源使用需求。例如：用户创建一个应用，传统方式需要虚拟机，但虚拟机本身就占用了更多的系统资源。又如，应用需要在开发和运维之间转移、协作，当开发和运维的操作环境不同时，也会影响结果。使用容器技术可将应用隔离在一个独立的运行环境中，该独立环境称之为容器，可以减少运行程序带来的额外消耗，并可以在几乎任何地方以相同的方式运行。

### 2) 云存储技术

云存储技术是基于传统媒体系统发展而来的一种全新信息存储管理方式，该方式整合应用了计算机系统的软硬件优势，可较为快速、高效地对海量数据进行在线处理，通过多种云技术平台的应用，实现了数据的深度挖掘和安全管理。

分布式文件系统作为云存储技术中的重要组成部分，在维持兼容性的基础上，对系统复制和容错功能进行提升。同时，通过云集群管理实现云存储的可拓展性，借助模块之间的合理搭配，完成解决方案拟定解决的网络存储问题、联合存储问题、多节点存储问题、备份处理、负载均衡等。云储存的实现过程中，结合分布式的文件结构，在硬件支撑的基础上，对硬件运行环境进行优化，确保数据传输的完整性和容错性；结合成本低廉的硬件的扩展，大大降低了存储的成本。

在分布式文件系统的支撑下，实现了通过云存储资源的拓展，辅助高吞吐量数据的分析，使得用户可以更加充分、全面地进行数据管理，实现用户上传信息的优化管理，满足了不同平台信息获取需要。另一方面，通过加强对云存储技术中相关数据的安全防护，实现信息存储过

程中的病毒防护和安全监控，确保信息存储应用的安全性。

### 3) 多租户和访问控制管理

云计算环境下访问控制的研究是伴随着云计算的发展而开始的，访问控制管理是云计算应用的核心问题之一。云计算访问控制的研究主要集中在云计算访问控制模型、基于 ABE 密码体制的云计算访问控制、云中多租户及虚拟化访问控制研究。

云计算访问控制模型就是按照特定的访问策略来描述安全系统，建立安全模型的一种方法。用户（租户）可以通过访问控制模型得到一定的权限，进而对云中的数据进行访问，所以访问控制模型多用于静态分配用户的权限。云计算中的访问控制模型都是以传统的访问控制模型为基础，在传统的访问控制模型上进行改进，使其更适用于云计算的环境。根据访问控制模型功能的不同，研究的内容和方法也不同，常见的有基于任务的访问控制模型、基于属性模型的云计算访问控制、基于 UCON 模型的云计算访问控制、基于 BLP 模型的云计算访问控制等。

基于 ABE 密码机制的云计算访问控制包括 4 个参与方：数据提供者、可信第三方授权中心、云存储服务器和用户。首先，可信授权中心生成主密钥和公开参数，将系统公钥传给数据提供者，数据提供者收到系统公钥之后，用策略树和系统公钥对文件加密，将密文和策略树上传到云服务器；然后，当一个新用户加入系统后，将自己的属性集上传给可信授权中心并提交私钥申请请求，可信授权中心针对用户提交的属性集和主密钥计算生成私钥，传给用户；最后，用户下载感兴趣的数据。如果其属性集合满足密文数据的策略树结构，则可以解密密文；否则，访问数据失败。

云中多租户及虚拟化访问控制是云计算的典型特征。由于租户间共享物理资源，并且其可信度不容易得到，所以租户之间就可以通过侧通道攻击来从底层的物理资源中获得有用的信息。此外，由于在虚拟机上要部署访问控制策略可能会带来多个租户访问资源的冲突，导致物理主机上出现没有认证的或者权限分配错误的信息流。这就要求在云环境下，租户之间的通信应该由访问控制来保证，并且每个租户都有自己的访问控制策略，使得整个云平台的访问控制变得复杂。目前，对多租户访问控制的研究主要集中在对多租户的隔离和虚拟机的访问控制方面。

### 4) 云安全技术

云安全研究主要包含两个方面的内容，一是云计算技术本身的安全保护工作，涉及相应的数据完整性及可用性、隐私保护性以及服务可用性等方面的内容；二是借助于云服务的方式来保障客户端用户的安全防护需求，通过云计算技术来实现互联网安全，涉及基于云计算的病毒防治、木马检测技术等。

在云安全技术的研究方面，主要包含：

- 云计算安全性：主要是对于云自身以及所涉及的应用服务内容进行分析，重点探讨其相应得安全性问题，这里主要涉及如何有效实现安全隔离，保障互联网用户数据的安全性，如何有效防护恶意网络攻击，提升云计算平台的系统安全性，以及用户接入认证以及相应的信息传输审计、安全等方面的工作。
- 保障云基础设施的安全性：主要就是如何利用相应的互联网安全基础设备的相应资源，有效实现云服务的优化，从而保障满足预期的安全防护的要求。

- 云安全技术服务：重点集中于如何保障实现互联网终端用户的安全服务要求，能有效实现客户端的计算机病毒防治等相关服务工作。从云安全架构的发展情况来看，如果云计算服务商的安全等级不高，会造成服务用户需要具备更强的安全能力、承担更多管理职责。

为了提升云安全体系的能力，保障其具有较强的可靠性，云安全技术要从开放性、安全保障、体系结构的角度考虑。①云安全系统具有一定的开放性，要保障开放环境下可信认证；②在云安全系统方面，要积极采用先进的网络技术和病毒防护技术；③在云安全体系构建过程中，要保证其稳定性，以满足海量数据动态变化的需求。

综上所述，云安全技术是新一代互联网中安全技术构架的核心内容，体现了当前快速发展的云计算的先进性，是未来的信息安全技术发展的必然趋势。随着云计算应用领域的拓展，云安全技术也必然会越来越成熟，能有效全方位保障广大互联网用户的 data 应用安全性，对于云计算的进一步推广与应用具有至关重要的作用。

### 3. 应用和发展

云计算经历十余年的发展，已逐步进入成熟期，在众多领域正发挥着越来越大的作用，“上云”将成为各类组织加快数字化转型、鼓励技术创新和促进业务增长的第一选择，甚至是必备的前提条件。

云计算将进一步成为创新技术和最佳工程实践的重要载体和试验场。从 AI 与机器学习、IoT 与边缘计算、区块链到工程实践领域的 DevOps、云原生和 Service Mesh，都有云计算厂商积极参与、投入和推广的身影。以人工智能为例，从前面提到的 IaaS 中 GPU 计算资源的提供，到面向特定领域成熟模型能力开放（如各类自然语言处理、图像识别、语言合成的 API），再到帮助打造定制化 AI 模型的机器学习平台，云计算实际上已成为 AI 相关技术的基础。

云计算将顺应产业互联网大潮，下沉行业场景，向垂直化、产业化纵深发展。随着通用类架构与功能的不断完善和对行业客户的不断深耕，云计算自然渗透进入更多垂直领域，成为提供更贴近行业业务与典型场景的基础能力。以金融云为例，云计算可针对金融保险机构特殊的合规和安全需要，提供物理隔离的基础设施，还可提供支付、结算、风控、审计等业务组件。

多云和混合云将成为大中型组织的刚需，得到更多重视与发展。当组织大量的工作负载部署在云端，新的问题则会显现：①虽然云端已经能提供相当高的可用性，但为了避免单一供应商出现故障时的风险，关键应用仍须架设必要的技术冗余；②当业务规模较大时，从商业策略角度看，也需要避免过于紧密的厂商绑定，以寻求某种层面的商业制衡和主动权。

云的生态建设重要性不断凸显，成为影响云间竞争的关键因素。当某个云发展到一定规模和阶段，就不能仅仅考虑技术和产品，需要站在长远发展的角度，建立和培养具有生命力的繁荣生态和社区。另外，云生态需要关注面向广大开发者、架构师和运维工程师的持续输出、培养和影响。只有赢得广大技术人员的关注和喜爱，才能赢得未来的云计算市场。

综上所述，“创新、垂直、混合、生态”这四大趋势伴随云计算快速发展。云计算对 IT 硬件资源与软件组件进行了标准化、抽象化和规模化，某种意义上颠覆和重构了 IT 业界的供应链，是当前新一代信息技术发展的巨大的革新与进步。

### 2.2.3 大数据

大数据（Big Data）指无法在一定时间范围内用常规软件工具进行捕捉、管理和处理的数据集合，是具有更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产。

#### 1. 技术基础

大数据是具有体量大、结构多样、时效性强等特征的数据，处理大数据需要采用新型计算架构和智能算法等新技术。大数据从数据源到最终价值实现一般需要经过数据准备、数据存储与管理、数据分析和计算、数据治理和知识展现等过程，涉及数据模型、处理模型、计算理论以及与其相关的分布计算、分布存储平台技术、数据清洗和挖掘技术、流式计算和增量处理技术、数据质量控制等方面的研究。一般来说，大数据主要特征包括：

- 数据海量：大数据的数据体量巨大，从TB级别跃升到PB级别（ $1\text{PB}=1024\text{TB}$ ）、EB级别（ $1\text{EB}=1024\text{PB}$ ），甚至达到ZB级别（ $1\text{ZB}=1024\text{EB}$ ）。
- 数据类型多样：大数据的数据类型繁多，一般分为结构化数据和非结构化数据。相对于以往便于存储的以文本为主的结构化数据，非结构化数据越来越多，包括网络日志、音频、视频、图片、地理位置信息等，这些多类型的数据对数据的处理能力提出了更高要求。
- 数据价值密度低：数据价值密度的高低与数据总量的大小成反比。以视频为例，一部1小时的视频，在连续不间断的监控中，有用数据可能仅有一二秒。如何通过强大的机器算法更迅速地完成数据的价值“提纯”，成为目前大数据背景下亟待解决的难题。
- 数据处理速度快：为了从海量的数据中快速挖掘数据价值，一般要求要对不同类型的数据进行快速的处理，这是大数据区别于传统数据挖掘的最显著特征。

#### 2. 关键技术

大数据技术作为信息化时代的一项新兴技术，技术体系处在快速发展阶段，涉及数据的处理、管理、应用等多个方面。具体来说，技术架构是从技术视角研究和分析大数据的获取、管理、分布式处理和应用等。大数据的技术架构与具体实现的技术平台和框架息息相关，不同的技术平台决定了不同的技术架构和实现。从总体上说，大数据技术架构主要包含大数据获取技术、分布式数据处理技术和大数据管理技术，以及大数据应用和服务技术。

##### 1) 大数据获取技术

目前，大数据获取的研究主要集中在数据采集、整合和清洗三个方面。数据采集技术实现数据源的获取，然后通过整合和清理技术保证数据质量。

数据采集技术主要是通过分布式爬取、分布式高速高可靠性数据采集、高速全网数据映像技术，从网站上获取数据信息。除了网络中包含的内容之外，对于网络流量的采集可以使用DPI或DFI等带宽管理技术进行处理。

数据整合技术是在数据采集和实体识别的基础上，实现数据到信息的高质量整合。数据整合技术包括多源多模态信息集成模型、异构数据智能转换模型、异构数据集成的智能模式抽取

和模式匹配算法、自动容错映射和转换模型及算法、整合信息的正确性验证方法、整合信息的可用性评估方法等。

数据清洗技术一般根据正确性条件和数据约束规则，清除不合理和错误的数据，对重要的信息进行修复，保证数据的完整性。包括数据正确性语义模型、关联模型和数据约束规则、数据错误模型和错误识别学习框架、针对不同错误类型的自动检测和修复算法、错误检测与修复结果的评估模型和评估方法等。

## 2) 分布式数据处理技术

分布式计算是随着分布式系统的发展而兴起的，其核心是将任务分解成许多小的部分，分配给多台计算机进行处理，通过并行工作的机制，达到节约整体计算时间，提高计算效率的目的。目前，主流的分布式计算系统有 Hadoop、Spark 和 Storm。Hadoop 常用于离线的复杂的大数据处理，Spark 常用于离线的快速的大数据处理，而 Storm 常用于在线的实时的大数据处理。

大数据分析与挖掘技术主要指改进已有数据挖掘和机器学习技术；开发数据网络挖掘、特异群组挖掘、图挖掘等新型数据挖掘技术；创新基于对象的数据连接、相似性连接等大数据融合技术；突破用户兴趣分析、网络行为分析、情感语义分析等面向领域的大数据挖掘技术。

## 3) 大数据管理技术

大数据管理技术主要集中在大数据存储、大数据协同和安全隐私等方面。

大数据存储技术主要有三个方面。①采用 MPP 架构的新型数据库集群，通过列存储、粗粒度索引等多项大数据处理技术和高效的分布式计算模式，实现大数据存储；②围绕 Hadoop 衍生出相关的大数据技术，应对传统关系型数据库较难处理的数据和场景，通过扩展和封装 Hadoop 来实现对大数据存储、分析的支撑；③基于集成的服务器、存储设备、操作系统、数据库管理系统，实现具有良好的稳定性、扩展性的大数据一体机。

多数据中心的协同管理技术是大数据研究的另一个重要方向。通过分布式工作流引擎实现工作流调度、负载均衡，整合多个数据中心的存储和计算资源，从而为构建大数据服务平台提供支撑。

大数据隐私性技术的研究，主要集中于新型数据发布技术，尝试在尽可能少损失数据信息的同时最大化地隐藏用户隐私。在数据信息量和隐私之间是有矛盾的，目前没有非常好的解决办法。

## 4) 大数据应用和服务技术

大数据应用和服务技术主要包含分析应用技术和可视化技术。

大数据分析应用主要是面向业务的分析应用。在分布式海量数据分析和挖掘的基础上，大数据分析应用技术以业务需求为驱动，面向不同类型的业务需求开展专题数据分析，为用户提供高可用、高易用的数据分析服务。

可视化通过交互式视觉表现的方式来帮助人们探索和理解复杂的数据。大数据的可视化技术主要集中在文本可视化技术、网络（图）可视化技术、时空数据可视化技术、多维数据可视化和交互可视化等。在技术方面，主要关注原位交互分析（In Situ Interactive Analysis）、数据表示、不确定性量化和面向领域的可视化工具库。

### 3. 应用和发展

大数据像水、矿石、石油一样，正在成为新的资源和社会生产要素，从数据资源中挖掘潜在的价值，成为当前大数据时代研究的热点。如何快速对数量巨大、来源分散、格式多样的数据进行采集、存储和关联分析，从中发现新知识、创造新价值、提升新能力的新一代信息技术和服务业态，是其应用价值的重要体现。

(1) 在互联网行业，网络的广泛应用和社交网络已深入到社会工作、生活的方方面面，海量数据的产生、应用和服务一体化。每个人都是数据的生产者、使用者和受益者。从大量的数据中挖掘用户行为，反向传输到业务领域，支持更准确的社会营销和广告，可增加业务收入，促进业务发展。同时，随着数据的大量生成、分析和应用，数据本身已成为可以交易的资产，大数据交易和数据资产化成为当前具有价值的领域和方向。

(2) 在政府的公共数据领域，结合大数据的采集、治理和集成，将各个部门搜集的信息进行剖析和共享，能够发现管理上的纰漏，提高执法水平，增进财税增收和加大市场监管程度，大大改变政府管理模式、节省政府投资、增强市场管理，提高社会治理水平、城市管理能力和人民群众的服务能力。

(3) 在金融领域，大数据征信是重要的应用领域。通过大数据的分析和画像，能够实现个人信用和金融服务的结合，从而服务于金融领域的信任管理、风控管理、借贷服务等，为金融业务提供有效支撑。

(4) 在工业领域，结合海量的数据分析，能够为工业生产过程提供准确的指导，如在航运大数据领域，能够使用大数据对将来航路的国际贸易货量进行预测分析，预知各个口岸的热度；能够利用天气数据对航路的影响进行分析，提供相关业务的预警、航线的调整和资源的优化调配方案，避免不必要的亏损。

(5) 在社会民生领域，大数据的分析应用能够更好地为民生服务。以疾病预测为例，基于大数据的积累和智能分析，能够透视人们搜索“流感、肝炎、肺结核和未病”的发病时间和地点分布，结合气温变化、环境指数、人口流动等因素建立预测模型，能够为公共卫生治理人员提供多种传染病的趋势预测，帮助其提早进行预防部署。

## 2.2.4 区块链

“区块链”概念于 2008 年在《比特币：一种点对点电子现金系统》中被首次提出，并在比特币系统的数据加密货币体系中成功应用，已成为政府、组织和学者等重点关注和研究的热点。区块链技术具有多中心化存储、隐私保护、防篡改等特点，提供了开放、分散和容错的事务机制，成为新一代匿名在线支付、汇款和数字资产交易的核心，被广泛应用于各大交易平台，为金融、监管机构、科技创新、农业以及政治等领域带来了深刻的变革。

### 1. 技术基础

区块链概念可以理解为以非对称加密算法为基础，以改进的默克尔树（Merkle Tree）为数据结构，使用共识机制、点对点网络、智能合约等技术结合而成的一种分布式存储数据库技术。区块链分为公有链（Public Blockchain）、联盟链（Consortium Blockchain）、私有链（Private

Blockchain) 和混合链 (Hybrid Blockchain) 四大类。

一般来说，区块链的典型特征包括：

- 多中心化：链上数据的验证、核算、存储、维护和传输等过程均依赖分布式系统结构，运用纯数学方法代替中心化组织机构在多个分布式节点之间构建信任关系，从而建立可信的分布式系统。
- 多方维护：激励机制可确保分布式系统中的所有节点均可参与数据区块的验证过程，并通过共识机制选择特定节点将新产生的区块加入到区块链中。
- 时序数据：区块链运用带有时间戳信息的链式结构来存储数据信息，为数据信息添加时间维度的属性，从而可实现数据信息的可追溯性。
- 智能合约：区块链技术能够为用户提供灵活可变的脚本代码，以支持其创建新型的智能合约。
- 不可篡改：在区块链系统中，因为相邻区块间后序区块可对前序区块进行验证，若篡改某一区块的数据信息，则需递归修改该区块及其所有后序区块的数据信息，然而每一次哈希的重新计算代价是巨大的，且须在有限时间内完成，因此可保障链上数据的不可篡改性。
- 开放共识：在区块链网络中，每台物理设备均可作为该网络中的一个节点，任意节点可自由加入且拥有一份完整的数据库拷贝。
- 安全可信：数据安全可通过基于非对称加密技术对链上数据进行加密来实现，分布式系统中各节点通过区块链共识算法所形成的算力来抵御外部攻击、保证链上数据不被篡改和伪造，从而具有较高的保密性、可靠性和安全性。

## 2. 关键技术

从区块链的技术体系视角看，区块链基于底层的数据基础处理、管理和存储技术，以区块数据的管理、链式结构的数据、数字签名、哈希函数、默克尔树、非对称加密等，通过基于P2P网络的对称式网络，组织节点参与数据的传播和验证，每个节点均会承担网络路由、验证区块数据、传播区块数据、记录交易数据、发现新节点等功能，包含传播机制和验证机制。为保障区块链应用层的安全，通过激励层的发行机制和分配机制，在整个分布式网络的节点以最高效率的方式达成共识。

### 1) 分布式账本

分布式账本是区块链技术的核心之一。分布式账本的核心思想是：交易记账由分布在不同地方的多个节点共同完成，而且每一个节点保存一个唯一、真实账本的副本，它们可以参与监督交易合法性，同时也可共同为其作证；账本里的任何改动都会在所有的副本中被反映出来，反应时间会在几分钟甚至是几秒内，记账节点足够多，理论上除非所有的节点被破坏，所有整个分布式账本系统是非常稳健的，从而保证了账目数据的安全性。

分布式账本中存储的资产是指法律认可的合法资产，如金融、实体、电子的资产等任何形式的有价资产。为了确保资产的安全性和准确性，分布式账本一方面通过公私钥以及签名控制账本的访问权；另一方面根据共识的规则，账本中的信息更新可以由一个、一部分人或者是所

有参与者共同完成。

分布式账本技术能够保障资产的安全性和准确性，具有广泛的应用场景，特别在公共服务领域，能够重新定义政府与公民在数据分享、透明度和信任意义上的关系，目前已经广泛应用于金融交易、政府征税、土地所有权登记、护照管理、社会福利等领域。

## 2) 加密算法

区块数据的加密是区块链研究和关注的重点，其主要作用是保证区块数据在网络传输、存储和修改过程中的安全。区块链系统中的加密算法一般分为散列（哈希）算法和非对称加密算法。

散列算法也叫数据摘要或者哈希算法，其原理是将一段信息转换成一个固定长度并具备以下特点的字符串：如果某两段信息是相同的，那么字符也是相同的；即使两段信息十分相似，但只要是不同的，那么字符串将会十分杂乱、随机并且两个字符串之间完全没有关联。

本质上，散列算法的目的不是为了“加密”而是为了抽取“数据特征”，也可以把给定数据的散列值理解为该数据的“指纹信息”。典型的散列算法有 MD5、SHA-1/SHA-2 和 SM3，目前区块链主要使用 SHA-2 中的 SHA256 算法。

非对称加密算法由对应的一对唯一性密钥（即公开密钥和私有密钥）组成的加密方法。任何获悉用户公钥的人都可用用户的公钥对信息进行加密与用户实现安全信息交互。由于公钥与私钥之间存在的依存关系，只有用户本身才能解密该信息，任何未受授权用户甚至信息的发送者都无法将此信息解密。常用的非对称加密算法包括 RSA、Elgamal、D-H、ECC（椭圆曲线加密算法）等。

## 3) 共识机制

在区块链的典型应用——数字货币中，面临着一系列安全和管理问题，例如：如何防止诈骗？区块数据传输到各个分布式节点的先后次序如何控制？如何应对传输过程中数据的丢失问题？节点如何处理错误或伪造的信息？如何保障节点之间信息更新和同步的一致性？这些问题就是所谓的区块链共识问题。

区块链共识问题需要通过区块链的共识机制来解决。在互联网世界，共识主要是计算机和软件程序协作一致的基本保障，是分布式系统节点或程序运行的基本依据。共识算法能保证分布式的计算机或软件程序协作一致，对系统的输入输出做出正确的响应。

区块链的共识机制的思想是：在没有中心点总体协调的情况下，当某个记账节点提议区块数据增加或减少，并把该提议广播给所有的参与节点，所有节点要根据一定的规则和机制，对这一提议是否能够达成一致进行计算和处理。

目前，常用的共识机制主要有 PoW、PoS、DPoS、Paxos、PBFT 等。根据区块链不同应用场景中各种共识机制的特性，共识机制分析可基于：

- 合规监管：是否支持超级权限节点对全网节点、数据进行监管。
- 性能效率：交易达成共识被确认的效率。
- 资源消耗：共识过程中耗费的CPU、网络输入输出、存储等资源。
- 容错性：防攻击、防欺诈的能力。

### 3. 应用和发展

当前，TCP/IP 协议是全球互联网的“牵手协议”。将“多中心化、分布式”理念变成了一种可执行的程序，并在此基础上派生出了更多的类似协议。然而，回顾互联网技术的发展，当前的互联网技术成功实现了信息的多中心化，但却无法实现价值的多中心化。换句话说，互联网上能多中心化的活动是无需信用背书的活动，需要信用做保证的都是中心化的、有第三方中介机构参与的活动。因此，无法建立全球信用的互联网技术就在发展中遇到了障碍——人们无法在互联网上通过多中心化方式参与价值交换活动。

从区块链技术研究角度看：①在共识机制方面，如何解决公有链、私有链、联盟链的权限控制、共识效率、约束、容错率等方面的问题，寻求针对典型场景的、具有普适性的、更优的共识算法及决策将是研究的重点；②在安全算法方面，目前采用的算法大多数是传统的安全类算法，存在潜在的“后门”风险，算法的强度也需要不断升级；另外，管理安全、隐私保护、监管缺乏以及新技术（如量子计算）所带来的安全问题需要认真对待；③在区块链治理领域，如何结合现有信息技术治理体系的研究，从区块链的战略、组织、架构以及区块链应用体系的各个方面，研究区块链实施过程中的环境与文化、技术与工具、流程与活动等问题，进而实现区块链的价值，开展相关区块链的审计，是区块链治理领域需要核心关注的问题；④在技术日益成熟的情况下，研究区块链的标准化，也是需要重要考虑的内容。

(1) 区块链将成为互联网的基础协议之一。本质上，互联网同区块链一样，也是个多中心化的网络，并没有一个“互联网的中心”存在。不同的是，互联网是一个高效的信息传输网络，并不关心信息的所有权，没有内生的、对有价值信息的保护机制；区块链作为一种可以传输所有权的协议，将会基于现有的互联网协议架构，构建出新的基础协议层。从这个角度看，区块链（协议）会和传输控制协议/因特网互联协议（TCP/IP）一样，成为未来互联网的基础协议，构建出一个高效的、多中心化的价值存储和转移网络。

(2) 区块链架构的不同分层将承载不同的功能。类似 TCP/IP 协议栈的分层结构，人们在统一的传输层协议之上，发展出了各种各样的应用层协议，最终构建出了今天丰富多彩的互联网。未来区块链结构也将在一个统一的、多中心化的底层协议基础上，发展出各种各样应用层协议。

(3) 区块链的应用和发展呈螺旋式上升趋势。如同互联网的发展一样，在发展过程中会经历热甚至泡沫阶段，并以颠覆式的技术改变融合传统产业。区块链作为数字化浪潮中下一个阶段的核心技术，其发展周期将比预想得要长，影响的范围和深度也会远远超出人们的想象，将会构建出多样化生态的价值互联网，从而深刻改变未来商业社会的结构和个人的生活。

#### 2.2.5 人工智能

人工智能是指研究和开发用于模拟、延伸和扩展人类智能的理论、方法、技术及应用系统的一门技术科学。这一概念自 1956 年被提出后，已历经半个多世纪的发展和演变。21 世纪初，随着大数据、高性能计算和深度学习技术的快速迭代和进步，人工智能进入新一轮的发展热潮，

其强大的赋能性对经济发展、社会进步、国际政治经济格局等产生了重大且深远的影响，已成为新一轮科技革命和产业变革的重要驱动力量。

### 1. 技术基础

人工智能从产生到现在，其发展历程经历了6个主要阶段：起步发展期（1956年至20世纪60年代初）、反思发展期（20世纪60年代至20世纪70年代初）、应用发展期（20世纪70年代初至20世纪80年代中）、低迷发展期（20世纪80年代中至20世纪90年代中）、稳步发展期（20世纪90年代中至2010年）、蓬勃发展期（2011年至今）。

从当前的人工智能技术进行分析可知，其在技术研究方面主要聚焦在热点技术、共性技术和新兴技术三个方面。其中以机器学习为代表的基础算法的优化改进和实践，以及迁移学习、强化学习、多核学习和多视图学习等新型学习方法是研究探索的热点；自然语言处理相关的特征提取、语义分类、词嵌入等基础技术和模型研究，以及智能自动问答、机器翻译等应用研究也取得诸多的成果；以知识图谱、专家系统为逻辑的系统化分析也在不断地取得突破，大大拓展了人工智能的应用场景，对人工智能未来的发展具有重要的潜在影响。

### 2. 关键技术

人工智能的关键技术主要涉及机器学习、自然语言处理、专家系统等技术，随着人工智能应用的深入，越来越多新兴的技术也在快速发展中。

#### 1) 机器学习

机器学习是一种自动将模型与数据匹配，并通过训练模型对数据进行“学习”的技术。机器学习的研究主要聚焦在机器学习算法及应用、强化学习算法、近似及优化算法和规划问题等方面，其中常见的学习算法主要包含回归、聚类、分类、近似、估计和优化等基础算法的改进研究，迁移学习、多核学习和多视图学习等强化学习方法是当前的研究热点。

神经网络是机器学习的一种形式，该技术出现在20世纪60年代，并用于分类型应用程序。它根据输入、输出、变量权重或将输入与输出关联的“特征”来分析问题。它类似于神经元处理信号的方式。深度学习是通过多等级的特征和变量来预测结果的神经网络模型，得益于当前计算机架构更快的处理速度，这类模型有能力应对成千上万个特征。与早期的统计分析形式不同，深度学习模型中的每个特征通常对于人类观察者而言意义不大，使得该模型的使用难度很大且难以解释。深度学习模型使用一种称为反向传播的技术，通过模型进行预测或对输出进行分类。强化学习是机器学习的另外一种方式，指机器学习系统制订了目标而且迈向目标的每一步都会得到某种形式的奖励。

机器学习模型是以统计为基础的，而且应该将其与常规分析进行对比以明确其价值增量。它们往往比基于人类假设和回归分析的传统“手工”分析模型更准确，但也更复杂和难以解释。相比于传统的统计分析，自动化机器学习模型更容易创建，而且能够揭示更多的数据细节。

#### 2) 自然语言处理

自然语言处理（Natural Language Processing, NLP）是计算机科学领域与人工智能领域中的一个重要方向。它研究能实现人与计算机之间用自然语言进行有效通信的各种理论和方法。自

然语言处理是一门融语言学、计算机科学、数学于一体的科学。因此，这一领域的研究将涉及自然语言，即人们日常使用的语言，所以它与语言学的研究有着密切的联系，但又有重要的区别。自然语言处理并不是一般地研究自然语言，而在于研制能有效地使用自然语言通信的计算机系统，特别是其中的软件系统。因而它是计算机科学的一部分。

自然语言处理主要应用于机器翻译、舆情监测、自动摘要、观点提取、文本分类、问题回答、文本语义对比、语音识别、中文OCR等方面。

自然语言处理（即实现人机间自然语言通信，或实现自然语言理解和自然语言生成）是十分困难的，困难的根本原因是自然语言文本和对话的各个层次上广泛存在着各种各样的歧义性或多义性。自然语言处理解决的核心问题是信息抽取、自动文摘/分词、识别转化等，用于解决内容的有效界定、消歧和模糊性、有瑕疵的或不规范的输入、语言行为理解和交互。当前，深度学习技术是自然语言处理的重要技术支撑，在自然语言处理中需应用深度学习模型，如卷积神经网络、循环神经网络等，通过对生成的词向量进行学习，以完成自然语言分类、理解的过程。

### 3) 专家系统

专家系统是一个智能计算机程序系统，通常由人机交互界面、知识库、推理机、解释器、综合数据库、知识获取等6个部分构成，其内部含有大量的某个领域专家水平的知识与经验，它能够应用人工智能技术和计算机技术，根据系统中的知识与经验，进行推理和判断，模拟人类专家的决策过程，以便解决那些需要人类专家处理的复杂问题。简而言之，专家系统是一种模拟人类专家解决领域问题的计算机程序系统。

在人工智能的发展过程中，专家系统的发展已经历了三个阶段，正向第四代过渡和发展。第一代专家系统以高度专业化、求解专门问题的能力强为特点。但在体系结构的完整性、可移植性、系统的透明性和灵活性等方面存在缺陷，求解问题的能力弱。第二代专家系统属于单学科专业型、应用型系统，其体系结构较完整，移植性方面也有所改善，而且在系统的人机接口、解释机制、知识获取技术、不确定推理技术、增强专家系统的知识表示和推理方法的启发性、通用性等方面都有所改进。第三代专家系统属多学科综合型系统，采用多种人工智能语言，综合采用各种知识表示方法和多种推理机制及控制策略，并运用各种知识工程语言、骨架系统及专家系统开发工具和环境来研制大型综合专家系统。

当前人工智能的专家系统研究已经进入到第四个阶段，主要研究大型多专家协作系统、多种知识表示、综合知识库、自组织解题机制、多学科协同解题与并行推理、专家系统工具与环境、人工神经网络知识获取及学习机制等。

## 3. 应用和发展

经过60多年的发展，人工智能在算法、算力（计算能力）和算料（数据）等方面取得了重要突破，正处于从“不能用”到“可以用”的技术拐点，但是距离“很好用”还存在诸多瓶颈。实现从专用人工智能向通用人工智能的跨越式发展，既是下一代人工智能发展的必然趋势，也是研究与应用领域的重大挑战，是未来应用和发展的趋势。

(1) 从人工智能向人机混合智能发展。借鉴脑科学和认知科学的研究成果是人工智能的一