中国人工智能学会发展报告系列

——数字社会的风险挑战与治理应对

CAAI 社会计算与社会智能主编

二〇二三年九月

《中国人工智能系列白皮书》编委会

主　　任：戴琼海

执行主任：王国胤

副 主 任：陈　杰　何　友 刘成林　刘　宏　孙富春　王恩东
　　王文博　赵春江　周志华

委　　员：班晓娟　曹　鹏　陈　纯　陈松灿　邓伟文　董振江
　　杜军平　付宜利　古天龙　桂卫华　何　清　胡国平
　　黄河燕　季向阳　贾英民　焦李成　李　斌　刘　民
　　刘庆峰　刘增良　鲁华祥　马华东　苗夺谦　潘　纲
　　朴松昊　钱　锋　乔俊飞　孙长银　孙茂松　陶建华
　　王卫宁　王熙照　王　轩　王蕴红　吾守尔·斯拉木
　　吴晓蓓　杨放春　于　剑　岳　东　张小川　张学工
　　张　毅　章　毅　周国栋　周鸿祎　周建设　周　杰
　　祝烈煌　庄越挺

《中国人工智能系列白皮书——数字社会的风险挑战与治理应对》编委组

主　　编：孟小峰　齐佳音

统　　稿：邓建高　余　艳

编写人员（按姓氏拼音排序）：

邓建高　傅湘玲　何潍涛　胡　俏　黄匡时　李瑾颉

李三希　刘和盛　刘可欣　刘鲁宁　孟小峰　彭松涛

齐佳音　石　慧　宋保振　宋美琦　王雪莹　吴　超

吴联仁　夏园强　宣　琦　余　艳　张明圣　朱宏淼

目 录

序言

关于数字社会风险应对与治理的探讨，从互联网出现至今就一直是学界和业界关注的焦点，但是随着云计算、物联网、大数据、人工智能、区块链和数字孪生等技术的不断发展，我们对于数字世界的认识从 Web1.0，发展到 Web2.0，发展 Web3.0……对于数字社会的认识也从网络社会，发展到虚拟社会，发展到依托数字化、网络化和智能化实现人们各类活动的平台和通行路径的数智社会。进入数智社会，迎接人类的不仅仅是更为便捷、更为丰富、更具创意的新模式、新业态、新产品和新服务，同时也要未雨绸缪防范可能的新风险，需要提前布局与之配套的新规则、新管理、新监管和新治理，才能让数字社会在平衡发展与风险中实现高质量提升。本专辑源于中国人工智能学会社会计算与社会智能专业委员会《数字社会的风险挑战与治理应对》战略研究报告（2022 年 8 月 12 日，杭州发布），但在单独成文时，又做了进一步的完善和扩充。

2022 年伊始，国务院印发《“十四五”数字经济发展规划》，提出要健全完善数字经济治理体系，加强重大问题研判和风险治理，着力强化数字经济安全体系，切实有效防范各类风险。期望本专辑对促进我国数字经济高质量发展起到积极作用。

第 1 章 社会数字化与数字社会

当前，新一轮科技革命和产业变革突飞猛进，科技革命与社会变革加速渗透融合。科学技术尤其是以互联网、大数据、云计算、人工智能和区块链等为代表的数字技术正与社会交往、社会服务、社区建设、社会治理等领域不断渗透融合，社会正在由人与环境构成的物理关系向“万物数字化”和万物互联的数字关系转变，社会关系和社会结构呈现出前所未有的新特征，社会运行的机理与方式正在发生深刻变革。人类社会正在经历一场由科学技术引发的社会数字化变革，并全面融入到一个迈向智能的数字社会。

“加快数字社会建设步伐”已经写入了《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》，包括提供智慧便捷的公共服务、建设智慧城市和数字乡村、构筑美好数字生活新图景等。数字社会的一幅幅美好生活图景正向我们扑面而来，但是学界对数字社会的发展演变、系统结构、行动逻辑和风险识别等理论研究明显滞后。为此，本研究将在系统梳理我国数字社会的发展演变的基础上，分析数字社会的基本结构，探究数字社会的行动逻辑，并对数字社会的主要风险进行识别。

1.1 数字社会的发展演变与未来趋势

数字社会起源于数字和符号的诞生，发轫于计算机的应用，形成于互联网的普及。我国数字社会的形成主要得益于计算机和互联网技术的发展，于 20 世纪 90 年代随着信息互联网的诞生而兴起，于 21 世纪初随着博客、QQ、微信等社交互联网的发展而得到快速

发展，到21世纪20年代随着5G和物联网的快速发展而逐渐形成。我国社会的数字化转型大致可以分为三个时期，即以信息数字化为主的数字社会早期、以社交数字化为主的数字社会中期、以万物数字化为主的数字社会晚期。

（1）信息数字化社会（1994-2004年）

从1994年中国实现了与国际互联网的全功能接入后，我国开启了信息资讯数字化进程，起初主要是科研单位主导互联网基础设施建设，后来随着雅虎（YAHOO）、新浪、搜狐、网易、腾讯、阿里巴巴、百度等门户网站、电子商务和搜索引擎等数字化企业的成立及快速发展，我国以信息资讯为主导的数字化社会得到了快速发展。有学者将这个时期定义为WEB1.0社会，或者定义为信息社会，还有人定义为网络社会，大致较好地描述了早期以信息为主导的数字化社会的早期特征。

（2）社交数字化社会（2005-2018年）

社会交往的数字化是数字社会形成的关键性标志。2005年我国互联网网民首次超过1亿人，意味着我国数字化进程进入了加速发展阶段，而且博客、校内网（2009年升级为人人网）的兴起，以及此后大量社交网站的兴起，比如博客中国、天涯社区、人人网、开心网、新浪微博和QQ空间，以及国外的Facebook、Twitter等，标志着我国网民不仅是一个信息资讯的接收者，而且是信息社会的创造者和传播者。更重要的是，网民已经开始通过互联网等数字化形式开展社会交往，通过数字化的社交软件和平台拓展自己的社会关系。

一旦社会交往数字化后，人类的数字化进程得以加速。2011年

腾讯推出了更具有社交功能的微信（Wechat）小程序。2012 年我国手机网民首次超过 PC 端网民，预示着移动互联网的爆发。在移动手机的普及和腾讯庞大 QQ 用户的推动下，腾讯的微信注册用户呈现爆炸式增长，2013 年年底注册用户超过 6 亿，如今微信用户已经超过 12 亿，覆盖了 14 亿人口大国的主要成年人口。而且微信小程序不断整合日常生活基本服务，日渐成为囊括了购物支付、出行打车、看病挂号、酒店住宿、餐饮快递和零售百货等传统衣食住行的全新的数字生活方式。有学者将这个时期定义为 WEB2.0 社会，大致较好地描述了以社交为主导覆盖生活和工作领域的数字化社会的中期特征。

（3）万物互联的数字化社会（2019 年至今）

当我国社会的数字化转型进入到 2019 年时，数字社会建设进入了一个全新的发展阶段。首先，2019 年三大运营商的 5G 牌照正式商用，为万物互联奠定了一个坚实的基础，意味着我国的远程会议、远程教育、远程医疗、无人驾驶、智慧安防和云游戏等场景应用进入了一个快速发展时期。其次，受到 2019 年年底爆发的新冠肺炎疫情的影响，我国社会的数字化进程明显加快并提速。最近三年，我国物联网载体的各个行业经历了快速的爆发期。到 2022 年 7 月，我国蜂窝物联网终端用户已达到 16.7 亿，与移动蜂窝电话用户数持平；从网络连接看，2022 年连接物的终端数目超过了连接人的终端数，我国物联网进入了里程碑意义上的加速发展期。此外，最近几年，我国的数字经济、数字政务、智慧城市和数字乡村得到了快速发展，不仅教育、医疗、养老、抚幼、就业、文体和助残等重点领域的基本公共服务数字化进程加快，而且购物消费、居家生活、旅游休闲

和交通出行等各类生活场景的数字化进程也不断加快，一幅智慧共享、和睦共治和全民畅享的新型数字生活的美好画面正迎面而来。

（4）数字社会的未来趋势：元宇宙智能社会

2021 年最为火爆的元宇宙（Metaverse）概念掀起了人类对未来社会的广泛想象。元宇宙虽然没有一个标准概念，但大致指的是利用区块链、5G/6G、人工智能、3D、VR/AR/XR、脑机接口等人类最尖端科学技术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。有人认为，元宇宙是互联网技术的最新演化；也有人认为，元宇宙是人类存在状况的最新征候；还有人认为，元宇宙是人类数字化生存的高级形态。但是不管如何，基本可以认为，元宇宙本质上是对现实社会的虚拟化、数字化过程，同时也包含着基于数字化环境的新型社会生态的涌现。元宇宙为我们理解未来数字社会的形态提供了一种丰富的畅想。

尽管人类对元宇宙社会没有一个统一的定义，但是代表了数字社会的未来发展方向。元宇宙社会不仅是一个智能社会，而且是一个虚拟社会和现实社会相互交织、相互融合的超级智能社会，或许与日本提出的超级智能社会“社会 5.0”一致，或许是“社会 6.0”，或者是“Web3.0”。

1.2 数字社会的基本结构

1.2.1 数字社会的虚实二元结构

数字社会是一个由虚拟社会和现实社会共同构成的社会形态。在数字社会中，虚拟社会和现实社会既相互交织又互相区别。一方面，数字社会中的虚拟社会看似虚拟，却又可通过手机、电脑和多媒体终端清晰可见，是客观存在的社会事实，而且虚拟社会中的很

多制度、规则、算法和伦理等又根植于现实社会，与现实社会有着千丝万缕的联系；另一方面，数字社会中虚拟社会又在很多方面与现实社会有不同的表现形式和运行逻辑，尤其是虚拟互动与现实互动的受众对象、互动规则及影响效果都存在明显的不同。数字社会中，这种既可区分又不可区分的虚拟社会和现实社会的二元特征使其与传统社会有明显不同的独特性。

事实上，如果将数字社会的虚拟性和现实性视为数字社会的一种二元分析框架，却可以较好地认识数字社会的基本结构。按照数字社会的虚拟与现实的二元结构的区分，可以识别出数字社会的三种典型类型：一是纯虚拟社会，比如游戏和电影中的纯虚拟社会；二是纯现实社会，又称物理社会，比如饮食聚餐和住宿出行；三是虚拟与现实相结合的数字社会，这可以区分出增强型的虚拟社会、沉浸式的虚拟社会、角色扮演式的现实社会等不同形式的虚实融合社会。

1.2.2 数字社会的系统结构

如果将数字社会看成一个系统，其结构至少包括四个方面的内容。

一是数字设施，即数字社会的基础设施，主要包括网络通信设施、数据存储设施、云计算设施、物联网设施和工业互联网设施等一系列支撑数字社会的硬件设施，这些好比数字社会这栋大楼的砖墙门窗、钢筋水泥、办公桌椅和网络电力等。

二是数字平台，即数字化社会的应用平台，主要包括教育、医疗、养老、抚幼、就业、文体和助残等不同领域的社会服务，这些服务构成了数字社会的一个个应用场景，很多综合性的数字化应用

平台已经开始将不同应用场景的平台进行整合，形成一个综合性的数字化平台。

三是数字算法，即数字社会的算法规则，主要包括自然语言、伪代码、流程图、drakon 图表和编程语言或控制表等不同形式的算法。由于每一种算法背后都有其经济、社会和文化背景，有时隐含着一定的主观性特征，甚至偏见、歧视乃至不公平。正因为如此，有些国家要求算法需要公开并备案。

四是数字互动，即数字社会的社会交往，主要包括浏览网页、观看短视频、发微博、微信聊天、转发朋友圈、点赞评论和网络会议等数字互动，这些数字互动本身会成为数字社会的行动痕迹的一部分，又是个体行动者与其他个体乃至全体行动者进行社会交往的各种形式。数字互动既有例行性的互动，又有突发性的互动；既可以是虚拟的、超越时空的互动，也可以是面对面的现实互动。

1.3 数字社会的行动逻辑

1.3.1 行动者的二元属性

在数字社会中，人的数字化是反映整个数字社会数字化水平的核心指标。人的数字化进程就是人从出生到托育、入园、上学、工作、结婚、生育、退休、养老乃至死亡的整个生命周期的数字化。因此，相较于人的社会化进程，人的数字化进程不仅要早，而且持续时间要长。人作为数字社会的行动者的独特性在于，人在数字化进程中形成了独特的数字虚拟性属性。在数字社会中，人是数字虚拟性和数字现实性相统一的行动主体。人的数字虚拟属性既是虚拟的又是现实的，既是数字的又是物理的。

人的数字虚拟性属性是人作为数字社会的行动者具有超越物理

现实参与到虚拟空间或者虚拟社会的属性，指的是人作为虚拟社会的行动主体可以以虚拟身份或者匿名身份在虚拟环境中开展一系列社会互动的特性。比如，以匿名身份参与到社区、微博、新闻事件的点评和讨论，以匿名的微信号参与到网络交友和群信朋友群讨论中，以及以虚拟身份参与到相关游戏活动中。在数字社会中的一系列社会互动构成了人的虚拟实践。

1.3.2 社会行动的数字化结构

社会行动蕴含着一定的行动逻辑，而这种行动逻辑是行动者所遵循的行动规则和行为规范。在数字化时代，行动者所遵循的数字化行动逻辑既是现实行动逻辑的数字化过程，又是超越现实行动的逻辑而呈现出新的行动逻辑。一方面，数字技术让人与人之间的互动超越了传统社会行动的地域、时间、人数规模、互动频次、传播速度和力度的局限；另一方面，数字技术的发展改变了群体组织的生成方式和互动方式，互联网诞生了大量的微信群、朋友圈和粉丝圈等。群体之间的互动模式、网民的互动结构、互联网的意识形态结构、虚拟社区的形态、家庭的数字化和虚拟化等深刻地改变了社会结构的内涵、特征和弹性，传统的垂直、单向、金字塔结构正向扁平、多向、椭圆结构转变。

数字社会中行动者所遵循的行动逻辑主要体现在数字化的结构性资源上，其中算法便是典型的数字化的结构性资源。数字化结构性资源是建立在整个社会规范基础上的知识体系。数字社会中的数字化结构性资源具有两个方面的特征，一是开放性，即数字化结构性资源基本是开放性的、免费的、透明的，只有这样人人才可以得以习得并遵循；二是参与性，即数字化结构性资源本身是共享的、

参与的，比如网络百科就是具有网民自发更新和完善的特征。当然，正是数字化结构性资源的这两个特征构成了数字社会行动的独特行动逻辑。

1.3.3 社会行动的数字化

数字化时代，互联网的普及使得传统线下的社会行动，比如捐赠行动、志愿行动、体育比赛、上访、游行和抗议等呈现数字化和虚拟化趋势。与传统纸质媒体传播时代不同的是，公民使用互联网来开展社会行动，不仅成本低，组织起来也相对容易，而且互联网的集体行动更加具有传播性和影响力，因此，社会的数字化过程根本上改变了社会行动的组织方式和表现形式，呈现出数字化和虚拟化特征。

1.4 数字社会的主要风险

（1）隐私无限数字化的风险

数字社会最大的风险是个体隐私丧失的风险。数字社会旨在更好地满足人民对美好生活的向往，更大程度地促进人的自由和解放。但是，数字社会是一个数字化高度发达的社会，是一个时时、事事、处处、人人都被数字监控的社会。在这个万物都处于监控的数字社会中，人类好比在摄像头面前“裸奔”，人的隐私存在被无限度数字化的风险。我国于 2021 年 11 月 1 日正式实施了《中华人民共和国个人信息保护法》，这对有效保护个人信息权益具有重要作用，但是在数字化时代，个人信息保护的广度和保护的力度以及效度都面临严峻的挑战。

（2）社会消失的风险

数字社会不能只有数字技术，而没有社会。数字社会的建设初

衷应该是让人类在新的社会形态中活得更加轻松、自由，与人交往更加方便，获得更好的社会交往体验。但是，数字社会的建设过程中往往容易导致过度强化数字社会的数字化特性，而忽略数字社会的社会属性，由此很容易让人产生孤独感，由此引发压抑、冷漠、抑郁、失落等心理健康问题。因此，在数字社会建设中，要更加强化群体互动和社会交往方面的技术和场景设计。

（3）人性消失的风险

数字社会可能会有大量的机器人存在，人类与机器人的互动将会更加频繁。在这个互动过程中，人类可能在与机器人的互动中慢慢地将人类看成机器人，甚至将部分人类看成比机器人还差的人群，长期如此，导致部分人类会异化为如同机器人般的人类，并逐渐丧失人性。因此，在数字社会建设中，要弘扬人性的光辉。

1.5 小结

回顾我国社会的数字化进程，从社会交往的数字化，以及数字社会的基础设施、数字化生态等综合来看，我国数字社会基本建成，并进入数字社会的中后期，且朝着智能社会方向发展。当然，从我国国民的数字素养和数字社会的制度建设，以及法律规范来看，我国数字社会建设依然还在路上。

伟大的社会实践呼唤伟大的社会理论。数字社会建设亟需研究分析数字社会建设的理论脉络、系统结构和行动逻辑。本文认为，数字社会是由虚拟社会和现实社会有机融合的新社会形态，为此提出了数字社会的虚实二元理论框架，识别出数字社会的三种典型社会类型——纯虚拟社会、纯现实社会和虚实融合社会。从数字社会的系统结构来看，至少包括数字设施、数字平台、数字算法和数字

互动四大支柱。从数字社会的行动逻辑来看，人作为数字社会的行动者是数字虚拟性和数字现实性相统一的行动主体，在现实社会或者虚拟社会中学习和掌握数字化的结构性资源，并在此基础上开展社会行动，而行动者本身的社会行动又在修改和更新数字化的结构性资源。

数字社会对人类发展而言既是机遇又是挑战。数字社会让人类生活在一个更加数字化和智能化的社会，与此同时，数字社会也让人类面临一些风险，比如隐私无限数字化的风险、社会消失的风险和人性消失的风险。因此，数字社会建设既要把握机会，更要回避风险。

第 2 章 数字社会世界观与数字社会风险治理的主、客体界定

随着大数据、云计算、人工智能、区块链、移动互联网和 5G/6G 等现代信息技术的迅速发展，以及其在生产、工作、学习和社会生活中的广泛应用，人类已经步入数字社会时代。数字社会是继农业社会、工业社会和信息社会之后，以信息技术的广泛应用为时代特点，集数据采集、数据存储、数据分析、数据挖掘和数据决策等为一体的更高级社会形态。然而，“技术从来就是好坏参半”，具有天然的中立性，它“既赋予我们创造性，也赋予我们毁灭性”。进入数字社会，迎接人类的不仅是更为便捷、更为丰富、更具创意的新模式、新业态、新产品和新服务，同时也要未雨绸缪防范可能的新风险，需要建立与数字社会相适应的数字思维观和数字风险观，需要提前布局与之配套的新规则、新监管和新治理，才能让数字社会在发展中实现高质量提升。

2.1 数字社会世界观

从物理世界到数字世界，人际关系、交易关系、生产模式和消费形态等都发生了根本性的变化。在人际关系上，全球范围内几乎每一个人都能通过数字化媒体进行广泛交流，数字化人际关系既“疏”又“远”，其无形中对诸多社会运动和社会理念产生了影响和挑战；在交易关系上，市场交易关系、服务交易关系、产品交易关系、劳动交易关系和货币交易关系等均因数字化而发生前所未有的改变，数字交易关系非常“便捷”，但又充满“不确定性”；在生产模式上，

生产设备智能化和生产过程自动化，数字化生产模式的改变既“高效”又“精确”，但数字化生产模式加速了人类的进化速度和“去社会化”程度；在消费形态上，数字化消费平台、数字化物流平台、数字化推广平台和数字化售后平台等为人类的产品消费和服务消费，创造了更加富有想象力的途径和形式，但也为人类创造了某种“陷阱”。因此，数字社会时代需要重新审视数字世界的方法论和风险观。

2.1.1 数字社会的网络资源观

数字世界方法论首先要树立互联网的资源观，将互联网不仅仅看作技术、看作平台，更应该看作资源、看作数字经济时代的独特资源，用来优化配置不受时空限制的各类资源，获得最大化的竞争能力。任何人或组织都可以使用数字世界的网络资源，寻求效用最大化。

2.1.2 数字社会的数据要素观

数据要素是数字世界与物理世界相互联动的纽带。第四次技术革命实质是以数据为要素通过数字孪生的形式实现物理世界与数字世界的双向赋能，实现数据要素的价值化和流动性。数字世界方法论要确立数据是生产要素的观念，制定数据作为生产要素的经济规则，以此激活各类网络资源和数字资产。

2.1.3 数字社会的算法规则观

数字世界基于数据，经由算法，载于软件，通过平台，提供社会化应用。数字化应用的规则体现于算法，数字世界的算法本质是物理世界的运行规则，但算法具有独有特性，算法的全面使用亦可建立新的规则。因此，算法背后是规则，规则背后是法制，数字世界需要建立算法规则观，需要建立与算法相配套的法制规则。

2.1.4 数字社会的软件设施观

数字世界天然运行在软件之上，软件是数字世界的“厂房”，基础核心软件是数字世界建造“厂房”的“地基”。谁建造“地基”，谁就制定标准，谁就在网络世界拥有最大话语权，并获得最大收益。因此，数字世界软件的“地基”属性、“载体”属性迫使人类建立软件设施观，需要对软件设施进行规范化、科学化管理。

2.1.5 数字社会的开放开源观

数字世界天然具有开放性，开源开放将逐步成为社会集体意识。工业社会围绕工业商品属性（独占、排他）而形成封闭排他的社会意识。未来的数字社会将围绕数字商品属性（共享、协作）而形成开放开源的社会文化。数字世界的开放开源，有助于人类社会文明共同进步，有助于人类科技发展水平提升。

2.2 数字社会风险治理主体

纵观理论界和实践界现有研究成果，学者们对数字社会的多元主体进行了有益的探索。学者危红波提出面对数字社会中隐私泄露、数字鸿沟、算法黑箱与歧视、大数据杀熟、身心健康损害、网络舆情危机和意识形态风险等众多问题，针对技术缺陷、利益驱动、矛盾爆发、心理问题、法制短板和监管不力等原因，需要建立现代化的数字社会风险治理主体架构。数字社会的形成促使人们树立新的数字社会世界观，新兴数字社会世界观催生现代数字社会风险观。传统社会治理在相对“封闭”区域中展开，其政治与权力组织形态是建立在科层制的物理空间之上，治理主体体现出一种“中心与边缘”结构。数字社会治理显著有别于传统社会治理，数字社会形态下的权力技术化、数据资产化，进而导致数字社会治理权力体系的

弥散化、数字社会权力形态的扁平化和数字社会权力的虚实交叠，数字社会传统的中心、边缘、半边缘权力格局受到一定程度的冲击，因此，迫切需要建立与数字社会相配套的现代化风险治理主体架构体系。

2.2.1 数字社会风险治理的“核心”主体——政府组织

数字社会风险治理是国家治理体系的重要构成。因此，数字社会治理的多元主体中，政府组织必然是其中重要的一员。在当前社会背景下，政府组织在数字社会风险治理发挥着无法替代的功能，但政府组织如何进行数字社会风险治理的角色定位，政府组织又如何科学实施数字社会风险治理，这是现代数字社会风险治理亟需解决的社会问题。

从数字社会的特征来看，数字社会具有传统线下社会的特点，又有不同于传统社会的显著特征。数字社会是众多机构、网民的交互行为，他们之间的互动同样需要遵循基本的社会制度和社会结构。但不同的是，数字社会具有时空跨越、多中心、虚拟化和“光速化”等特征，上述特征致使现有政府组织的角色、介入模式无法在数字社会空间进行简单的复制、移植和实施。当前，数字社会风险治理呈现出悖论现象，一方面数字社会高速发展，人们在数字社会中的卷入程度持续演化，数字社会“乱象”丛生，数字社会风险绵延不断；另一方面，在数字社会风险治理中政府组织角色不清、介入方式不明、介入工具不全，政府组织在数字社会风险治理中显得“力不从心”“手忙脚乱”。因此，探索政府组织在不同场景下参与数字社会风险治理的模式、角色、路径、工具、机制和效应等问题，有待于广大学者和产业界深入的思考和研究。

2.2.2 数字社会风险治理的“技术”主体——网络运营者

数字社会时代人类的信息浏览、社会交往、数据通信、交通出行和网络购物等基本活动均与各类网络平台紧密关联，数字社会时代存在的网络平台呈现出高度技术化特征。在网络运营者提供服务和享受收益的基本逻辑前提下，无论是在服务供给、技术掌握和规则建立，还是数据持有和隐私保护等方面，网络运营者既是数字社会风险的责任主体，又是数字社会风险治理的技术主体。与西方发达国家一样，我国的网络运营者涵盖网络平台所有者、管理者，以及利用他人所有或者管理的网络平台提供相关服务的网络服务提供者，包括国家基础电信运营商、网络信息服务提供商和提供网络服务的各类信息系统运营商。数字时代网络运营者法律责任和管理责任涵盖了信息发布与传播责任、网络消费交易秩序维护责任、数据安全保护责任等，这些责任既体现了提供网络服务时承担的管理义务，也体现了我国数字社会风险治理的行政权力色彩，在一定程度上体现了我国数字社会风险治理多元主体共同治理理念，体现了数字社会风险治理中治理模式的创新。因此，探索网络运营者在数字社会不同场景下风险治理中的权力、义务、角色、响应机制、激励机制、管理责任和行业公约等，是我国数字社会风险治理的重要研究内容。

2.2.3 数字社会风险治理的“专业”主体——社会组织

现有学者关于社会组织参与数字社会风险治理问题进行了有益探索。如徐顽强等学者论述了非政府组织参与数字社会治理的模式、职责和路径。系统分析诸多不同主体参与数字社会治理的研究成果，发现我国政府、社会组织、网络运营者、新旧媒体、网民等多元主

体，参与数字社会风险治理逻辑不清晰、治理依据不明确、参与模式不成熟、参与路径不明朗，多元主体参与数字社会风险治理的权责同样不清晰、工具不完备、协作方式不科学、绩效难计量，数字社会风险治理急需对这些问题背后的关键科学问题开展系统性研究。

在数字社会发展浪潮高速推进过程中，社会各领域的利益诉求和社会矛盾在网络空间释放和扩散，不同区域和不同领域的社会组织成为数字社会风险治理体系的“专业”性主体，在法律解析、心理疏导、矛盾调解、科普推广、信息溯源和观念引领等众多领域发挥了重要作用。社会组织扎根民间，与社会矛盾基层民众正面接触，可以及时发现征兆、识别苗头、理性沟通和传达民意，在数字社会风险出现前、出现中、出现后均可发挥重要价值。当前我国社会组织相关法规不健全，社会组织内部建设不完善，社会组织总体发展不成熟；特别是社会组织与政府的合作机制、沟通机制、信息共享机制、激励机制，缺乏有效的法律保障。虽然社会组织在公益性救助上取得良好成效，但是在数字社会不同领域的风险治理上，参与的广度、深度和效果等各方面有待于持续优化。

2.2.4 数字社会风险治理的“不可或缺”主体——公民

数字社会兴起，微博、博客、论坛、网络社交、网络评论和视频分享等网络服务深入百姓生活，表层上看是公民自主享受网络服务，但从社会学和政治学角度看，网络是多元主体参与数字社会风险治理的重要途径和渠道，为数字社会风险治理多元主体融合治理模式发展创造条件和可能。数字社会网民参与社会治理具有多重含义，首先，众多普通网民在网络环境中自由创造、讨论、评论和转发，对社会治理形成强大的影响力，进而演化成公民参与数字社会

风险治理不可或缺的重要主体；其次，公民在网络环境下的风险揭露、意见表达、权益维护、网络监督，实质为我国健全民主监督、舆论监督机制的发展和完善；另外，公民通过网络就公共事务阐述观点、发表意见、交流看法，致使数字社会环境下的网络空间成为思维汇聚和公共舆论形成的重要阵地，在众多数字社会风险治理中发挥了重要作用。

当前，我国数字社会处于高速发展过程中，广大网民是数字社会风险治理中众所周知的、不可或缺的重要主体。但是，多数情况下仍是一种自发式的、相对“无序”的参与模式，不同场景下数字社会风险治理的公民参与模式、参与路径、参与机制、保护制度、激励方式等有待于学界和产业界不断深入研究。特别是在网络参政、网络议政、网络监督、个体矛盾和社会矛盾化解、谣言识别、网络互助等领域，需要在相关法律、行政法规、部门规章、司法解释、规范性文件等方面进行系统性建设，规范和保障公民参与数字社会风险治理的义务和权力。

2.3 数字社会风险治理客体

现代信息技术塑造了一种全新的社会适应机制，为人类工作和生活提供了极大的“便利”，数字生产、数字出行、数字健康、数字消费、数字金融、数字社区和数字家居等众多数字场景似乎使人们进入了一个自我管理、自我实现的高度自由的数字社会。由此引发了法律价值上数据正义观、代码正义观和算法正义观对人类正义认识的颠覆，引发法律关系上权利关系的根本性重塑和结构性转向；传统的行政法规、部门规章、司法解释、政策文件和规范性文件难以应对数字社会需要；数字社会中的网民主体、基础设施、系统平

台、算法规则和数字内容等因素均成为新的风险客体。

数据社会框架中的软件算法规则、硬件网络资源、数据要素内容和数字社会法规体系有别于传统社会形态的认知，众多元素均可成为数字社会的新兴风险来源。如数字内容本身蕴含的潜在数据伦理风险；产业数字化、数字产业化过程形成的数字资产，以及数字资产交易存在的未知风险；数字社会形态形成过程中，社会治理过程数字化、社会规则算法化和社会媒体智能化过程中的算法裁决过度风险；数字社会发展过程中各类硬件、软件和数据等基础设施可能存在的风险；数字社会化进程中催生的各大平台可能出现阶段性和行业性数据垄断，数据垄断本身及行业数字平台潜在的失控风险；数字社会形态中的组织和个体享受数字社会红利的同时，从“用户”属性悄然演化成为“产品”属性，期间造成难以避免的公域、私域过度信息采集和信息泄露风险；还有数字社会形态下数据资产创造者是全民，但数据资产所有权并非公民，数字社会看似抹平了信息鸿沟但又重构了新形态的信息不对称，进而形成新的数字红利不平等和社会分化风险。数字社会不同风险客体的机理、特性和规律尚未得到充分的解释，有待于广大学者深入探索和研究。

2.4 小结

数字社会伴生新兴世界观和风险观，数字社会风险治理是国家经济社会发展的重要议题。本文对数字社会的数字世界观进行了初步分析，并进一步以数字社会风险治理的主体和客体进行了归纳和分析。研究结论可为数字社会风险治理领域学者和产业界提供借鉴和参考。

第 3 章 数字社会的风险挑战

3.1 失真数字内容传播风险与耦合网络沟通机制

数字内容是二十世纪中期以来数字技术、数字媒体持续发展与变革的产物，是指以数字形式提供的产品或服务，包括文本、数据、图像、音频及视频等。近年来，得益于互联网及网络基础设施的普及与进步，数字内容市场发展迅猛，海量数字内容不断生成、传播及消费，涉及数字阅读、在线学习、网络游戏、数字音乐、网络视频、网络直播等。随着 Web2.0 时代网络数字内容体量的激增，社会公众通过 Web 网页、移动应用软件等平台访问并观看文字、图像或视频的数字内容消费越来越流行，数字内容创作的速度和传播效率明显加快。数字科技的广泛运用不仅改变了公众接收信息的内容形态，也改变了数字内容在公众间的传播逻辑。一方面，数字科技被广泛应用于人们的生活与工作中，互联网和在线社交网络等线上平台逐渐成为主要的传播渠道，并融合移动传播成为主要传播形态。另一方面，相比于传统内容，数字内容经历了从格式转换到延展创造的发展。数字内容不仅能够提供有意义的信息，也搭载着传播者与受众的社会联系。数字内容实质上是内容与传播策略的组合体，即数字内容本身已经蕴含与人的关系。不同于传统内容生产中的受众需求判断与满足，数字内容不仅供媒体等在单向传播时使用，还可以为社会公众在多向交互传播时广泛应用。其受众可以即时参与再传播及多次传播，也可以二次创作、延伸创作或成为阐发讨论的切口与接口，并且在交互传播中对受众反馈、转发、评论的呼应和对原发内容的进一步补充都已成为数字内容生产及其传播过程中的

重要组成部分。

在数字内容分发与其影响形成的过程中，每一公众个体都可能参与传播，受众地位与功能发生重要变化，其已经成为数字内容首次传播及再传播的重要力量。在此情况下，人们普遍对互联网环境下数字内容质量产生忧虑，主要原因在于大量失真数字内容存在于互联网中，对数字内容接收者主观感知、决策制定和相关行为造成严重误导，并对消费者的知情决策带来潜在威胁。社会公众在失真数字内容的误导下所作出的各种行为决策可能会引发严重的后果和悲剧。例如，数字时代下电子药店的崛起将供应处方药的权利交给了市场，从而对公众产生了诸多不利影响，这种负面效应包括给消费者造成了接收具有误导性的失真健康信息、不当使用药品等风险和伤害。此外，频发使用社交媒体进行在线交流也会对个体之间的线下社交和亲密关系的形成产生重要影响，我们认为线上社交是否有助于线下面对面直接交流失真数字内容取决于人们使用社交媒体的方式以及线上社交与线下社交的关系。病态化使用社交媒体在线交流某一失真数字内容很可能会对个体间线下交流该失真数字内容的错误观点和理念存在抑制作用。因此，刻画和研究线上线下多个信息渠道交互作用下的失真数字内容传播机理和演化规律，对数字经济的发展具有重大积极意义。

兴于世纪之交的网络科学在过去二十年间取得重要进展。网络科学中的耦合网络理论为深入研究上述问题提供了崭新的视角。耦合网络是对单一传播渠道构成的单层复杂网络的进一步拓展，也是多路传输网络中的一种特殊形式。耦合网络重点关注真实复杂系统中各元素之间在多个相互作用的不同层次的交互影响，将同一组元

素间的多种不同作用方式区别对待，以开展更为精确的复杂系统理论与仿真研究。相关研究的主要思路是将社会公众抽象为节点，并将公众通过线上线下多种信息渠道中的其中一种渠道传播扩散失真数字内容形成的关联关系抽象为一种类型的连接，在此基础上从耦合网络理论的视角出发，将公众通过以上多种不同渠道传播失真数字内容构成的复杂系统看作一个耦合网络，即一组各层节点均相同但各子网络层的边均不同的多层耦合网络。通过研究这一耦合网络的多个子网络层间失真数字内容传播的交互影响机制，来探讨公众通过线上线下多种信息渠道互动如何交互作用影响失真数字内容的传播，探索基于线上线下多渠道的失真数字内容传播机理，为提出有效的预防及治理政策提供必要的理论基础和可靠的数量依据。

3.1.1 失真数字内容

数字内容产业作为数字经济的重要组成部分，不断创造出新的市场和商业机遇，逐渐成为知识经济和信息社会的重要推动力。后疫情时代更是加速了数字内容产业的发展，数字内容迎来了显著增长的消费需求。与此同时，随着数字技术的进步，广大民众都拥有了采、写、编、评、发的能力，对自己所要表达的内容和意义拥有更多的自主权和支配权，由此引发了大规模的用户内容生产。数字工具的简单易行使数字内容的生产者进一步泛化，用户自主生产、自由转发、自助评论、自动搜索，UGC、PUGC 愈发成熟，即使专业的内容生产机构往往也要@用户账号，来带动内容的传播。故而数字内容是一种聚合式的云传播模式，内容是不同生产者间合作生成的动态演化，来源不可具体区分间离。普遍而言，数字技术赋予了数字内容生产公众性，使得个人可以基于兴趣、意愿、机会、能

力或激励进行内容生产。任一用户都能在数字平台上创造并实时分享数字内容，并且这种行为可以看作一种维系和拓展社会关系的社交行为。然而，数字技术提供的数字环境增加了社会公众利用信息资源的同时，也为失真数字内容的传播创造了有利环境。

失真数字内容是指个体相信了客观上并不完全准确的数字内容的情形。与失真数字内容紧密相关的一个概念是谣言。谣言是在一定时间内特定群体对其关心的、未经证实的话题信息的扩散与传播。基于该定义，谣言可以被证实也可以被证伪。笔者认为，谣言的传播虽然容易造成负面的群体性结果，但由于谣言的时效性较短，且有被证实的可能，所以对个体产生的深层次影响较为有限。然而，失真数字内容由于客观上的不准确性，对个体的认知、心理、信念和情感等都有可能产生深远的影响。因此，二者在概念的内涵和影响程度上都具有显著差别。与失真数字内容有关的另一个概念是伪数字内容，或称为虚假数字内容。笔者认为失真数字内容未必完全是伪数字内容，真实数字内容在传播过程中由于信息的解构和重组、部分内容丢失、外围噪音等因素的影响也可能造成一定的数字内容失真。基于此，笔者主张在数字内容搜寻语境中使用“失真数字内容”一词指代与现有客观科学证据有一定偏差，对数字内容接受者主观感知造成误导的各类数字内容，同时使用“伪数字内容”指代与现有客观科学证据完全相反、对数字内容接受者直接造成极其严重的后果的各类虚假数字内容。本文主要研究失真数字内容在传播中的风险极其预防和治理策略。

失真数字内容的主要危害在于其误导性，使社会公众产生偏差认知，令公众对相关数字内容可信度的判断产生失误，从而轻信失

真数字内容中所宣称的信息与观点，继而进一步影响决策与行为。越来越多的证据表明，通过大众媒体或社交网络传播的失真数字内容会对社会公众产生显著负面影响。譬如，营养方面的失真数字内容使消费者不合理地夸大某些食物的益处或危害，而导致不均衡的营养摄入。再如，急救医学上的失真数字内容使人们错误地估计特定急救技术的收益与风险。又如癌症及重大疾病的失真数字内容对人们的不恰当寻医行为具有较强解释力。综上，预防及治理失真数字内容传播所带来的社会风险显得尤为重要，而理解基于多种信息渠道的失真数字内容的传播机理与演化机制，则是实施干预的前提和关键。

3.1.2 失真数字内容传播动力学模型

耦合网络上的动力学主要探索局部节点或连边的行为是如何在整个耦合网络中扩散的。在建模中先将节点进行必要的分类，如失真数字内容传播者、失真数字内容接收者和失真数字内容免疫者等。在此基础上，依据节点在单位时间内通过多种线上及线下渠道中的其中一种渠道交流失真数字内容的次数进一步对他们进行分组，如在微博平台中单位时间内交流 k 次失真数字内容的传播者、在微信朋友圈中单位时间内交流 k 次失真数字内容的传播者等。在此基础上，基于多层耦合网络理论，建立多种信息渠道构成的多层耦合网络中考虑层间网络结构关系的失真数字内容传播动力学模型。模型中的主要参数包括：各渠道构成的各层子网络中的失真数字内容传播率系数、失真数字内容自主产生率系数、失真数字内容暂时遗忘率系数以及失真数字内容免疫率系数等。

另外，考虑接收者在接受并内化吸收失真数字内容时会思考该

信息内容实施起来所需的成本和所获收益，拟将各层子网络中的失真数字内容传播率，分别设为公众个体认为自身采取该失真数字内容所表述的行为时所需的平均成本与所获得的平均收益的非线性函数。同时，为了刻画并研究各渠道间失真数字内容传播的交互影响，在建模中需映射出多重耦合网络中各子网络层间网络结构的相关性。基于以上传播机理分析，可以构建失真数字内容在多渠道构成的多层耦合网络中的传播动力学模型，即各种不同类型节点数量随时间变化的微分方程模型。

3.1.3 多渠道融合治理

数字科技全面应用于内容生产和传播领域，不仅改变了传媒生态和传播格局，也深刻改变了内容形态、传播工具和传播规律等影响数字内容传播运行的重要构成元素。信息基于多渠道的交互传播是互联网时代产生的诸多传播现象的最重要特征，也是数字科技赋予传播的最重要能力。互联网成为信息载体后，包括网站、移动应用、社交平台在内的各种互联网应用都提供了交互功能，使用户不仅可以获得相关资讯、信息或服务，还能实现用户与互联网应用、用户之间的多种交互。如社会化媒体是一类基于 Web2.0 思想和技术的互联网应用，它允许人们进行深度交互，如创建、分析、交换数字内容等。社会化媒体的实际形式包括微博和博客、社交网络站点（SNS）、虚拟社区、内容（视频、音乐等）分享站点、协作内容生产（如维基百科）等多种类型，它们之间的界限越来越模糊，功能越来越综合，成为一类重要的数字内容交互平台。比如人们可以在交易网站、论坛和博客上写下大量他们的评分和观点，消费者在购买产品或者服务前，通常会阅读与产品或服务相关的评价以对要购

买的产品有一个粗略的认知，再来决定是否购买。基于此，在交互传播思维下，本文提出融合使用多渠道的失真数字内容治理思路和策略。例如，对微博、虚拟社区、维基百科等多种线上传播渠道及线下传播渠道中失真健康信息的传播与扩散进行均衡、协同、适度地预防及治理。

稳定性理论主要研究时间趋于无穷时微分方程解的性态，在自然科学、工程技术、环境生态、社会经济等方面有着广泛的应用。可以从动力系统的零平衡点的稳定性分析中验证以上分析中所得的传播阈值的准确性，也可以从其正平衡点的存在性中给出验证。本研究可以从所建立的基于耦合网络的失真数字内容传播动力学模型的零平衡点的稳定性分析来计算传播阈值。传播阈值是一个重要参量，其取值大小可以用于判断某一所考察的失真数字内容能否在整个耦合网络中扩散开来，或是逐渐消亡。如失真数字内容以一定的概率从一个节点传递到另外一个节点时，传播概率一般存在一个临界阈值，如果传播概率高于该临界阈值，初始少量失真数字内容会迅速扩散到整个网络；反之，失真数字内容在传播过程中会迅速消亡（与初始状态无关）。

利用上述建立的模型和阈值，对所建立的模型进行数值模拟与仿真实验，预测失真数字内容能否在所考察的社会群体中持续传播，以及最终传播规模，即传播的范围。针对某一关键参数，在其他参数取值不变的前提下给出该参数的取值变化对失真数字内容传播的影响，以及对各类节点的数量随时间延续的变化有何影响，从而提出相应的预防和治理失真数字内容传播干预策略，同时比较各参数变化所对应的各种策略的有效性。

此外，“餍足心理”理论指出，个体长时间与某种刺激物接触后对该刺激物的敏感度将会逐渐降低乃至反应消失，转而会对另一种新的刺激产生兴趣。因此，在治理失真数字内容所拥有的成本和资源均有限的前提下，如果集中所有资源仅对某一种线上传播渠道中失真数字内容的传播进行管控，在抑制部分公众使用该渠道进行失真数字内容传播的同时，反而更容易激发传播者通过其它可替代传播渠道继续扩散失真数字内容。本研究在模型研究、案例分析和数值模拟实验的基础上认为，仅针对一种传播渠道进行治理，对该失真数字内容传播的管控效果并不能产生理想的治理效果，并且该失真数字内容在社会系统中的最终传播规模可能明显变大，持续扩散的时间也可能明显更久，这将对社会系统的正常运转产生更大负向影响。而如果合理分散资源和成本对线上和线下等多种渠道中失真数字内容的传播进行融合治理，将更有效地抑制部分公众利用各种渠道传播失真数字内容的行为，降低部分公众利用各种渠道进行失真数字内容传播的可能性。更易于缩小该失真数字内容通过线上线下等多种渠道在社会公众中的传播范围，并且该失真数字内容在社会系统中也将更快地消失。

3.1.4 小结

为了对数字时代下失真数字内容传播这一社会风险进行有效预防及治理，本文引入耦合网络理论与传播动力学模型，建立考虑线上线下多种渠道及其之间相互影响的失真数字内容传播动力学模型。通过模型分析与仿真实验得出相应的治理策略，即不应仅对线上线下其中一种特定渠道中失真数字内容的传播进行防治，而应该均衡分散资源对线上线下多种渠道同时进行预防和治理，这样做可以以

更低成本、更小代价、更有效地降低失真数字内容传播的扩散度，使其在社会公众中迅速消亡。

网络中存在着一些具有较强信息交互能力的用户，其影响力的覆盖范围和作用强度均有别于普通用户，成为该网络中信息传播的关键力量，并且从信息的扩散广度、扩散速度、扩散深度等维度对信息传播形成不同影响。因此，探究失真数字内容交互关系网络中的关键力量，有助于理解如何增强关键力量影响力，从而提升失真数字内容治理效率。鉴于此，本文另外给出以下两点策略建议：

（1）针对具有较高网络权力的失真数字内容传播者的认知纠偏策略。失真数字内容传播形成的关系网络中具有较大度的节点将对整个耦合网络中失真数字内容的扩散产生重要影响。这类节点可以根据其他节点的行为活动迅速将不合理的内容反馈传播开来，最终左右失真数字内容走势。笔者在以上研究中发现，应在多渠道均衡治理模式的基础上，积极对以上基于网络结构的具有较大传播影响力的关键节点进行认知纠偏，使其在内心深处改变自己的认知并形成正确认知，可以更有效地对失真数字内容的传播进行管控和治理。

（2）对具有极高单次传播效力的超级传播者进行有针对性的管制。超级传播者即在单次互动中具有极高传播能力的节点。超级传播者参与失真数字内容传播时，其传播能力和破坏力也很强，可以通过少量沟通导致更多的普通传播者和信息接收者受到影响，从而增大该失真数字内容的传播范围，给治理带来极大阻碍。因此，笔者同样建议对这些具有极高单次传播效力的超级传播者进行有针对性地管控和治理。

本文将耦合网络理论及其在沟通机制中的应用引入到数字时代

下的失真数字内容多信息渠道传播机制研究中，为现有研究提供了崭新的理论视角，并提出有效预防和管控失真数字内容通过线上线下多个信息渠道在社会系统中扩散的有效策略，进一步促进数字技术的平稳发展和普及。本研究主要从传播机理上阐述了深入研究线上线下多个渠道交互影响下失真数字内容传播规律的具体思路和研究结论，未来可进一步获取实际数据以更好地验证失真数字内容传播模型和相关治理政策的有效性。

3.2 数字货币发行中的风险、机遇和挑战

近年来，区块链技术、以区块链技术为基础的比特币以及其他数字货币的快速发展引起了大家的广泛关注。尽管支票和现金等业务依然扮演着重要的角色，但是新兴技术对传统支付方式和原有支付系统构成一定挑战。数字时代区块链技术及电子支付迅速崛起，世界各国的中央银行都在高度关注和思考如何趋利避害地推动数字货币。中国央行早在 2014 年就开始着手于数字货币的研究和发行工作，并成立了数字货币研究所，集中开展数字货币发行及相关技术的研究工作。现今，数字人民币试点工作正在稳步推进。

3.2.1 数字货币的主要风险

数字货币是以数字形式存在，并基于网络记录价值归属和实现价值转移的货币。广义上的数字货币包含虚拟货币、电子货币和加密数字货币三个部分。虚拟货币如游戏网站积分和 QQ 币，其仅在发行方规定的范围内流通，并且不与法币直接挂钩，属于非货币。电子货币是存在于电子账户中，通过银行等金融机构实现电子化支付的货币，如网上银行和支付宝支付。我们通常说的数字货币一般是指加密数字货币，其又分为私人数字货币和法定数字货币。私人

数字货币由非国家主体发行，在被法律支持、大众广泛接受的支付环境中使用，如比特币和以太币。法定数字货币是基于国家信用且一般由一国央行直接发行的数字货币，是法币的数字化形式。

目前比特币等加密数字货币的问题主要聚焦在：一是缺乏中心化的管理机构，因为像比特币这样的数字货币没有任何中心机构来发行和维护；二是比特币这些数字货币提供了更多的匿名性；三是数字货币系统的易受攻击性；四是数字货币的监管还处于初级阶段；五是数字货币的发行和流通都是基于互联网的，容易在全球流通。正是因为加密数字货币存在这些问题和特点，导致的主要风险有犯罪风险、信用风险、金融风险、技术风险、声誉风险、盗窃风险和投机风险。

①犯罪风险。由于数字货币的匿名性，可能存在为洗钱、毒品交易、偷漏税等提供便利，而又无法追踪导致犯罪风险高发。随着数字货币的快速发展，数字货币导致的敲诈、洗钱、偷税漏税和非法交易的风险已暴露出来。

②信用风险。传统的货币由国家信用背书，现有私人数字货币的去中心化特征导致其货币缺乏法定机构和政府支持，数字货币的获得过程没有任何发行机构参与，无任何锚定和信用担保，一旦交易出现争议，缺少传统金融领域的仲裁手段，容易引发信用风险。

③盗窃风险。数字货币的去中心化和匿名性，用户的钱存在钱包里，而不是存在戒备森严的银行里，一旦发生账户盗窃，根本无法跟踪和识别账户的归属。因此，比特币等数字货币丢失和被盗的风险远高于现实世界的真实钱包。

④技术风险。区块链技术刚刚起步，技术仍不够成熟，分布式

记账无法支撑全球商品货币的巨大交易量，加密数字货币存在交易安全性差的缺点，再加上基于区块链技术的交易没有严格的时间框架，存在技术风险。

⑤声誉风险。数字货币持有者容易把数字货币价格波动导致的问题和国家及央行的监管联系起来，进而导致政府和央行的声誉受损，存在声誉风险。

⑥金融风险。数字货币的全球易流通性，且监管比较困难，再加上加密数字货币市场相较于其他传统资产市场承受着更高的尾部风险容易对现有金融体系造成冲击，进而导致金融风险。

⑦投机风险。由于数字货币的价格波动较大，容易引发市场炒作和投机行为，交易加密货币可能会吸引那些表现出投机更严重的赌徒，随着数字货币价格的上升，一些投资者倾向于持有更多的数字货币，由于数字货币价格的波动又容易给投资者造成巨大损失，导致投机风险。

3.2.2 中国发行法定数字货币的机遇

尽管比特币等数字货币的去中心化所导致数字货币存在监管空白、投机风险、价格波动大、价值稳定性差、洗钱犯罪等一系列的弱点，但以区块链技术为基础的数字货币是经济社会发展特别是数字经济时代的必然产物，数字货币的去中心化和匿名性等特性也大大降低了交易的成本，提高了效率，从而受到欢迎和追捧。但大规模的价值交易如果得不到秩序保障，必然会损害到不特定的群体。货币是一致同意下的社会共识，私人数字货币目前还难以在社会共识基础上获得普遍认可。比特币等私人数字货币当下依然缺乏内在价值基础，价格波动剧烈，难以有效发挥货币职能，由国家主权货

币来背书的数字货币才有机会得到广泛的社会共识。法定数字货币是法定货币在数字世界的延伸和表现，是法币的数字化形式，是一种新的货币形态，是基于国家信用且一般由一国央行直接发行的数字货币。2018 年国际清算银行从发行主体（是否央行发行）、形式（实物或数字货币）、可获得性（广泛的或受限制的）、技术（是否点对点）等四个货币主要属性分析央行法定数字货币，提出如下“money flower”模型（图 3-1）。

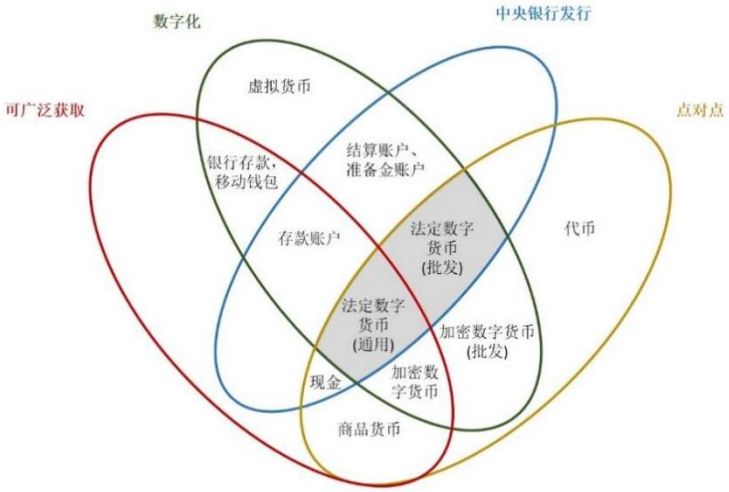

图 3-1 “money flower”模型图

数字货币是货币形态发展的一次重大革命，是与数字经济相匹配的新时代支付变革，央行发行法定数字货币势在必行。

（1）非现金交易快速上涨的市场机遇

全球支付方式偏好继续由现金和信用卡向数字钱包转移，非现金交易量正在迅速增长，数字钱包占据主导地位。2021 年，数字钱包在中国电子商务总交易额的占比近 83%，仅自 2020 年以来，这一份额就上升了超过 10 个百分点，其中 2021 年亚太地区 92%的数字

钱包交易额来自于中国；中国的电子商务市场预计在 2021 年至 2025 年将增长 55%，达到全球规模最大的 3.2 万亿美元；而现金支付份额将持续下滑，预计至 2025 年在销售点总交易额中的份额将自 2021 年的 10%降到 3%。特别的，近年来的新冠肺炎疫情，进一步助推了非接触的线上交易的需求，线上消费不断升级，带动互联网平台业务大幅增长。虽然国际货币的使用具有较强的网络效应和路径依赖，一般情况下主导货币的地位难以被动摇；但货币的发展史显示货币的价值属性不断向交易的便利性妥协，数字货币可借助互联网渠道实现数字货币的普适性和泛在性。线上交易的快速上涨和社会对线上交易便利性的市场需求为数字货币的发行和推广提供了广阔的市场机遇。

（2）中国数字金融全球领先的历史机遇

用产品周期理论看中国生产的很多产品，如手机、汽车等大多数日常使用的产品都是在发达国家完成研发的；但令人振奋的是，数字金融的很多原创实验和创新都在中国，在数字金融革命中，中国走在了世界前列。2021 年，微信月活跃账户数超过 11 亿，支付宝全球客户超 12 亿；天猫“双 11”，支付宝自主研发的分布式数据库 OceanBase 每秒处理峰值达到 6100 万次，展示了巨大的数字金融运算能力。中国的移动支付规模是美国的近百倍，中国的电子支付系统已经全球领先。我国在数字货币领域的基础、技术、实践均处于世界领先地位；2022 年一季度，人民币在全球外汇储备中所占比重为 2.88%，已连续多个季度攀升，不断刷新记录；人民币占全球外汇储备比例的提升，意味着各国官方愿意以人民币形式来持有金融资产，表明这些国家对人民币的高度信心；数字人民币将为人民币国

际化提供更便捷的渠道，数字货币领域实现弯道超车的历史机遇已经到来。

（3）重塑全球金融生态的时代机遇

习近平总书记关于“世界百年未有之大变局”的内涵丰富，其核心是一个“变”，本质是世界秩序重塑。数字化的全空域、全流程、全场景、全解析和全价值与金融业的不断融合，将颠覆传统、重塑全球金融生态，引领金融业进入一个全新的时代。目前有两方面的力量可能挑战美元在国际货币体系中的中心地位，重塑全球金融生态：一个是技术的变革，另一个是主观的动机；技术方面，以智能化、网络化和数字化为核心的新一轮工业革命将重塑国际竞争格局，为后发国家提供赶超的“窗口期”；数字货币将给国际货币竞争增加数字技术和应用的新维度。主观方面，在大数据和区块链技术的驱动下，构建一个新的全球清结算系统已经成为很多国家的共识，目前已有二十多个国家政府投入建设分布式记账系统，通过区块链试水跨境支付。

“物质扩张”能催生货币国际化的需求，也是国际货币化的基础，2021 年中国 GDP 总量为 177340.6 亿美元，依旧是世界第二大经济体，增长达 8.1%；这为人民币的流通提供了“物质扩张”的基础。此外，人民币加入 SDR 货币“篮子”，已经有超过 60 个国家和地区将人民币纳入外汇储备；再加上“一带一路”建设为数字人民币的发行和流通提供了市场准入机会。这些“百年未有之大变局”为全球金融生态重塑提供了可能，也为中国数字人民币的发行、推广和人民币的国际化提供了难得的“窗口”机遇。

3.2.3 中国发行法定数字货币的挑战

中国人民银行早在2014年成立了研究团队就数字货币的发行和业务框架等问题进行研究，2017年1月成立数字货币研究所。2020年初，疫情作为黑天鹅事件，由于纸币、硬币存在携带病毒的风险，市场普遍关注疫情是否会成为一个契机促使数字人民币加快落地，数字人民币目前正处于扩大实地试验范围的阶段，也面临着诸多挑战。

（1）风险将与虚拟加密数字货币长期共生

快速发展的虚拟加密数字货币，不能一禁了之，与其相伴的各类风险不会因为国家禁止而社会安然无恙。在美国股市的历史上，第一次熔断是发生在2008年金融危机中；但是在2020年疫情中，在3月9日至3月18日短短的10天内，美国股市连续熔断了4次。在这种极端的全球经济环境下，比特币的价格在3月13日一度下挫到4106.98美元，相比3月9日的最高值8177.79美元，下降了50%。在这种极端经济环境下，大家普遍认为比特币不再是“数字黄金”；但是到2020年4月6日，比特币价格又强劲地拉升到了7271.78美元，几乎回到了股市熔断前的价格。近两年来，虽然比特币价格时有波动，但一直高于2020年熔断时的水平，这显示比特币在市场中的强劲生命力。这些私人数字货币的强势存在及其区块链等支撑技术既能为中国发行数字货币提供有益参考，同时对中国法定数字货币的推广应用，特别是国际化交易等形成挑战。

匿名性和去中心化在成就加密数字货币的同时，也成为其“藏污纳垢”、“野蛮生长”的诱因，导致对加密数字货币的有效监管成为世界性的难题。我国虽然采取强硬的“监管”措施，但难以从

根本上杜绝私人加密数字货币的交易、挖掘和持有等活动，由于世界各国监管力度松紧不一，我国依然困扰于通过加密数字货币实现洗钱、犯罪、投机等，因此迫切需要通过加强对加密数字货币监管技术的研究，配合现有的监管政策，才能更加有效地遏制加密数字货币的“野蛮、无序”态势；为数字人民币的发行奠定基础。针对我国当下监管需求，可重点研究：①加密数字货币交易异常识别；②加密数字货币市场操纵识别；③加密数字货币 FoMO 情绪识别；④加密数字货币中的“空气币”识别。最终为我国从技术层面有效监管加密数字货币和发行数字人民币提供理论支撑。

就数字人民币的发展而言，虚拟加密数字货币与其存在着竞争与合作并存的复杂关系。数字人民币的重要意义在于保存数字货币优势的同时强化货币监管与控制，其作为官方发行的数字货币与虚拟加密数字货币之间存在着市场竞争，数字人民币等法定数字货币的应用与推广将对虚拟加密数字货币产生冲击，二者将作为货币市场的重要竞争者长期存在。但是，虚拟加密数字货币的技术积累与应用经验对于数字人民币的发行与推广有着重要的借鉴价值，其得以在世界范围广泛使用的相应机理与机制对于数字人民币的应用和国际化推广有着现实意义。自 2008 年中本聪提出比特币以来，各种“虚拟加密数字货币”在世界范围内快速扩张，其扩散速度与波及范围远超传统货币。这种在世界范围快速扩散推广的经验对于深入参与世界市场的中国而言有着重要的价值。一国货币的影响与其应用的范围和广度有着紧密的联系，“虚拟加密数字货币”的扩散规律和应用机理与相关机制将可能为“数字人民币”的国际化应用与推广提供智慧。为此，针对“虚拟加密数字货币”与“数字人民币”

之间既竞争又合作的关系，不仅要研究“虚拟加密数字货币”监管技术，以完善监管机制，还应当深入探讨“虚拟加密数字货币”扩散规律与其广泛使用的原因及相应机理机制。通过借鉴这些经验完善我国的“法定数字人民币”，使我国的“法定数字人民币”能够在未来与各国“法定数字货币”及“虚拟加密数字货币”的竞争中占据优势。

（2）各类金融货币生态中蕴含潜在风险

2022 年第一季度，全球外汇储备总额为 12.55 万亿美元，其中美元资产为 6.88 万亿美元，所占份额为 58.88%，仍是全球最广泛持有的储备货币;但人民币在全球外汇储备中所占比重暂时仅为 2.88%，而世界银行公布的数据显示，2021 年全球经济规模 96.1 万亿美元，美国经济规模为 22.996 万亿美元，全球占比为 23.93%；中国经济规模为 17.727 万亿美元，全球占比 18.45%，人民币的外汇储备和中国 GDP 权重严重不对称。中国的互联网、大数据和金融科技的发展与美国的差距，要比中国金融业与美国金融业的差距小得多；中国法定数字货币某种程度上肩负着“人民币国际化”的重任。比特币在某种程度上已经证明了区块链技术的优越性；同时在 Libra 项目的设计和推动过程中，Facebook 试图结合区块链的理念和技术搭建强大的“货币互联网”，从而使项目本身具有强大的颠覆性和竞争力。然而，零售所要求的高并发对区块链技术的计算速度形成了严峻挑战，一定时间内能够处理的交易规模有限。

各类数字货币之间以及其与市场的深度作用机制需要加快研究，具体包括：①政府监管政策对虚拟加密数字货币的影响效应研究；②虚拟加密数字货币与中央银行数字货币在市场竞争中的演化博弈

研究；③中央银行数字货币发行对宏观经济的影响；④中央银行数字货币对整体金融系统的生态效应；⑤多形态货币形态共生的均衡条件与监管工具开发；⑥中国数字货币如何应用区块链技术助推人民币的国际化。

（3）国际竞争博弈与国家安全风险

数字金融成为各国在数字经济时代的重要利益竞争场，数字货币是重构国际支付体系的重大历史机遇，是大国在数字经济时代博弈的重要金融工具。在这一背景下，应特别关注用于跨境支付的数字货币研究，具体包括：①研究数字货币对国际贸易、国家主权货币、国家金融体系稳定的影响；②研究数字货币在国际贸易中的各国相关法律适用问题及监管问题，设计可行的国际贸易非传统货币交易体系；③研究跨国家间的数据共享和数据安全保护，研究如何利用数字货币实现国际间公平贸易；④研究支持各国跨境贸易的数字货币定位、目标、技术实现框架与实施推进路径；⑤研究跨境数字货币的使用边界、实物贸易与数据贸易的差异、强势法币国与弱势法币国之间的均衡；⑥研究高并发情形下的交易效率等国与国之间的贸易问题和技术方案的可信问题。

此外，基于数字货币的跨境贸易将给国际贸易带来全新的支付方案，在提高支付安全、提升贸易效率、创新贸易方式的好处之下，也会冲击到国家的税收政策、法币的稳定性、贸易的可监管、外汇的可控性等其他方面，需要深入研究。

3.2.4 小结

数字经济蓬勃发展，数字新金融也势不可挡，将来的国家间的战争形态，将从高精尖武器争霸转换为以信息战、金融战二大战争

为主线。从当年的伊拉克信息战到如今的欧美制裁俄罗斯发起乌克兰战争的金融战、舆论战，都是现实的案例。面对数字货币带来的机遇与挑战，我国要提前布局、积极试点、不断验证和纠错，才能够在未来的国际金融竞争中占得先机。

3.3 数字社会算法裁决过度风险与公平计算

数字经济的核心是数据，大数据、人工智能等技术对经济社会中各种数据广泛和实时的采集和处理，使数据像工业时代的石油一样成为数字经济时代最重要的生产要素。算法是强制给定的有限、抽象、有效和复合的控制结构，在一定的规则下实现特定的目的。随着人工智能技术的广泛使用，一个算法社会正在到来。算法在带动经济增长、提高经济效率、丰富和便利生活的同时，在信息传播、个人隐私、弱势群体利益等方面的负面影响和风险也暴露出来，需要高度重视并加强算法规制。数字社会算法过度裁决造成的负面影响和风险主要包括以下四个方面。

第一，侵害用户隐私。一方面，大多信息的获取没有得到消费者的许可，属于非法采集；另一方面，个人虽然是数据的生产者和所有者，但是这些数据一旦被互联网公司获取后就脱离其最初所有者的控制，个人不但无法知晓自己的信息被用于何处，而且在后续的使用中可能会威胁个人的隐私，甚至可能对个人的财产和人身安全造成损害。

第二，造成算法歧视。在传统的商品市场上，商品的功能和质量高度标准化，价格不同很容易被发现，而互联网服务通常是基于特定时间和特定场景的，时间和场景不同，市场上的供需关系就不同，并由此形成不同的市场价格，这就造成“大数据”杀熟更难被

察觉，也更难被举证。

第三，形成“信息茧房”。人们被算法圈定在“信息茧房”里，被动地接受算法让人们看到的信息。在社会层面，“信息茧房”还会形成用户观点的极化，造成不同群体之间的交流障碍，甚至由于思想的偏狭引致群体间的误会，催生极端行为，引发社会矛盾和冲突。

第四，损害弱势群体。数字经济的发展和算法的广泛使用不会自动地平等惠及每一个人，弱势群体反而会成为算法的受损者。

3.3.1 数字社会算法风险

（1）算法裁决过度风险

2020 年 9 月，自然杂志子刊《机器智能》（Machine Intelligence）刊登了一篇文章《我们赖以生存的算法》，尖锐批评英国政府运用算法预测学生高考成绩，认为用学生平时成绩预测高考成绩是对既有社会分化的固化和极化。人们似乎认为，算法是人类共同面对的、前所未有的敌人，算法之恶已令社会深恶痛绝。为此，如何应对算法的利弊，让算法更好地服务于人类社会？

数字社会的算法与社会的强互动形成算法对生产、生活和治理的实时参与，适合德国社会学家马克斯·韦伯（Max Weber）提出的权力分析框架。在数字社会的算法呈现其社会影响力时，计算社会学权威大卫·拉泽尔（David Lazer）呼吁人们关注算法的影响，并提出“社会算法”（social algorithm），指出计算程序有能力掌握、评估我们的期待并提供个性化的体验。其实这是一种误导，把算法提供的建议当成强制，以为算法掌握了某种社会权力，甚至是霸权。

近些年，随着中美欧出台的算法治理相关法律的落地和算法治理文献数量的爆发式增长，北京大学中国社会与发展研究中心主任

邱泽奇在归纳既有文献基础上，认为数字社会算法过度裁决带来的风险主要有三类。

一是算法风险。在个体层面，算法是强化着“信息茧房”、带来人的认知窄化风险；在市场层面，算法遵循商业逻辑，酝酿着监控资本主义的风险；在国家层面，算法隐藏着被特定利益集团用于社会控制和政治权力再生产的政治风险。

二是算法侵害。现实生产和生活中已经出现了算法歧视、算法偏见、算法共谋、算法垄断、算法黑箱、算法遮蔽、算法短视、算法霸权、算法操纵和算法剥削等与算法关联的侵害，给人类社会生活、经济，甚至政治带来了负面影响。

三是，除算法带来的风险和侵害，还将算法影响上升到制度层面，认为算法权力和算法规则是触发规则竞争、权力竞争，甚至规则垄断、权力垄断的根由。

（2）算法裁决与公平

算法作为数字化时代的产物之一，成为帮助组织快速进行决策及判断的重要工具，被广泛应用于组织管理中。由于算法决策在现实中日益普遍的应用，并对组织和身处其中的成员产生重要影响，研究者们也开始越来越关注成员对算法决策的感知和反应。关于员工对算法的感知，现有文献中存在两种观点。一种是算法欣赏（algorithm appreciation），认为算法更加高效、准确、客观，人们会更加信任算法决策。另一种则是算法厌恶（algorithm aversion），认为人工决策可以允许员工在决策过程中有更高的参与和控制感，人们会更倾向于选择人工决策。

与这两种观点紧密相关的是算法决策（相对于人工决策）对员

工公平感的影响。近年来，算法决策如何影响员工的公平感也受到关注，但并未呈现一致的研究结果。具体来说，一些研究发现，相对于人工决策，算法决策会带来更高的公平感，因为算法决策具有更高的一致性、无偏性，以及使用的信息更加准确等。然而，另一些研究则发现算法决策未必会导致更高的公平感，因为人们认为算法没有考虑质化信息、不够情境化、无法进行主观评价。这些不一致的研究发现，意味着在决策者类型（算法决策 vs.人工决策）与员工公平感的关系中可能存在重要的调节因素，使得两者的关系在不同情境下呈现出不一样的结果。

探讨这些潜在的调节因素对于我们更全面、具体地理解算法决策对员工公平感及其后续行为的影响非常重要。根据归因理论，员工会对涉及自身的决策进行解释，而这种解释会影响他们对于该决策的反应。特别地，人们面对有利和不利的结果，会进行不同的归因。关于自我服务归因偏差（self-serving attribution bias）的研究发现，当面对有利的决策结果时，员工更倾向于归因于自身因素，因此其公平感较少受外部因素（如决策者特征）的影响。相反，当决策结果不利时，员工更倾向于归因于外部因素，如决策者。此时，由于员工对于算法和人工这两类决策者的无偏性、一致性、准确性等方面有不同的感知会产生不同程度的公平感。因此，基于归因理论，有学者认为决策结果的有利性可能是影响决策者类型与员工整体公平感之间关系的一个重要调节变量。决策者类型与决策有利性的交互作用不仅会影响员工的公平感知，还会影响员工后续的行为结果。其中特别值得关注的是员工的偏差行为。员工在应对算法的过程中表现出的针对组织或社会的偏差行为已经引起了广泛关注。

关于公平的研究也发现，当员工感到不公平时，会展现出更多的偏差行为。因此，魏昕等提出了下图 3-2 的研究模型，探讨决策有利性如何调节决策者类型，通过员工公平感对员工偏差行为产生的间接影响。

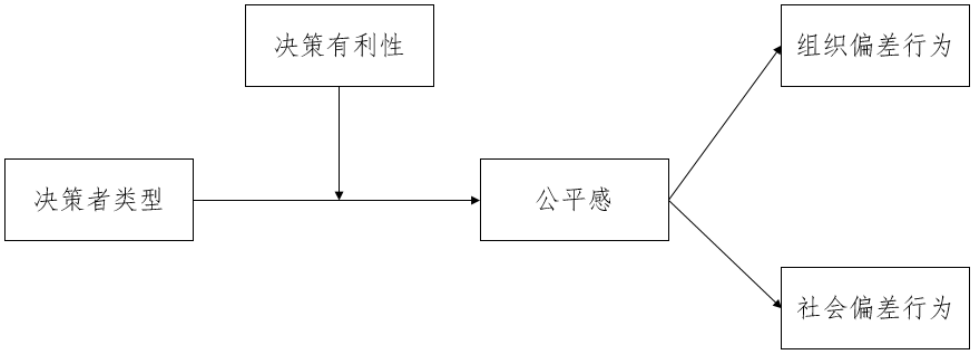

图 3-2 研究模型

3.3.2 算法公平与公平计算

（1）算法公平

2022 年 8 月 12-14 日，在浙江杭州举行的 CAAI 第七届全国大数据与社会计算学术会议上，中国人民大学信息学院孟小峰教授进行了题为《算法公平与公平计算》的报告，从历史的角度揭示算法公平（AF）的发展过程，并提出从社会偏见、数据偏见和模型偏见三个维度认识算法公平，揭示其相互作用关系。2022 年，马特乌斯-多拉塔（Mateusz Dolata）等在国际信息系统期刊 Information Systems Journal 发表文章 A Sociotechnical View of Algorithmic Fairness（关于算法公平的社会技术观点），其中算法公平被认为是一种新出现的技术，它可以减轻自动决策中的系统性歧视，为改善信息系统（IS）的公平性提供机会。他们基于对 310 篇文章的系统分析，首先对当前关于算法公平的讨论中的基本假设提出了问题；其次，通过将算

法公平理论化为一种社会技术结构来回应这些假设；最后为信息系统研究人员提出了方向，以通过追求独特的对社会技术性算法公平的理解。呼吁并采取全面的办法来处理算法公平。

（2）公平计算

孟小峰教授指出，从广义上来看，一切使用算法手段在社会矛盾纠纷中实现公平的方法都可算作算法公平范畴。并将实现广义算法公平的过程称之为“公平计算”（Fairness Computing）。公平计算应该成为算法治理的重要手段，其核心要素包含公平定义与量化、公平监测预警、公平方法权衡等内容。2021 年梅赫拉比（Mehrabi）等发表论文“A Survey on Bias and Fairness in Machine Learning（关于机器学习中的偏见和公平的调查）”，提出了，数据、算法和用户互动反馈回路中的偏差框架（下图 3-3）。

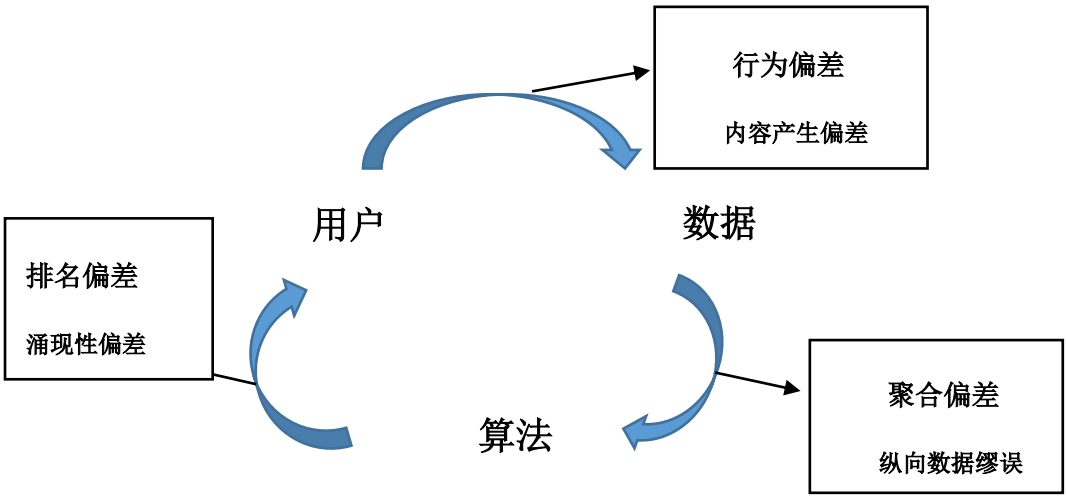

图 3-3 数据、算法和用户互动反馈回路中的偏差框架

希拉-米歇尔（Shira Mitchell）等和马特乌斯-多拉塔（Mateusz Dolata）等提出了公平计算的数学模型。从跨领域的角度给出了算法中偏差的处理方法：

预处理（Pre-processing）。预处理技术试图改变数据，以便消除潜在的歧视。如果允许算法修改训练数据，那么就可以使用预处理。

内处理（In-processing）。内处理技术试图修改和改变最先进的学习算法，以便在模型训练过程中消除歧视。如果允许改变机器学习模型的学习程序，那么在模型的训练过程中，可以通过在目标函数中加入变化或施加约束来使用内处理。

后处理（Post-processing）。后处理是在训练后，通过访问模型训练过程中未涉及的保留集来进行的。如果算法只能把学到的模型当作一个黑箱，没有任何能力来修改训练数据或学习算法，那么只能使用后处理。模型最初分配的标签在后处理阶段会根据一个函数重新分配。

3.3.3 算法治理

（1）算法治理法律法规

算法对人类社会生活影响的普遍性和深刻性使算法成为堪比自然环境的人工环境，算法影响的利弊两面性，以及算法侵害的不断出现将人类推入算法治理时代。邱泽奇指出，中美欧在算法治理领域的探索实践呈现不同格局，美国从防范算法侵害入手，形成了政府和第三方的问责模式；欧盟从数据保护入手，逐渐与美国的问责模式汇流；作为数字时代三方国际力量之一的中国，从 1994 年制定第一部相关行政指令《计算机信息系统安全保护条例》到目前，共出台了 60 多部相关法律法规和行政指令。从时间书序来看，可以认为我国的算法治理是从总体安全入手实施治理，但尚未形成有法理逻辑和明确操作路径的算法治理模式。表 3-1 列出了对算法治理起到关键影响的法律法规。

表 3-1 算法治理代表性法律法规

时间	法律法规	核心
2009	《互联网信息服务管理办法》	是行政许可对服务内容相关的约定，属于数据经营监管类的行政指令。
2017	《网络安全法》	奠定了算法治理的基本方向，强调数字基础设施、网络信息的安全与保障。
2019	《电子商务法》	聚焦交易安全，部分内容涉及对算法结果的治理。
2019	《儿童个人信息网络保护规定》	虽聚焦于个人信息，强调了信息安全，对儿童信息相关运营商的责任约定。
2020	《网络信息内容生态治理规定》	在关于内容服务的第十二条涉及了算法治理，是从内容安全出发的。
2021	《数据安全法》	将网络安全进一步延伸到数据领域，指导简历健全数据安全治理体系。
2021	《个人信息保护法》	涉及算法治理，出现了与欧盟 GDPR 相似的内容，安全指向依然清晰明确。
2022	《互联网信息服务算法推荐管理规定》	第一部针对算法的管理规定，针对不正当竞争等算法侵害的规范。

（2）算法治理的框架

基于欧美在算法治理中的经验教训，并结合目前我国算法治理存在的问题，清华大学公共管理学院曾雄和中国信息通信研究院胡坚波等学者就算法治理框架的构建提出以下方向性的政策建议。

①在治理目标上，实现算法可问责与算法经济高质量发展。算法治理的根本目标在于促进算法经济的高质量发展，在日益复杂的国际竞争形势下，提高自主技术创新的能力和水平是国家安全发展的根本出路。②在治理对象上，建立一套共性的规则和标准后，统筹考虑多元应用场景基于对拓宽算法治理对象的考虑，可以制定一部《算法问责法》实现算法的综合治理，并为算法问责建立一套完

整的责任机制，包括明确问责主体、被问责对象、问责方式和程序以及问责事项等。③在治理手段上，补强司法救济和技术治理措施推行“遵循伦理的设计”机制可以确立算法的底层伦理标准，实现事前干预。利用算法评估提前研判安全风险，利用技术手段预警日常风险隐患，利用行业标准明确算法运行底层逻辑。④在治理模式上，积极构建多元主体协同的治理模式。发挥政府在算法治理中的主导作用，规范企业切实履行算法治理主体责任，引导行业组织积极参与算法治理，鼓励公众参与算法治理，完善社会监督，深化算法治理国际合作。

3.3.4 小结

在社会规则算法化的过程中，针对算法运行流程本身，算法设计人员会将自己的主观意识，如问题的定义、数据的收集、模型的选择都会有意无意融入到算法规则，这对于算法的结果会造成影响。同时算法过度决策会给公众造成巨大的困扰，比如公众只注意自己选择的东西和使自己愉悦的通讯领域，从而形成信息茧房；算法继承了来自于社会本身的偏见问题，并且进一步放大；在做出判断和决策的时候，被算法操纵等等。这一系列风险在使用人工智能和大数据算法时，是不可避免的，在使用算法时，我们需要对风险识别、形成机理、评估和预警，以及治理等问题开展研究，为人工智能、大数据等新技术平稳安全落地，服务人类社会。

3.4 数字社会基础设施风险与新兴安全技术发展

数字社会已经成为我国的新型社会形态，与之匹配的“新型基础设施建设”工作也正在稳步推进。“新基建”是整个数字社会的基础，蕴含着中国经济发展的新趋势，其引发的热潮受到了广泛的关

注，各方的视角普遍聚焦于“新基建”带来的新动能和新机遇。然而，在“新基建”的建设和发展之际，其安全问题和风险应对需要提前部署。

数字社会既是技术革新，也是社会革新。现实世界与虚拟世界以数据为信息桥梁，使得它们之间的联系愈加紧密，这也令世间万物能更直观地表示、更深度地推断和更有效地调控。虚实空间融合的时代背景，众多新兴技术与数字社会的融合发展，也伴随着众多安全和风险问题从虚拟世界威胁到了现实世界。在宏观层面上，我们从数据安全、算法安全、系统安全三个角度探讨数字社会基础设施发展带来的跨界风险；在微观层面上，我们关注海量异构终端互联带来的安全短板，聚焦数据安全问题引发的 AI 信任危机，审视应用场景多元化伴随的风险多元化，并深入挖掘各要点的发展现状、面临问题、创新技术。

综上，本文围绕数字社会“新基建”涉及的 5G、人工智能、数据中心、工业互联网、云计算、区块链等为代表的众多新兴技术的应用和迭代，从以下四个方面来展开讨论新兴技术下的安全问题和风险，及其已有的发展和突破。

3.4.1 虚实空间融合导致攻击的跨界威胁

数字技术将现实世界产生的海量信息数字化，在虚拟空间中进行处理和计算，再通过具体的控制指令反馈到现实世界，建立起了虚实空间之间的桥梁。“新基建”的发展使得虚拟空间和现实空间的边界变得模糊，虽然进一步加快了两者的融合，但也让虚拟世界的安全问题“渗透扩散”到了现实世界。针对新兴技术载体的攻击将会直接或间接地影响现实世界，危及人身安全和公共安全，造成不

可挽回的损失。

（1）发展现状

自提出推进“数字中国”建设以来，众多数字产业在国内蓬勃发展。正如习近平总书记所说：“数字技术正以新理念、新业态、新模式全面融入人类经济、政治、文化、社会、生态文明建设各领域和全过程，给人类生产生活带来广泛而深刻的影响。”近年来，大数据、互联网、云计算和人工智能技术的应用为数字社会发展提供了强劲动力，5G和区块链网络的建设推广为数字经济提供了重要支持，以及一系列新兴技术的创新开发为数字时代的到来奠定基础。

在虚实空间融合的时代背景下，数字技术带来便利的同时，各种以新兴技术为载体的风险威胁也从统一、多层次、规范化的现实世界延伸到多元、去中心、开放化的虚拟世界，在数据、算法、系统三个层次上存在一系列问题，具体情况如图 3-4 所示。

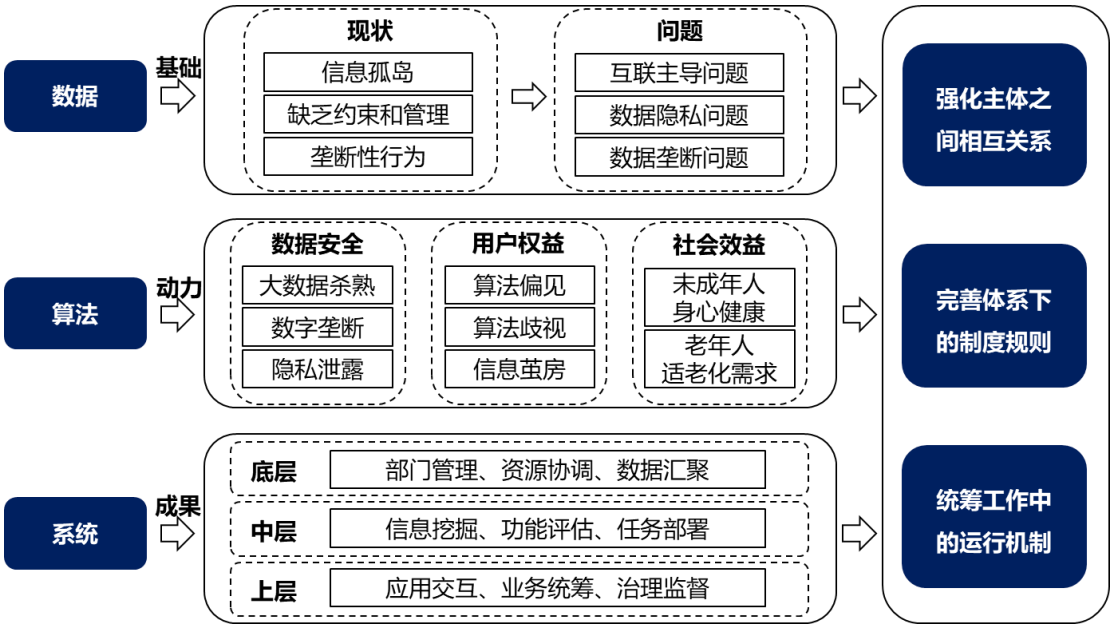

图 3-4 虚实空间融合下的多层次跨界威胁

数据是虚实空间融合的基础。万物互联的数字化社会的发展治理，首先需要明确数据的互联主导问题，避免信息孤岛的形成；其次数据隐私问题，采集机构门槛的模糊和采集信息范围的扩大，表明目前缺乏对数据强有力的约束和管理；最后数据信息垄断问题，企业平台掌握海量个人、企业、政府的社会数据信息，这种垄断性行为在运营、技术、道德等风险面前，都会造成不可估量的损失。随着数字社会的深入发展，数据层面所存在的威胁也会不断频繁化和剧烈化。

算法是虚实空间融合的动力。算法赋予数据差异化的应用场景和多样化的利益价值的同时，也被从技术、运营、效益等方面进行严格规范。数据安全角度，有违合理规范，比如大数据杀熟，即根据消费者的习惯偏好等特征，在交易上实行不合理的差别对待；用户权益角度，有违公平公正，比如算法偏见、算法歧视、“信息茧房”等不良影响；社会效益角度，有违群体权益，比如诱导未成年人沉迷网络、过度消费等。综上，虚拟世界需要有与之相适应的制度规则作为支撑。

系统是虚实空间融合的成果。不同层面的数字系统有机协作，为现实世界的提供特定的功能服务。在底层，资源协调能力不足，需实现部门管理、资源协调、数据汇聚；在中层，算法评估制度缺乏，需实现信息挖掘、功能评估、任务部署；在上层，多元共治局面待完善，需实现应用交互、业务统筹、治理监督。因此，系统是虚实空间融合中发展水平的体现。

（2）治理方案

面对虚实空间融合引发的种种问题，不同层面响应积极的治理

政策与策略。数据层面，强化主体之间相互关系，比如国家标准化管理委员会等五部门印发《国家新一代人工智能标准体系建设指南》，提出人工智能领域安全与隐私保护标准；算法层面，完善体系下的制度规则，及时响应新兴技术发展出现的问题，比如国家互联网信息办公室等四部门出台《互联网信息服务算法推荐管理规定》，推动算法安全可信和技术创新之间的有效平衡；系统层面，统筹工作中的运行机制，比如中共中央办公厅、国务院办公厅印发《关于加强科技伦理治理的意见》，对加强科技伦理治理制度保障、深入开展科技伦理教育和宣传等方面进行具体部署。

综上所述，面对虚实空间融合导致攻击的跨界威胁，应充分发挥了虚拟世界对现实社会的支撑作用，同时，兼顾现实社会对虚拟世界的约束作用，实现新兴安全技术的创新利用和稳固发展。

3.4.2 海量异构终端互联带来的安全短板

5G“新基建”、人工智能、数据中心等技术快速发展，带动了国家电网、制造业、商业、医疗等垂直行业迈向数字化、智能化。各个垂直行业海量的异构、计算能力不均的设备终端不断接入互联网，这些安全性低、性能有限的设备极大地限制了已有安全工具的开发、运行和维护，使得安全问题解决方案存在滞后，缺乏统一的安全防护体系，难以应对未知的威胁，具体情况如图 3-5 所示。

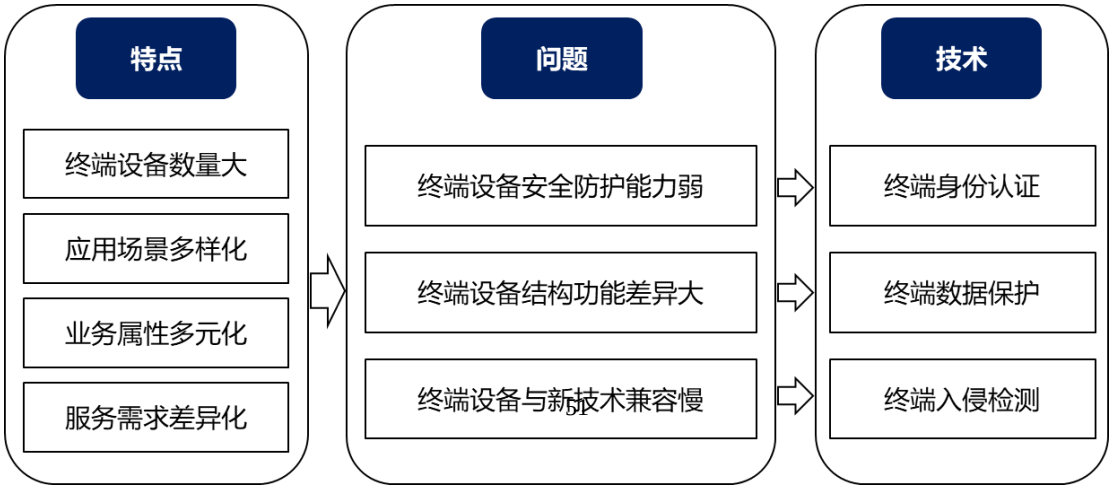

图 3-5 异构终端互联的问题及相关技术

（1）安全短板

异构终端互联网络有着终端设备数量大、应用场景多样化、业务属性多元化以及服务需求差异化等特点。近年来，针对智能家居、摄像头、工业物联网设备等终端设备的安全事件频发，究其原因，海量异构终端互联主要存在以下短板：

一是终端设备安全防护能力弱。因设备的数量和场景差异需求大，所以以低功耗、低成本设备为主，没有足够空间采用复杂的安全机制，同时，设备的关联性使得业务网络存在“单点攻破，全局瘫痪”的风险。

二是终端设备结构功能差异大。终端设备类型复杂多样、软硬件环境各异、匹配和防护差异明显；终端设备数量大、分布面广，组网结构复杂，均增加了管理和防护的难度。

三是终端设备与新技术兼容慢。移动边缘计算在提升数据处理效率的同时，也增加了可选择的网络攻击对象；IPv6 使海量终端设备拥有专属的地址成为可能，但也使设备直接暴露在互联网上；容器技术在提升开发效率的同时，也使得数据泄露和关联攻击风险增大。

（2）技术研究

针对异构终端存在的安全问题，已提出众多行之有效的技术。

终端身份认证技术主要分为基于用户账户密码、基于 MAC 地址、基于用户公开身份、基于用户持有的令牌和基于用户生物特征这五种验证机制，同时，基于区块链的身份验证技术成为未来的重要研究方向之一。

终端数据保护技术主要包含数据脱敏和数据传输两个方面。不同的场景、不同的对象、不同的阶段，在数据保护技术上都应有所差异，比如防火墙、身份匿名、数据扰乱、模糊化位置等技术。

终端入侵检测技术主要检测终端互联网络中异常通信行为。通过对网络的实时监测实现对内部攻击和外部攻击的保护，具有实时性、主动性等特点。

3.4.3 数据安全问题引发的 AI 信任危机

随着“新基建”建设加快推进，智慧工厂、智能交通、智慧城市等项目在全国如火如荼的展开，这些“智”的背后离不开人工智能技术，它早已渗透到生产、生活的方方面面。然而，人工智能极度依赖大数据平台和深度学习技术来帮助形成机器智能，因此外在的数据安全、内在的“黑箱模型”不可解释以及脆弱性漏洞，以及因为深度伪造等技术造成的低成本造假等，也让人们对 AI 技术从最初的全面肯定转变为信任危机，具体情况如图 3-6 所示。

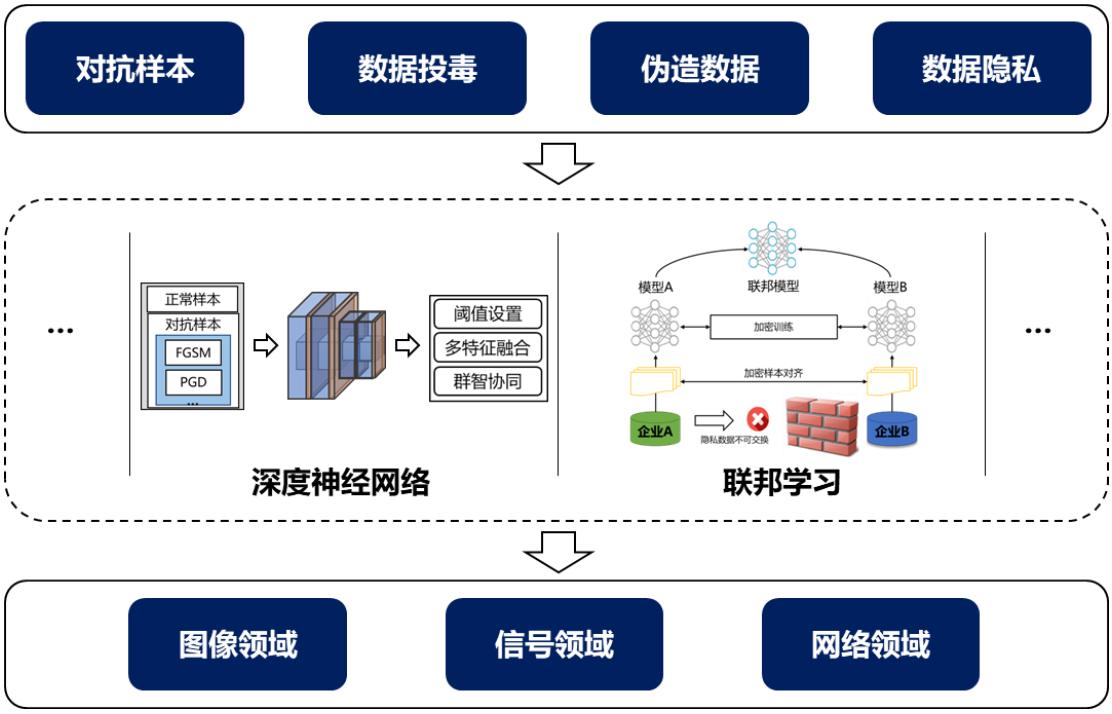

图 3-6 数据安全问题下的多领域 AI 攻防

（1）数据问题

新兴技术与现实生活中经济、生活、文化等方面的交叉融合，产生了海量数据，这些数据的运用将成为未来竞争和发展的基础。特别地，随着深度神经网络的崛起，使得基于深度学习模型的视觉识别、信号处理以及网络分析等研究领域取得了极大的进展，人工智能技术作为大数据挖掘的利器重新走到前列。然而这项造福人类的技术，却频频出现各种风险危机。基于数据的 AI 安全风险，主要有以下四类：

一是对抗样本，对输入数据以一定策略添加细微的干扰，导致 AI 模型输出错误结果，例如机器视觉中的车牌识别误导；

二是数据投毒，是指干预深度学习的训练数据，导致预测模型

输出错误，例如通过控制对话内容影响可训练问答式机器人发表敏感言论；

三是生成伪造数据，指 AI 无中生有自动生成图像、视频、音频等数据，例如深度伪造（Deepfake）技术的 AI 伪造视频、音频等；

四是隐私问题，AI 技术从数据采集、标注到数据分析和挖掘阶段，都存在着信息滥用、泄露等隐私风险，例如以获取 AI 模型隐私为目的的窃取攻击。

（2）技术研究

AI 数据安全性的方法研究方面，主要针对和数据相关的对抗攻击与防御开展。深度学习的攻防研究也最早在图像领域中开展，对抗样本和数据投毒的危害尤为突出，例如：身份认证攻击、人脸检测攻击、场景识别攻击和目标检测攻击等。在电磁空间领域，SDR 在信号攻防领域应用日趋广泛，在频谱传感应用、发射器识别、认知干扰和抗干扰等领域也取得了一系列进展，常见的信号攻击方法有：反向工程信号、重放攻击、信号干扰、GPS 欺骗等。在网络空间领域，图神经网络作为深度学习在网络分析应用中的一类重要模型，同样容易受到对抗扰动的攻击，常见攻击方法一般通过重连边、增删节点、修改节点特征等方式实现。

鉴于对抗攻击将带来重大的人身伤害和财产损失，甚至威胁到国家安全，当前针对性的防御策略也获得越来越多的关注。对抗攻击的防御措施主要沿着三个方向发展：在学习中使用改进的训练或在测试中使用修改的输入；修改网络，例如通过添加更多的层/子网络、改变丢失/激活功能等；用外部模型作为网络附件对未知的样本进行分类。

在应用领域方面，目前对抗防御方法已在图像、信号和网络等领域研究开展并应用于一些行业场景，为数据安全问题引发的 AI 信任危机提供更多的可行方案，对抗环境下数据安全和隐私保护，依然是提升 AI 可信的严峻挑战之一。

3.4.4 应用场景多元化伴随着风险多元化

大量业务随着数字化改革迁移到数字基础设施中，5G、人工智能、云计算、区块链等技术的普及将促进新兴业务的快速发展，深刻影响人们的办公、家居、出行等。然而，“新基建”技术的多元化应用，也使得技术本身的安全性问题渗透到具体的应用场景。区块链金融安全问题、智能交通运输安全问题、工业互联网安全问题、智慧城市安全问题等等也将影响到整个数字经济的正常运行，波及到政府、企业和个人。

（1）工业互联网

工业互联网作为国家“新基建”战略的重点领域之一，是新一代信息技术与工业经济深度融合的新型基础设施、应用模式和工业生态。安全作为工业互联网的发展的前提和保障，事关经济发展、社会稳定和国家安全。当前工业互联网主要面临以下问题：

● 法规政策响应度低。近些年已构建安全管理体系，并完善了工业互联网的安全顶层设计，但各领域在实施和相应上较为缓慢。

● 新旧共生系统较脆弱。因数字化程度不一，众多领域、行业技术在实现互联互通的同时，新旧系统共生的情况增加了工业互联网的脆弱性。

● 数据安全问题严重。工业互联网所采集数据覆盖面广、体量大、结构复杂，却又缺乏严格加密，使得数据隐私问题一直是攻克

的重点和难点。

● 外界攻击方式繁杂多样。工业互联网集成大量设备，进行通信与控制等行为，但也为攻击者提供大量的控制端点，实现多样的攻击方式。

（2）智慧城市

智慧城市是新型基础设施与传统基础设施融合发展的重要领域。新型智慧城市建设能推进产业升级，激发经济发展内生动力，提升城市的综合承载能力，为城市新发展格局提供强大支撑。新冠疫情加速了行业数字化转型，也带来了智慧城市规划和建设的深刻反思，让人们意识到新型智慧城市建设在支撑城市健康高效运行和突发事件快速智能响应方面发挥的重要作用。由于智慧城市系统复杂、参与方多、演进较快等特征，智慧城市建设过程仍存在诸多问题和挑战：

● 基础设施建设缺乏规划性。智慧城市建设涉及交通、医疗、电网等诸多领域，设施应适应技术的发展，技术应可持续发展和长期运营迭代。

● 信息安全顶层设计薄弱。针对信息产品和系统漏洞的信息安全犯罪屡见不鲜，同时，公民安全意识薄弱。可见信息安全治理及政策的不足。

● 统筹协调力度不强。智慧城市建设的各部门间缺乏有规划的合作交流，信息不对等、协调滞后，缺乏完整的组织管理系统及产业链。

● 标准体系设计不统一。概念体系、设施技术、行业应用的不断发展导致标准内容交叉重复，难以形成统一，因此，标准体系设

计要满足跨行业、跨产业、跨组织的有机协同的标准需求。

（3）区块链

区块链为“新基建”构建信任基石。区块链作为一种底层的分布式数据存储技术，其本身具有数据公开透明、开放性、独立性、去中心化等特性，可用来解决可信身份、数据篡改等问题，对金融、物联网、公共服务等众多领域的安全稳定有重要保障作用，但仍存在下列风险：

● 缺乏第三方保护。利益冲突的逐渐出现表明需要规范法律和社会问题，形成对社会秩序的冲击，第三方介入能实现利益保护、权责追溯等作用。

● 核心技术及机制不完善。智能合约本是一种旨在传播、验证和执行的协议，却被恶意利用而进行注入、交易回滚等攻击手段。

● 产业生态中存安全问题众多。主要攻击分：基于对等网络，比如日蚀攻击、女巫攻击等；基于共识和挖矿，比如算力攻击、时间劫持等。

● 用户自身潜在安全风险大。例如私钥被盗、病毒侵入、用户交互风险等，与之相对应的是技术风险的普及和用户风险意识的培养。

（4）治理策略

产学研用各方需凝聚共识，应自顶向下形成政策支持-体系完善-技术保障的新兴技术安全发展局面，共同为多元场景下的安全保障体系建设贡献力量。

加强顶层管理，强化政策制定与部署落实。构建规范有序的数字化治理体系，建立政府监管、企业履责、行业自律、社会监督的

安全多元共治局面，实现政府、行业、个人间的有机互动与有效结合，比如《智慧城市标准化白皮书》的编制为政府、企事业单位等智慧城市相关方的标准化工作提供指导。

完善法治体系，维护新兴技术的健康环境。针对技术融合发展中发现的问题，及时研制有效的应对方案，并健全监管治理的法律体系，实现技术的健康发展，比如面向区块链技术、平台、应用生态，研制相对应的安全技术要求、安全标准。

增强技术保障，促进新兴安全技术发展。进一步加强监管与治理方面的技术探索与应用。同时，发挥安全技术在新技术领域的支撑保障作用，强化对区块链、智慧城市等行业的安全风险检测和应对技术，打造安全和发展并重的技术体系。

3.4.5 小结

“新基建”加快推进数字社会的发展，各类新技术新业务新应用不断涌现，国家、政府、企业和个人都将面临更加复杂多面的虚拟和现实世界。在关注“新基建”带来的新动能和新机遇，享受“新基建”带来的社会红利的同时，更应该警惕数字化高速发展背后的隐患和风险。通过提前对这些风险做出应对方案，“新基建”将更好地赋能数字社会发展。

3.5 数字平台生态失控风险与新型反垄断制度设计

作为我国经济发展的重要动能，数字平台保持着高速发展的势头，并且已经引起了政府的高度重视。《平台经济与竞争政策观察（2021 年）》中的数据显示，从 2015 年到 2020 年，我国市值在 10 亿美元以上的数字平台的市场价值以高达 35.4%的复合增长率增加。一方面，数字平台的迅速发展从多个领域对社会经济产生了影响；

另一方面，数字平台独特的竞争方式也对传统市场的运行秩序形成了挑战,从而引起了政府对于平台规范发展的重视。从2021年开始，为营造良好的数字平台生态，各部门陆续发布《关于平台经济领域的反垄断指南》《互联网平台分类分级指南（征求意见稿)》《互联网平台落实主体责任指南（征求意见稿)》《关于推动平台经济规范健康持续发展的若干意见》等文件，逐渐完善数字平台治理体系。

要坚持推动平台经济规范健康持续发展的一个重要任务，是明确当前数字平台生态可能存在的各种风险，并进一步采取相应措施来加以防范。在多种风险中，数字平台垄断风险是一个至关重要的方面，在当前反垄断实践的基础上进一步构建出有针对性的、更为合理的反垄断制度具有重要意义。

3.5.1 数字平台生态失控风险

随着平台经济的不断发展，数字平台逐渐在经济社会的多个领域扮演了重要角色，加剧了数字平台失控风险可能带来的不良影响。

（1）内部失控风险

平台内部失控风险包括两类，一部分风险在传统经济中就已经存在，但会被数字平台进一步放大；另一部分是由数字平台带来的新风险。

在传统经济中就已经存在、并且由数字平台进一步放大的风险，主要包括侵权假冒、虚假宣传、虚假促销、传播违法信息和虚假广告等。其中部分风险在网络交易类平台中有明确体现。根据国家知识产权局与中国消费者协会的相关数据，在侵权假冒方面，2017 年电子商务领域专利执法办案量为 19835 件；在虚假宣传方面，2017 年的网络购物虚假宣传投诉有 3546 件；在虚假促销方面，2017 年有

6.10%的商品在“双 11”进行了虚假促销。其他风险则在信息内容类平台中更加明显。根据《互联网平台治理研究报告（2019 年）》的数据，在违法信息方面，2017 年的违法和不良信息有效举报受理量为 5263.9 万件；在虚假广告方面，2017 年的虚假违法互联网广告查处量为 14904 件；在版权侵权方面，2017 年的“剑网”专项行动共关闭侵权盗版网站 2554 个；在内容低俗方面，2017 年的“净网”行动处置淫秽色情等有害信息 455 万件。上述传统风险虽不由数字平台直接产生，但是在数字平台的作用下能够突破空间限制实现广泛传播，对社会产生了更加严重的影响。

此外，由于数字平台具有与传统经济不同的运行模式，数字平台中还可能出现新的内部冲突风险，具体包括平台与内部经营者的冲突风险，以及平台与劳动者的冲突风险。

平台与内部经营者的冲突风险主要表现为数字平台向平台内经营者收取了过高的费用，以及自我优待问题。滴滴司机的“高抽佣”是平台向平台内经营者收取高费用的典型案例。滴滴披露的 2020 年的数据显示，当滴滴对司机和乘客端都不进行补贴时，佣金率高达 25.6%；当滴滴对司机和乘客端都进行补贴时，佣金率也有 11.2%。滴滴的“高抽佣”成为了其与内部经营者冲突风险的具体形式，也成为了滴滴司机数量近年来明显降低的一个重要原因。平台自我优待的典型表现是数字平台利用其平台服务提供者的身份，在自家与其他的服务提供者的竞争中，通过特定的商业手段，给予自家的产品或服务优惠待遇。《数字市场竞争状况调查报告》指出，自我优待主要有如下表现形式，一是搜索引擎在搜索时将自家的搜索结果放置于更显眼、更有利的位置，或将竞争者从索引中移除；二是应用

商店扮演安装软件时的“守门人”，优待自家的应用程序；三是减少竞争对手对于自家服务的访问权限，从而降低用户体验，实现自我优待。

平台与平台劳动者的冲突风险主要表现在针对零工经济劳动者的保护不足，其中外卖数字平台就是典型案例之一。北京大学博士后陈龙的田野调查显示，外卖数字平台在运行过程中不断利用多种方法来试探骑手极限，以获取更高收益，具体包括不断追踪记录骑手的相关数据，并以此为基础规划取餐、送餐以及定价安排，从而实现对骑手的全面控制，以降低支付给骑手的费用。此外，外卖数字平台还会通过系统调整的方式来试探骑手的送餐速度，并进一步压缩配送时间。外卖数字平台利用数据和算法不断压榨骑手劳动，而没有有效保护零工经济劳动者的利益，体现出了平台与平台劳动者的冲突风险。

（2）垄断风险

根据《中华人民共和国反垄断法（2022 修正）》对于垄断行为的界定，结合数字平台的相应特点，可以判断数字平台的垄断风险主要集中体现在经营者集中风险、滥用市场支配地位风险和垄断协议风险三个方面。

杀手并购是数字平台进行经营者集中的一个重要现象，指大型数字平台为了减少竞争和维持市场优势地位而收购新兴平台。杀手并购存在两种形式，一是大型数字平台并购规模小但有巨大潜力的创新型初创企业，以防止其产生威胁；二是大型数字平台并购初创企业后，迅速将其市场份额做大，从而使该行业的其他初创企业消亡。杀手并购的对象规模小、相关市场界定较为复杂的特点往往为

监管带来了困难。

滥用市场支配地位主要包括“二选一”和“大数据杀熟”两种风险。“二选一”主要表现为部分数字平台禁止平台内经营者在其他竞争性平台开店或参加其他竞争性平台的促销活动，并辅以多种奖惩措施来保障“二选一”顺利实施。数字平台强制“二选一”行为限制了平台内经营者的经营自主权，损害了平台内经营者的利益；同时也限制了平台经济的创新发展，损害了消费者利益。“大数据杀熟”具体有四种形式，一是对新老用户制定不同价格，会员用户反而比普通用户价格更贵；二是对不同地区的消费者制定不同价格；三是提高多次浏览相关页面的用户面临的价格；四是利用复杂的促销规则和算法，吸引计算真实价格困难的消费者。“大数据杀熟”往往具有隐蔽性强、方便区分用户和诱导消费等特点。

数字平台的垄断协议风险主要是指平台的“算法共谋”，即数字平台利用数据和算法实现多方协同，从而对市场竞争产生影响。一般而言，“算法共谋”可以分为四类，一是“信使类共谋”，各个数字平台在达成协议价格之后设计定价算法，保证共谋者背离协议价格的行为会受到严厉惩罚，从而来顺利实施共谋；二是“轴辅类共谋”，数字平台与多个经营者实现纵向共谋之后，利用平行式算法来自动决策多个有竞争关系的经营者之间的定价策略，以实现经营者之间的横向共谋；三是“代理类共谋”，数字平台利用算法来持续监控竞争者的定价，并根据市场数据来自动调整自身定价，从而实现价格共识；四是“自主类共谋”，算法能够通过深度自我学习来对市场主体的行为做出预判并形成最优定价策略，从而在不需要人类干预的情况下达成共谋。

（3）数据风险

数据是数字经济时代的重要生产要素，是数字平台提高运行效率的最重要资源之一。然而，在数据周期的每个阶段都存在隐私泄露和危及数据安全的可能。导致数据风险的主要原因包括数据产权不明晰、数据隐私保护不完善及数据垄断等。

数据要素的确权是数据收集、使用和共享的前提。当前数据产权界定的困难主要在于数据的复杂性和非排他性。数据包括原始数据、衍生数据和数据产品等多种形式，个人、企业和政府三个主体在不同数据中扮演的角色各不相同，因此数据同时会存在多主体使用和边界模糊的特点，具有复杂性。此外，由于同样的数据能够方便地被多个经济主体使用，具有非排他性，因此数据不能沿用传统有形资产产权立法，需要特别确立产权与交易规则来更好地实现经济效益。

数字平台的隐私保护存在两个问题，一是个人数据的收集与隐私保护存在冲突。数字平台需要大量个人数据才能为消费者提供更好的服务，而大量数据的收集又伴随着数据泄露和滥用风险。二是隐私保护的数据与商业情境下的大数据所关注的对象不等同。个人信息保护制度构建是以单一数据为对象的，在很大程度上能够保障单一数据安全，而个人数据的价值释放是以大数据为基础的，部分没有被隐私保护关注的数据被泄露和非法利用也仍然会影响个人隐私。

数据垄断指数字平台利用独占数据进行了垄断行为，可能具有如下危害，一是限制市场自由竞争，扭曲资源配置，降低市场效率；二是可能侵害消费者个人隐私安全；三是利用大规模数据排除潜在

竞争者，形成市场进入壁垒，降低了市场创新和研发活力。

（4）资本无序扩张风险

平台的资本无序扩张风险主要表现为进入金融领域的数字平台在追求规模扩张时处于的“无序”状态，主要包括如下五个方面。

一是垄断和不公平竞争。大型数字平台借助技术、信息、数据，以及客户资源优势，能够在竞争中占领优势，从而形成市场主导地位，并可能强化为市场垄断。

二是产品和业务边界模糊。一方面，开展金融业务的数字平台可能利用科技公司的身份来逃避金融服务需要的准入要求和业务监管；另一方面，数字平台同时提供的多个种类的产品服务打破了传统框架下的产品和业务边界，进一步加大了风险。

三是数字平台的相关信息技术在可控性和稳定性方面还存在着风险。具体包括监管机构难以识别并处置前沿信息技术中的隐含风险，数字平台的信用评估和风险控制中的数据类型不够完整、规模不够大、质量不够高、噪音难消除，数字平台征信的资源整合与共享机制不完善等。

四是数字平台存在欺诈消费者和数据泄露的风险。一方面是数字平台与金融消费者存在信息不对称和利益不一致的冲突，从而可能导致数字平台存在虚假操作和欺诈行为的风险；另一方面金融行业的数据信息丰富，使用金融科技的数字平台可能会由于自身出现技术故障、遭受黑客攻击和员工盗窃等原因泄露大量用户数据，成为信息泄露的重灾区。

五是数字平台可能导致金融体系的系统性风险。其一，大型数字平台进入金融领域后，可能会逐渐变得“大而不能倒”；其二，数

字平台服务对象宽泛，包括了传统金融机构无法覆盖的长尾人群，从而可能在市场波动时造成系统性风险；其三，进入金融领域的数字平台混业经营并提供交叉性金融产品，其顺周期性显著，形成了更具有隐蔽性和破坏性的风险。

（5）在非经济领域的风险

数字平台生态失控风险不仅局限于经济领域，还进一步扩展到了非经济领域。具体而言，数字平台在非经济领域的风险主要体现在如下四个方面。

数字鸿沟指由于不同人群在数字化进程中，对于数字技术的占有和应用程度不同而造成的落差。数字沟通能力与社会阶层、教育程度、工作状况等因素息息相关，处于社会分层顶层的群体往往能够更容易地通过数字技术跨入数字社会。随着数字技术应用范围不断扩张，数字鸿沟具有加剧社会分化的风险。

“信息茧房”风险指由算法主导下的数字内容分发模式放大和加强了被标签化的信息输出和推送，从而引发个体“自我封闭”危险。在算法的作用下，人们接受到的内容都是自己希望看到和听到的，从而在长期的重复和自我证实中得到自身观点的“茧房”，难以接受不同的信息和观点。

“后真相”指一种非常态化的舆论生态，与事实决定真相的基本规律不同，“后真相”在“先有观点再有事实，让事实为观点服务”的逻辑下，通过碎片化的事实来引导大众情绪和舆论走向，让受众难以将关注点聚焦在事实和真相上。

最后，数字平台，尤其是内容类平台，在社会价值导向中起着重要作用，从而也具有错误引导消费者的潜在风险。“饭圈”文化就

是典型代表之一。央视网评曾描述道，在“饭圈”的生态系统中，部分平台资本构建了复杂而又专业的“游戏规则”，刺激粉丝的竞争心理并诱导粉丝形成群体来参与冲榜与刷量，将“饭圈”变为了一种商业型平台。此外，“饭圈”文化还在不断发酵中逐渐超出经济领域，延伸到了价值观领域，甚至出现了部分无底线追星行为，具有严重的不良社会价值导向。可见，如果数字平台没有尽到相应的管理责任，甚至还错误地引导用户群体，将会带来社会风险。

（6）平台治理不规范的风险

对数字平台实施不规范的治理，可能会对数字平台发展产生严重的负面影响。其一，部分旨在规范数字平台的政策对市场准入设置了不必要的障碍，涉嫌行政垄断，违反公平竞争审查制度，不利于全国统一大市场的形成。其二，在规范数字平台时具有以运动式行政手段代替规范法制化手段的风险，从而导致不同行政部门出台的制度与指导意见存在冲突，并进一步使得数字平台在发展的过程中无法明确发展规则和方向。其三，在实施执法后缺乏后续妥善处置，在短期内为数字平台带来明显负面影响，针对数字平台的每一次反垄断措施都可能对市场情绪产生重要引导作用，如果没有妥善处置后续影响，将不利于数字平台长期健康发展。

3.5.2 新型反垄断制度设计

数字平台生态的一系列失控风险，对相关的防范和控制措施提出了新的要求。对于数字平台生态的内部失控风险，需要强化平台内部治理，并积极落实《关于维护新就业形态劳动者劳动保障权益的指导意见》等法律法规，以维护平台用户权益；规避数据风险则对数据合理确权、隐私有效保护和破除数据垄断提出了相关要求；

克服资本无序扩张风险，可以从树立合理的市场准入规则并注重有关政策的动态调整与落实入手；为了防范非经济领域的风险，数字平台应当落实主体责任，倡导良好价值导向；在平台治理不规范方面，要推动数字平台管理的法制化，并警惕行政权力的滥用。

在多种数字平台生态失控风险中，垄断是一个至关重要的方面。反垄断机构应当充分理解数字经济时代下的竞争特点，从宏观内涵到微观执法的各个层次，对于数字平台的反垄断有较之于传统经济不同的认识，并构建出适用于数字平台的反垄断制度。

（1）具体辨析“垄断”内涵，重视长期效率

数字平台呈现出“动态竞争”的特性。在数字平台之间的竞争中，产品服务以及市场上的主导企业更迭速度较快，传统观念里象征着垄断的高市场份额和高市场集中度都可能是不可持续的，进一步的，持续的高市场份额和高市场集中度现象也有可能是有效竞争的结果。

因此，数字平台的反垄断不能根据平台的指标表象武断地认为“凡是垄断必须干预”，而应该综合网络效应强度和其他经济特征来考虑。事实上，反垄断应当以提高竞争效率或长期社会福利为最终目标，在进行反垄断执法时明确数字平台运行机制，并辨明平台行为的实际效果，避免对于单一结构性指标的不当依赖。

（2）构建动态反垄断体系

数字平台的“动态竞争”特性，要求反垄断机构进一步构建动态反垄断体系。具体而言，动态反垄断体系包括动态效率“分析框架”、动态指标“工具箱”和动态执法“政策库”三个部分。

在建立动态效率“分析框架”方面的一个有益尝试是，在通用

的静态分析框架的基础上阐述数字平台竞争的动态演化过程。梳理出数字平台获取盈利机会所需的重要资源，并从资源入手来分析数字平台的垄断效果，一是考虑平台自身利用关键资源的能力，分析平台是否存在垄断动机；二是考察关键资源的可获得性，以理解平台是否能够在事实上造成垄断后果。

建立动态指标“工具箱”首先需要明确的是，通用的反垄断工具虽然不能直接被套用于数字平台，但是仍然具有初步圈定市场边界的大致范围，并排除明显错误的作用，不应全盘摒弃。在动态指标的设计上，在实践中比较合理的方法是，在通用的反垄断工具的基础上增加部分动态性指标，将数字平台绩效的刻画从单一产业维度扩展至更加全面的维度，以确保在执法时具有足够的灵活性。此外，反垄断执法不应只将数字平台行为或政府干预行为的短期影响作为考虑目标，而需要在一个足够长的时间跨度内进行综合考量。

建立动态执法“政策库”要求适当突破惯常的执法思路，除去采用高额罚款、禁止各种潜在反竞争行为和抑制特定的集中化市场结构等手段以外，还需要辅以平台主动合规、教育培训、社会监督等措施，进一步增加有助于促进竞争的在位企业和潜在进入者的数量。

（3）建立内外协同的反垄断格局

由于数字平台提供的产品和服务大多数都牵涉了众多相关利益主体，具有一定的复杂性，因此对于数字平台的反垄断监管还需要各方利益主体的通力合作，共同构建内外协同的反垄断格局。

内部治理包括平台自治、消费者反馈与商家举证。一是引导平台自主合规，通过引导数字平台相互监督和行业自查、互查的方式

来加强平台自律，提高平台治理能力；同时建立和完善围绕正当理由展开的抗辩制度，在观点碰撞中实现反垄断监管的最优化。二是完善平台用户的反馈评价功能，疏通消费者投诉举报机制，完善消费者公益诉讼制度，为消费者维权提供有力支持；同时政府和第三方信用机构要辅以相关的备案、处理、反馈与公示工作，及时处理消费者集中反映的问题。三是为商家提供有效的申诉渠道，鼓励商家在反垄断调查中积极举证，明晰抗辩制度的具体程序、监管形式、救济方式以及举证责任，打通平台内经营者的发声渠道。

外部监督包括政府执法、社会组织协助与公众监督。一是强化政府执法，明确反垄断执法机构与其他监管机构的目标分工，让反垄断目标回归经济本身，并将存在冲突的不同政策分别委托于相对独立的执法部门，并配以与多任务相匹配的激励结构；二是加快第三方信用机构建设，统一平台信用技术标识和认证流程，发布信用标识认证，使用户能够快速准确判断数字平台的信用状况；三是重视公众监督，反垄断执法从接到举报开始就需要在每个节点上尽量保持公开透明，并接受社会监督。

（4）倡导反垄断执法“类型化”

鉴于数字平台和相关业态具有的独有特征，应当慎用设定行为不是被允许就是被禁止的本身违法原则，进一步强调“类型化”。一方面，“类型化”有助于增强数字平台反垄断执法的“精准性”。反垄断工作量大面广的特点意味着必须突出重点、分清主次，要将过往存在高风险、近期受到违规举报较多、预期未来经营困难的数字平台，以及关于民生安定、社会稳定、国际国内竞争力的重点行业放在重要位置，将综合研判与专题分析相结合。另一方面，“类型化”

有助于提升数字平台反垄断的“可操作性”，数字平台反垄断越能分类细化，就越有利于精准施策，应当重视“一事一议”，用“一把钥匙开一把锁”代替“一刀切”。

数字平台反垄断的“类型化”需要更加广泛的法律主体共同参与。一是分专题深入梳理《中华人民共和国反垄断法》《中华人民共和国反不正当竞争法》《中华人民共和国电子商务法》等相关法律与平台治理的契合点，并根据平台特征作出相应释法。二是对于与特定类型存在冲突的法律及时修订，避免“政出多门”。三是坚持开门立法的基本原则，一方面广泛征求有关部门、各类市场主体和专家学者意见；另一方面与国际主要司法辖区反垄断执法机构围绕重点难点问题深入交流，充分借鉴反垄断领域前沿理论研究成果。

3.5.3 小结

作为迅速发展的一种新兴经济形式，数字平台在深入影响社会各个领域、为生活带来巨大便利和剧烈变革的同时，还不可避免地存在一系列失控风险。并且随着数字平台在社会范围内影响力的扩大，数字平台失控风险也逐渐覆盖了平台内部治理、平台垄断、数据、资本无序扩张，以及非经济领域等多个方面。要实现数字经济的规范健康发展，仍需要对伴随着数字平台发展而产生的相关风险有进一步认识，并针对各类风险提出有针对性和现实意义的制度建议。

3.6 数字社会数据隐私保护与隐私技术发展

在数字经济时代，数据已成为一种新的生产要素。我们每个人每天都在产生新的数据，同时我们的私人数据也在不经意间被泄露和利用。数据改善着人们的生活水平，使工作变得容易，但同时也

对个人隐私造成极大破坏，高度信息化使得私密信息泄露风险增加。在国务院印发的《“十四五”数字经济发展规划》中明确提出，需加快构建数据要素市场规则，培育市场主体，完善治理体系，促进数据要素市场流通。此外，人工智能技术的突飞猛进对数据所有权和用户隐私保护也提出了新的挑战。因此，研究数据保护、确权、定价、交易和公平性机制，不仅对拥有数据的主体意义重大，可以保障其在数字化时代的基本权利，而且对整个社会的经济发展和公平分配也是基础性的。

加强个人数据隐私保护，防范数字化权力风险，开创数据“可用不可见”的新局面，是加快构建数据要素市场规则的迫切需求和重要内容，从隐私保护技术研发的角度来看，该领域研究问题众多。

3.6.1 分布式隐私计算与建模

在大数据时代中，数据价值的体现必然依赖于数据之间的相互流通，但数据的流通又必然导致个人隐私安全的破坏，所以如何在实现数据流动的同时，有效防止敏感信息泄露，保护数据隐私安全是当前大数据应用技术发展的一个重要研究问题，隐私计算便由此产生。

隐私保护技术通俗来说即是实现数据的“可用不可见”，数据可用性即开放性，指的是充分利用各种数据，让数据对外开放，服务于决策；数据不可见性即不共享性，指的是数据不离开机构（如政府、互联网企业、运营商等）或个人，保证数据不对外直接共享。要实现隐私保护的以上要求，就必须对数据进行加密，且在数据被访问时，采取技术手段防止数据中敏感信息被访问者以某些方式“逆向”获取，从而造成用户敏感信息被泄露和滥用。在数据密集型计

算范式时代，如何妥善、安全地获取和使用数据成为迫切需要解决的问题。然而，数据自身具有分散性和非排他性。不同于一般实物，数据可以同时或非同时地为多个主体所使用，且只有在使用中才会产生价值，因而数据也很难做到中心化管理，谁都可能获得和使用数据。此时，分布式计算和建模成为数据隐私保护的更好选择，在这一范式下可依赖的关键技术包括联邦学习（federated learning）、安全多方计算（secure multi-party computation，SMPC）、区块链（blockchain）和差分隐私（differential privacy）。

（1）联邦学习

在现实生活中，除了政府和极少数大型互联网企业能够拥有海量优质的客户数据，绝大多数公司都面临数据量少、数据质量不高的问题，缺少支撑人工智能技术的基本前提条件，这些分散的数据往往会形成孤岛。联邦学习的产生便是为了解决这些数据孤岛问题。

联邦学习是一种新型的机器学习设定，其目的是在保证数据隐私安全，以及合法合规的前提下，实现各方共同建模，将模型训练的过程由中心转移到各个数据拥有者手中，而不需要集中收集数据。在联邦学习中，许多客户端可以在一个中央服务器的协调下共同训练模型，在保证各客户端节点独立训练模型的同时，又能实现不同节点之间的数据共享。

近年来，由于各国法律法规的约束，我们已经无法像以前一样直接粗暴地收集客户数据，然后用以完成机器学习任务。根据现行法律法规对用户个人隐私的保护，大多用户数据都必须保留在用户本地，虽然这些举措有效地保护了个人隐私，但同时也不方便实现数据交换和整合，大大制约了机器学习能力的进一步提高，因此隐

私安全的保护和人工智能模型能力的提升两者之间形成了矛盾。联邦学习便是人们在此情形下探索出的一种机器学习新模式，其可以在不交换本地原始数据的前提下，仅通过模型参数或中间结果的传递来实现全局模型的构建，从而很好地解决了隐私保护和数据共享之间的矛盾。也可以说，联邦学习是一种“数据可用不可见”“数据不动模型动”的应用新范式。

（2）安全多方计算

安全多方计算是指在无可信第三方参与的情况下，拥有数据的多方在确保数据不泄露的同时，利用隐私数据参与保密计算，共同得到的一个计算结果。安全多方计算主要利用到下面四个技术。

①不经意传输：不经意传输协议是一种可保护隐私的双方通信协议，用以保护信息发送者和接收者的隐私。信息发送者从一些待发送的消息中发送一部分给接收者，但不知道发送了哪些信息（对接收者的隐私性）；同时，接收者也只能获得那一部分信息，而无法获取其他的任何信息（对发送者的隐私性）。

②秘密共享：在秘密共享系统中，秘密被参与者群体合理分割，只有多于特定个数的参与者合作，才可以恢复或计算出秘密，参与者个数少于特定值则无法获取秘密。攻击者想要获取密钥就必须同时获得一定数量的秘密碎片，这样就能提高系统的安全性；此外，当某些秘密碎片遗失或者损坏时利用其他参与者掌握的信息依然可以获得秘密，提高了系统的可靠性。

③混淆电路：混淆电路可以用于解决安全计算问题，其核心技术是将两方参与的安全计算函数编译成布尔电路形式，然后将真值表加密打乱，从而在不泄露参与者信息的基础上实现电路的正常输

出。相比较于其他安全计算技术，混淆电路具有更高的通用性，因此发展空间很大。

④零知识证明：零知识证明也是安全多方计算的一种常用的技术手段。零知识证明指的是示证者在向验证者证明某项问题时，在不暴露任何有用相关信息的前提下，使验证者相信某个论断是正确的。所以，如果将零知识证明成功的应用于实际，那么就可以很好地保护隐私安全。

（3）区块链

区块链是一种去中心化、公开透明的防篡改账本。在中心化情况下，一些企业为了逃避法律责任，往往会篡改数据或者直接删除对自己不利的数据。传统的机械技术（例如硬盘数据恢复、日志查询、IP 追踪等）没有从根本上解决问题，而区块链的出现使得数据具有了不可篡改性，是一种重大突破。同时，区块链技术具有高可靠和高可用性，数据被分布式存储，冗余备份，任何单个节点的崩溃都不会导致整体数据的丢失。

智能合约是基于区块链技术的一种计算机协议，是一个在可信的执行环境下，由计算机语言取代法律语言记录条款，并由程序自动执行的合约。简单来说，智能合约就是传统文本合约的数字化形式，并且可以在脱离人为监控的情况下由程序自动执行。与传统合约相比，智能合约具有三个特点：①开放性。智能合约完全部署在区块链上，所以合约的内容自然是公开透明的。②安全性。由于区块链的特性，智能合约的内容无法被少数人修改。③永久运行。只要区块链存在，智能合约就可以被所有的网络节点共同维护，因而可以一直运行下去。

在隐私计算中，我们可以通过区块链记录所有的数字交互过程，并通过智能合约记录数据建模和分析的过程，实现计算过程的安全可信。

（4）差分隐私

差分隐私是针对数据库的隐私泄露问题提出的一种新的隐私定义，是为了解决差分攻击而引入的一种解决隐私保护模型。其原理是在原始的查询结果中添加干扰数据，再将结果返回。差分隐私可以在最大化实现数据查询准确性的前提下，最大限度地减少识别其记录的机会；即在保留统计学特征的前提下，去除个体特征以保护用户隐私。我们可以通过对目标函数、梯度和输出结果添加噪声，实现差分隐私和机器学习的结合。加入干扰后，用户便无法通过查询结果反推出准确的信息，从而达到保护隐私的目的。

3.6.2 基于现代产权理论的数据确权

数据确权即是数据产权的确定，其目的是保护数据权利人对数据财产的直接控制和支配的权利，本质是在大数据时代中，数据生产者对于社会资源分配的一种主张。2020 年 4 月 9 日，中共中央、国务院发布《关于构建更加完善的要素市场化配置体制机制的意见》，首次将数据与土地、劳动力、资本、技术等传统要素并列为生产要素之一，提出要加快培育数据要素市场，包括推进政府数据开发共享、提升社会数据资源价值，以及加强数据资源整合和安全保护三方面工作，明确了数据作为一种新的生产要素的地位。

既然数据已经被定义为新的生产要素，那么就必须要深入研究数据确权的机制。清晰的所有权归属是数据资产交易的前提与基础。数据资产的权利主要包括所有权、使用权和收益权等，其中所有权

是核心。一直以来，各国均在探索依靠法律通过“赋权-维权”的传统模式为数据产权保护提供依据，并取得了一定实际成果，例如欧盟最早颁布的《通用数据保护条例》（GDPR）、美国加州的 CCPA，以及我国通过的《中华人民共和国个人信息保护法》等。然而现有的法律规定在实际应用过程中仍面临掣肘，其中，GDPR 被指出可能会从根本上改变大数据分析的方式，使其成为次优且低效的保护方式；同时在国内也面临当事人提起诉讼的案件稀少而且胜诉率极低的问题，数据所有权保护并未随着立法的快速推进而达到预期效果。

在现代产权理论启发下，可以将数据确权的目标等价于最大化数据要素产生的价值。基本思想是合作中形成的产权应归属于对合作后产出贡献最大的一方。与交易成本理论相比，这一理论为垂直整合提供了新思路，回答了整合过程中的“由谁整合”的顺序问题。在数据要素整合过程中，数据要素的产权或者在用户协议之外的剩余控制权，应向起到关键作用的平台倾斜，以此激励平台投入更多资源促进数据市场高效运转。而作为贡献更显著的平台方，也即整合用户数据的一方，只须在交易过程中时向用户支付“赔偿”，即可实现现代产权理论背景下的数据产权交易。

3.6.3 数据定价和公平性

建立数据要素市场的另外一个难点就是数据资产的定价问题。数据的定价，尤其是消费数据的定价，是一个关乎未来数据市场是否公平的关键性问题。人们作为数据的生产者，不停地生产着数据，如果其他利用这些数据赚钱的人可以不支付成本，那么就会导致严重的社会不公平现象。在未来，数据就如同一种原材料，因此，对

其进行合理定价是很必要的。

“没有交易成本的世界，就像没有摩擦力的物理世界一样奇怪”，在借助机器学习对数据实现定价策略时，各方需要协同完成一个学习任务，因此基于合作博弈的定价策略才能科学地解决问题。一般来说，数据集（或机器学习中的节点）可以对应合作博弈中的参与者，机器学习训练产生的模型可以认为是合作收益，那么计算每个数据集的贡献，就转换成了合理分配利益的问题。我们可以引入夏普利值（Shapley value）来量化各数据集提供的贡献值。

夏普利值原本是解决博弈论中分配问题的一种方法，现在也可以用来解释机器学习中各特征对结果的贡献度。因具备对称性（合作者的顺序编号不影响合作获利的分配）、有效性（各合作方获利总和等于合作获利）、冗员性（无贡献的成员不参与最终获利的分配）和可加性等优良性质，天然地符合解决实际问题的要求，因而在数据定价中被广泛应用。例如联邦学习中评估每个参与者贡献的问题，可以等价于求解合作博弈中各个参与者的夏普利值问题。基于以上定价规则，可进一步构建数据交易规则。

由于大数据的生产者往往是大众，其产生的经济效益理应被普通大众分享，而非完全被平台公司私有化，否则就会再现“遍身罗绮者，不是养蚕人”的悲剧。除了上述的联邦学习+博弈定价之外，还可以采用数据银行、数据信托和数据 B2G（business to government）等方法对数据收益进行公平分配。

3.6.4 小结

我们正处于新工业革命背景下数字经济发展的大数据时代，数据已经和劳动力、土地、资本和信息等并列为一种新的生产要素，

人类社会、物理世界和信息空间深入融合所形成的三元空间，以前所未有的广度和深度映照人们工作、生活和生产的规律和模式，数据作为重要的经济社会发展资源的价值愈发得到凸显，不断地促进社会飞速发展，并逐渐成为一个社会的核心资产。但同时，大量数据资源的挖掘也成为了一把“双刃剑”，高度信息化对个体隐私安全造成前所未有的破坏，因此，如何平衡数据市场化和数据隐私安全的矛盾成为亟待解决的社会问题。为了解决这个问题，需要我们同时做好隐私技术的研发和数据权力的保护。其中，隐私技术作为保护个人隐私的重要手段，必须不断深化理论研究，优化算法和模型；而数据权力的保护作为数据要素市场化中的核心逻辑，其发展需要我们对数据确权和隐私保护、数据定价和交易、数据收益公平分配等基础问题投入更多的资源和人力，做长期深入的研究。

第 4 章 数字社会风险的智能治理

4.1 数字时代的技术差异赋权及其风险治理

互联网、人工智能和大数据等技术在改变人们生产生活方式、推进数字政府治理与改革、提升企业生产经营效率的同时，也因国家、企业和个人获取，以及运用数据信息的水平能力差异，引发数字技术对不同主体的差异赋能和参差赋权。在浩荡的数字经济大潮中，数字技术被全面、深入、广泛地运用，它在改变社会主体活动能力、方式和机遇的同时，引发深刻的社会结构与社会关系变革。这诱使数字技术的非均衡赋权超越资本运作和政权治理规律，不断加剧不同社会主体间权利/权力的结构性失衡。此时，一种“技术+法律”的分析与治理模式就成为必然选择。

4.1.1 数字技术赋权及其差异化样态

（1）数字技术赋权的特性

第一，从赋权对象看，数字技术赋权针对国家、组织和社会多主体，并具有交互性。相对于传统赋权主要面向社会弱势群体，目的在于扩大边缘群体进行平等资源分配的机会与能力，数字技术赋权则具有主体多元性。该多元性也构成“技术赋能”和“技术赋权”的本质差别。“技术赋能”强调新兴数字技术对公共管理的效率提升作用，主要将对象局限于政府等公权力部门的治理行为；而“技术赋权”则指向所有主体的社会参与和治理协同价值，强调新技术传播与扩散对政治、经济和文化发展所产生的影响。

第二，从赋权方式看，数字技术赋权包括重组架构和增强话语两种。两种赋权方式在当下大数据、人工智能技术赋权中均有体现。

前者既体现为数字技术对公共部门的赋能，以及对公民个人的增权，通过双向赋权，改变之前金字塔式的权力模式和单向的权利/权力结构，进而塑造一种线上线下、虚实相生、去中心扁平化的社会结构；又体现为不再仅由权力机关操控技术治理，个人也可通过赋权机制享有一定的社会参与权。后者明确指出，所赋权利/权力本质上是主体交流中的话语权。

第三，从赋权效果看，数字技术赋权重在所引起的社会结构与社会关系变革。“数字化生存天然具有赋权的本质，这一特质将引发积极的社会变迁。”而且，数字化通讯技术在提升现代主体之间交流效率的同时，也促进了主体的权能。在当今社会治理中，大数据、物联网早已超越纯粹技术，深刻影响到社会结构和社会关系调整，促使多主体围绕某一阶段的社会公共福利，展开多层次、宽领域的整体性、交叉性治理。“网络强国”“数字中国”“智慧社会”等建设要求，从生产力和生产关系角度看，本质就是建设数智化、信息化的社会生产体系、生活体系与交往体系。

（2）数字技术差异赋权的主要表现

首先，基于社会治理目标，数字技术对政府进行了最大化赋权。在数字经济时代，政府活动的数字化和信息化程度也成为衡量政府现代化水平的一个重要指标。以“政务云”平台及平台运营依赖的上百种系统和小程序为支撑，国家从数据收集、风险预警、科学决策、指挥调度，以及社会服务全方位进行数字政府和电子政务建设。在此方面，我国已然走在了世界前列，并促使政府社会治理活动正在经历前所未为的数字化转型。即以获取、共享和分析数据为基础，以面向和经由数据的治理为机制，重构政府、市场和社会关系，努

力形成“用数据说话、用数据决策、用数据管理、用数据创新”的治理体系。此时，数字技术必然与政府治理密切相关。

其次，在经济效益驱动下，数字技术对企业的赋权也不断强化。在互联网、人工智能和大数据等新技术推动下，我们迎来一个快速发展的数字经济时代。中国在此轮技术革命中更是成就斐然，无论是车联网、人工智能、大数据、云计算和 VR/AR 产业规模，还是相关科技企业融资与盈利，均位于世界先进水平，实现了数字经济时代的“弯道超车”。数字技术发展之所以在商业活动中获得最快发展，既是因为二者有着共同的数字经济发展目标，还是因为相比政治领域，经济领域提供了更强的灵活性和更大的试错空间。在数字红利驱动下，一些大型技术公司凭借强大的数据收集、存储和处理能力，顺利地进行着政企合作，并借助所获取的关键数据，最大限度地发挥数据的价值。

再次，受我国“大政府”体制影响，数字技术对个人的赋权实效并不乐观。依据技术赋权理论，数字技术赋权个人重在提升其在公共活动中的话语权。但现实中，公共参与却并非数字技术赋权个人的最主要内容。一方面，虽然互联网提供了民主参与的渠道，但政府的“信息控制者”角色却全程监控该民主活动；另一方面，由于数字信息获取使用需要一系列的主体及环境条件，该数字参与只是拓展了少数人或精英群体参与社会活动的便捷性，对于广大社会公众而言，他们对数字技术的运用仍止于便捷生活，并不关心所涉权利问题。

4.1.2 数字技术差异赋权引发的现实风险

第一，塑造强大的“数字利维坦”。随着数据时代的到来，人们

的生活方式、工作方式因技术变革而日渐便捷，使人们越来越深切地感受到某种“受缚于数字”的无奈感，这一现象背后隐藏数字技术滥用的新型危机，即“数字利维坦”。这是数字技术差异赋权带来的最直接风险，主要存在于国家与个人关系处理场域，是盲目追求“全景式监控”所产生的必然结果。当技术发展到大数据和人工智能时代，数字技术全面深入到社会生活方方面面，其影响也从经济领域扩展到政治领域，一旦数字技术被国家所俘获用于社会控制和政治权力再生产，“数字利维坦”必然演化成为国家利维坦的新形式。一方面，国家“利用人工智能技术的价值和工具理性编织新型的权力网络，国家意志通过算法制定得以展现，以此加强监控能力和社会管理能力。”另一方面，国家又借助其权力结构，以及与高科技企业的密切“合作”，获得最高的赋权契机，拓展新型数字化权力。此“数字利维坦”的影响可分为直接和间接两种。所谓直接影响主要指的是数字监控。在数字化生存中，政府通过数字治理使“压服”变得直接有效，具体表现为，借助强大的监控和数据分析系统，威胁公民的隐私保护和合法权益；依靠对数字技术及算法分析结果的盲目依赖，导致数据独裁并剥夺人类的自由意志选择；固化信息产生和传输路径，引爆国家安全问题；提供极端主义的温床，加剧社会破裂化风险。相比之下，间接影响主要指的是话语权控制。由于数字政治的非均衡，政府与个体在技术与信息的掌握和使用方面天然不对称。不断变化的新技术不仅最先由国家精英团体掌握并操控，而且还因公权力的社会治理特性，使国家继续获得新兴技术的解释权。

第二，导致平台控制与垄断信息资源。相比“数字利维坦”，该

风险主要存在于拥有信息控制权的互联网企业与个人的关系处理场域，是数字技术赋权企业与个人之矛盾体现。尽管相对于国家，企业和个人同属“弱势一方”，但企业组织尤其是大型数据控制企业在数据信息的获取、加工和使用中优于个人已是不争事实。大数据时代的经济日趋成为平台经济和流量经济，当数据信息作为全新生产要素投入到企业生产经营的各环节，趋于“理性经济人”的盈利性本质，个人想获得某些有价信息就必须付出相应的隐私或资本对价。在利益导向下，平台私权力作为“巨大的力量倍增器”，不仅名正言顺地行使一些公权力职能，而且在定向推送中，剥夺了公民的选择权。该信息资源垄断与控制主要源于自动化决策过程中，互联网企业对技术赋权的过度使用。现代信息技术诱惑人类把判断交给数据，并进一步依靠算法进行决策。在数据“喂养”下，算法又不断实现对传统决策歧视现象的复现、加剧与新增。从当下信息实践及学者关注来看，算法歧视其实纳入了两类构形相近但性质迥异的歧视类型。其一是反垄断语境下的价格歧视，该类型最典型的就是“大数据杀熟”。其核心问题是经营者依靠大数据和算法决策进行差别对待，进而威胁社会公平。其二是平等权语境下的身份歧视，主要呈现为基于性别、残障、种族，以及传染病原携带等特定集体身份实施的区别对待或造成的区别影响。

第三，加剧数字不平等和公民离散。该风险主要存在于数字技术向社会公众赋权过程中，因公民个体获取和运用数据信息的能力及水平不同，所产生的数字技术赋权差异是一种数字不平等。“数字不平等”产生于数字技术对不同公民个体的非均衡赋权，其最终结果是公民主体间的数字红利差异，以及参与社会活动时的公民离散。

数字技术设计之初，我们预设公民具有大体相当的信息接受与处理能力，但实际上，公民个体之间却客观存在不容忽视的技术能力差别，而且伴随社会活动数字化程度不断深入。相比前两者，该类型差异赋权风险较为隐蔽，但却不可被忽视。当前，我们对大数据、人工智能等技术的评价仍是采用一种“低伦理维度”标准，未能从数字红利平等和数据要素分配正义角度来考量。数字不平等和公民离散主要展现在两个层面。一是数字技术本身蕴含运用中的实质不平等。对此，最典型的就是算法本身。现代算法之运用大都蕴含着一个前提——所面对的主体都乐意而且有能力参与到信息收集中，此时社会弱势群体就会被刻意忽略。在信息分类和识别中，弱势群体的回应往往会被计算模型当作杂音或无用数据，因而自然成为数据收集中的暗点或盲点。二是“信息茧房”效应造成知识碎片化，进而影响人们的理性思考。

4.1.3 数字技术风险法律规控的具体展开

（1）明确公民信息权利及框定公权力监控边界

在数字技术差异赋权的法律规控中，首先要处理的就是国家公权力与个人私权利之间的结构性失衡。借助数字技术实现的便捷赋能，国家往往会秉持社会治理目的，扩大公权力监控的边界。在数字时代的虚拟社会里，公权力监控的核心目的主要有二个，一是鼓励创新；二是保障安全。这两方面目的，也大体设置了信息处理活动中公权力监控的边界。前者是为了维护科技的进步，避免我们在本轮数字信息革命中处于劣势而再次饱受欺凌；后者是为了证成安全对于政治社会的重要意义，通过国家权威拱卫自己的“自由领地”。

然而现实中，该公权力监控的边界却很容易僭越。由于我国非

常强调政府信息保密工作，以及全面推行开来的实名制，均使得一个一个的“秘密花园”成为市政广场的建设用地。当由上至下的监管目标易于达成，而由下至上的监督目的易于失效，克服社会治理纯技术依赖带来的主体地位差异，以及避免精英阶层权力的延伸失控与失灵矛盾的最有效手段就是借助重新赋权，将被智能技术“索取”的权利重新赋予公民，积极保障公民的信息权益，在强化对个人信息主体赋权的同时严格信息处理者义务。具体到操作中，主要就是借助新兴权利理论，塑造并保护公民个人的信息权利。该个人信息权利重点集中于两点，第一，通过确认数字身份重建社会关系。第二，界定公民信息权利。公民信息权利作为一种类型权利，由多项子权利构成。该权利设置主要针对政府、互联网企业的信息垄断和信息控制行为，其在理论上并不必然与特定的部门法相联系。因此，当前越来越多的学者将信息权利作为领域法中的权利，立足整个信息社会背景来反思。

（2）遵从数字资本运作逻辑规制自动化决策

不同于政府只是将数字化技术作为社会治理的工具，企业直接将数字化技术作为自身发展的内容，它们全面参与数字经济建设及数字技术开发，甚至将数字化程度作为企业的生命。此时，数字活动在企业间就悄然生成一套以知识、信息和数据为核心要素的数字资本运作逻辑。该逻辑是在已有法律设定的规范框架内，通过具有数据处理能力的平台组织生产，通过数字劳动生产社会关系，通过风险投资来实现资本最大化，以及通过交易和消费行为进行大数据统计。其核心在于借助不断完善的“关系化”，形成一般数据并依此做出决策。

此时，如何在保证数据充分流动和交易前提下，维持个人信息权利保护与企业发展创新之间的均衡，才是问题解决之本。基于数字资本运作逻辑，当下学者对自动化决策的法律规制主要集中于两点，即警惕中立立场被贴上“多数人同意”标签，以及避免利益相关者间的商谈成为摆设。围绕这两方面，学者们要么从完善算法程序、确保“知情-同意”、设定自动决策模型等方面出发，进行算法歧视的内部规制；要么从歧视的一般性理解出发，借助基本民事立法、经济立法，以及劳动与社会保障法，基于相关权利保障中的责任义务设置进行外部规制。其目的都是保证公民在公共信息及社会决策面前的知情、同意、参与和监督权利，建构起算法规制的法律框架。从当下来看，该自动化决策的法律规制主要还是诉诸立法。除了制定《个人信息保护法》《网络安全法》等为数字化生活中的公民信息权益保障提供直接依据外，而且针对平台信息垄断和算法推荐两种最典型行为，国务院反垄断委员会及国家互联网信息办公室、工业和信息化部、公安部、国家市场监督管理总局四部门也分别出台《国务院反垄断委员会关于平台经济领域的反垄断指南》和《互联网信息服务算法推荐管理规定（征求意见稿)》，对平台算法权力监管作出尽可能具体化的规定，从行政角度出发设置主体责任和处罚标准。

（3）通过“公-私”法体系保障公民的平等参与权

该保障主要针对“数字不平等”及公民离散问题，是对公民主体之间差异赋权的法律回应。在宪法上，该权益也常被界定为公民平等参与权的内容。实践中，该平等参与权的法律保障遵从下述逻辑。

第一，明确法律原则。该原则也是弱势群体权利保障的指引思

想。主要包括两点，首先，平等保护原则。即在“强弱并存”的数字社会，我们对“数字弱势群体”之平等参与权的回应。其次，倾斜保护原则。相比平等保护原则主要强调形式正义，该原则更多地是从实质正义层面考量。

第二，完善法律体系。体现在当下已有的法律规范中，并未针对数字技术赋权的差异性，以及信息公平等问题专门立法，有关公民数字素养和企业数字伦理的规范也只是停留于政策层面，但我们仍可以从已有立法中探寻保障各社会主体信息公平及数据信息权益的规范依据。比如，在《老年人权益保障法》《残疾人保障法》及《无障碍环境建设条例》中，都通过设定国家的救助义务，满足该弱势群体的信息获取与运用权益；在《个人信息保护法》《数据安全法》等新兴立法中，针对不同社会主体在信息处理活动中的不平等地位，设定劣势一方权益保障的规范条文，并通过法解释将其适用于司法实践。

第三，确定主体义务。该义务既是确定法律责任的依据和前提，又是化解平等主体非均衡社会参与的保障。数据信息蕴含的信息利益和人权价值，既可为实现不同社会群体之间的信息公平提供公法/私法双重保护，又有助于确定阻碍信息公平实现的信息控制主体，明确其在具体监管、保障及协助方面的责任义务。结合数字技术非均衡赋权样态，该义务主要包括国家的当然义务和平台的特定义务两类。前者包括，强化基础设施建设和普及，保证公共信息便捷获取和使用，制/修订智能产品运用及平台监管法规，承担弱势群体培训与教育，以及就高频服务事项设置线下办事渠道等；后者包括，进行产品适老化改革并提供有效备选，基于算法审核抑制不平等的

信息处理，探索多元化信息投放和均衡化信息分配，以及将非歧视设为人工智能开发的行为规范等。

4.1.4 小结

伴随数据信息成为基本社会资源和重要生产要素，数字技术也以新理念、新业态、新模式全面融入人类经济、政治、文化、社会建设各领域和全过程。面对现代数字技术差异赋权引发的现实风险，我们必须在已有的技术性回应基础之上，探寻必要的制度性回应方式，其中最主要的就是对此差异化赋权进行法律规控。此时，如何借助法律规范的制定与实施，规制数字技术对国家、数据控制企业及个人的非均衡赋权难题，保证信息在不同社会主体间公平分配，让更多社会主体共享数字红利，已成为数字社会公平正义的应有之义。从最终效果来看，该法律规控既可以促进完善有关数据权益及个人信息保护的规范依据，又可以通过明确国家、大数据平台的法律义务，实现社会主体之间的数字红利互补。

4.2 数字社会风险治理的数智化逻辑

当前，我们已经进入了以数字化、网络化和智能化为显著特点的发展新时期，以 5G、物联网、大数据、人工智能和区块链为代表的数智化技术，正在引领并构筑新的社会治理体系。党的十九届四中全会强调必须加强和创新社会治理，强调科技支撑以完善社会治理体系，提高预测预警预防各类风险能力是社会治理的重要内容。在数智化使能的全新社会治理体系中，其主体构建、组织方式和运作机制呈现新特点，通过新兴技术逐步实现对社会运行的精确感知、公共资源的高效配置、社会风险的及时预警和突发事件的快速处置。数智化时代下的治理理念已发生变革，从经验治理向数据治理、从

被动响应型治理向主动预见性治理转变，风险治理的科学化、精细化和智能化水平将得以大幅提升。

4.2.1 大数据驱动风险治理逻辑变革

（1）风险治理原则与模式转变

社会治理原则植根于平衡社会秩序与活力。纯粹的自上而下的治理方式将导致活力不足，而盲目的自下而上的治理方式将导致秩序缺失，因此通过智能算法，以及大规模、全要素社会模拟，在双向迭代中寻找最优次序与活力的动态平衡点，最大化收益，最小化风险。

新一代大数据、智能技术与各层面（国家、社会、社区和组织等层面）治理体系的深度融合，通过提供全景视图和实时分析，提供主动地、有针对性地识别和分析各类社会风险问题，提升各层面治理体系的系统性和统筹性，进而使得社会风险治理从以往以应急主导性的末端治理，向以智能预警与预判为主导的源头治理模式转变。尤其面向重大决策的风险治理模式是源头治本、预防为主、科学决策、权责统一，兼顾公平与效率。

（2）治理方法从社会化向全面数智化转型

陈国青等学者指出围绕信息技术研究的方法论范式可以分为模型驱动与数据驱动两大类，在模型驱动范式下，研究者基于观察抽象和理论推演建立概念模型和关联假设，再借助解析手段对模型进行求解和优化；或利用相关实证数据对假设进行检验；数据驱动范式则借助于统计分析、数据挖掘和机器学习等手段，从数据入手，直接发现特定变量关系模式，形成问题解决方案，进而凝练规律和理论。相关研究形成从自动化、集成化和数据化再到数智化的跃迁。

同理，治理方法从社会化向全面数智化转型，从经验治理转向精准治理。在数智化过程中，社会像素向更细粒度发展并极大提升了数字成像，大数据技术通过对海量数据的快速收集与挖掘、分析与共享，为社会风险治理的科学决策和准确预判提供了有力手段；风险识别不仅基于历史数据和经验，风险识别与预警机制由静态向动态实时转变，面向舆情监测、公共服务、公共卫生和应急事件响应等风险治理逐步精准化、精细化。

进一步，数智化社会风险治理体系通过平台构建愈加可靠和有效，社会风险治理的关键信息和渠道移至数字化风险治理平台，对风险识别与应对的流程进行变革与优化，对特定边界内外的风险治理进行数字化重塑，实现网络化风险协同管控与治理——防御主动化、运营智能化、操作实战化、修护韧性化。

4.2.2 数智化赋能风险治理体系

数智化时代的风险治理应基于数据基础设施，将人工智能、空间计算和复杂科学方法融会贯通，实现适应新时代的风险识别、预警与应对方法，构建由“数”到“智”的风险治理体系。这套体系由数据智能、空间智能和社会智能三个层面组成，相互交错，螺旋上升，以应对从社区至国家不同层次的风险。

（1）基于大规模数据智能的风险治理

随着数字社会程度的推进，对数据的使用也提出了更高的要求。数据智能旨在对大规模数据进行获取、处理和分析，挖掘数据蕴含的信息或知识，使数据拥有智能。因此，基于大规模数据智能的社会风险治理要建立在大规模数据的融合与管理，以及动态数据的实时分析基础上。大数据技术通过对海量数据的快速收集与挖掘、分

析与共享，成为支撑全息社会治理与风险管控的有力手段。

通过社会智能感知的数据具有两方面的重要特征，一方面，数据具有跨媒体、跨领域和多源异构的特点，对数据与知识的智能发现与分析融合带来了挑战；另一方面，数据收集的深度、广度和粒度不断增大，其蕴含的数据价值不断累积，如此海量的数据对社会计算的研究创造了丰富的社会条件，有利于数字社会与复杂系统的模拟与仿真，以发掘人类社会的活动规律与问题，实现基于稳态社会的风险治理。

数智化时代背景下，个人和社会数据源源不断产生，对这些数据进行实时交互分析，数据智能融合提供可粒度缩放、可跨界关联、可全局融合视图的全景式数据融合机制，有助于社会异常现象的发现与应对。例如，新冠疫情背景下社会新问题的发现与应对，以及社会关系的动态演化，这就要求构建面向动态变化的风险治理体系。而随着社会关系与对象的动态演化，增加了数据分析融合过程中的实体和关系的判别难度，以及数据的分析难度，因此从“大”数据中发现“小”概率的社会风险事件仍然具有很大挑战性，人类智慧与经验依然要发挥其作用。数智社会的风险治理需要“人在环中”，当新的社会异常产生后，应能对它进行及时且合理的验证，以便该社会风险能得到有效的应对与处理。

（2）基于多尺度空间智能的风险治理

空间计算综合了复杂性和复杂系统、信息论和数据科学、数据建模与仿真，以及信息可视化、赛博空间、本体论、语义计算等技术，以支撑社会计算急需解决的问题。空间计算是通过不断深化空间计算理论，通过智能化方法来形成一套社会风险治理体系，典型

场景包括防灾应急和传染病防御等。基于空间智能的社会风险治理体系，其优势在于空间计算在数据获取及数据处理方面，为社会风险识别与分析提供了重要的数据源。

数智化时代下，空间数据呈现多源异构形态，灯光、POI 和手机信令等各类地理空间数据已被用于刻画社会情境，可生成宏观、中观和微观等多尺度的关键变量，尤其在收集、处理和分析面向大尺度的社会科学问题时具有显著优势。多源空间数据和空间计算方法（例如空间相关性、空间可视化和地理编码等）为理解社会发展和社会风险另辟蹊径，有助于发现社会潜在风险的时空分布及其演变，从而解决社会风险治理体系的复杂自适应性问题，赋能面向空间计算的社会风险治理。同时，多尺度伸缩的空间智能依赖于可信赖智能计算技术，例如计算安全性、隐私保护、算法可解释性和公平性等。

（3）基于复杂系统的社会智能与风险治理

在数智化过程中，社会像素向更细粒度发展并极大地提升了数字成像，大数据技术通过对海量数据的快速收集与挖掘、分析与共享，为面向社会智能体系中的科学决策和准确预判提供了有力手段。数智化促进了人工社会的形成，人工社会是数字社会的仿真模型，由无数称作为“主体 agent”计算单元构成。物理社会与人工社会之间不断交织，构成了巨大的、更为复杂的社会物理系统。因此，数智化下的社会风险治理应具有复杂的适应性，面向不同主体和群体挖掘其演化规律，并提供与主体适配的智能解决方案。

复杂系统由大量个体构成，由于个体之间的相互作用，复杂系统不是个体性质的简单之和，而呈现关联、合作和涌现等集体行为，

具有非线性和动态性、非均衡、非周期性和开放性等一系列特征。复杂性科学是运用非还原论方法研究复杂系统，从而产生复杂性的机理及其演化规律的科学，其兴起和发展代表了一次重要的科学思维变革，并充分体现了学科交叉融合的特点，涵盖了非线性科学、混沌理论、分形学、模糊学、信息论、控制论、自组织理论、系统论和耗散结构论等不同分支学科的内容。复杂科学常见建模方法，包括基于主体的建模、元胞自动机、蚁群系统、进化算法、神经网络、机器人学、分析学、图形学和群体智能等。

随着数据科学和人工智能技术的快速发展，复杂性科学与人工智能相得益彰，可以将已有的建模方法与人工智能相融合，通过数字孪生建立高维仿真系统，在增强社会风险数据处理能力的同时，提升社会风险模型的拟合程度，以此形成具有高水平的社会风险事件预测、预判能力，以及风险识别能力的社会智能。在社会、经济、管理领域的数字风险系统演化、高阶复杂网络建模与分析、数字社会下人机交互风险等，也恰恰是社会计算与社会智能面向复杂社会系统的目标。例如，结合图表示与深度学习算法，对数智化下的大规模社会风险体系进行网络化建模与学习，智能识别数智社会中的规律和潜在风险，并赋予及时有效的干预策略。

4.2.3 数字社会风险治理方法与机制创新

（1）面向动态复杂网络的风险治理方法创新

数智化时代的风险治理体系从社会化转向全面数智化转型，从经验治理转向精准治理。在数智化过程中，社会像素向更细粒度发展并极大提升了数字成像，大数据技术通过对海量数据的快速收集与挖掘、分析与共享，为社会风险治理的科学决策和准确预判提供

了有力手段，风险识别不仅基于历史数据和经验，风险识别与预警机制由静态向动态实时转变，面向舆情监测、公共服务、公共卫生和应急事件响应等风险治理逐步精准化、精细化。

在大数据环境下领域情境、决策主体、理念假设和方法流程等决策要素都在发生变化，在领域情境方面，支持风险治理与决策的信息多源异构；在决策主体方面，人机交互与协同成为趋势；在理念假设方面，通过假设来概括不可测现象的情形受到冲击；在流程方法方面，传统的是从近线性的、分阶段的决策过程向复杂交错互动的非线性模式转变。

在复杂系统相关研究中，风险分析不再仅局限于对风险源相对独立地加以辨识、评估和控制，复杂系统内风险源及其关联关系交织形成风险网络，复杂系统风险分析相关研究采用基于图论、复杂网络理论、贝叶斯网络和神经网络等多种方法。进一步探索高阶网络的表示学习方法，结合图表示与深度学习算法，对数智化下的大规模复杂社会系统进行网络化建模与学习，智能挖掘数智社会中的风险并赋予有效的干预策略。同时，社会风险治理过程须嵌入有效的因果推断。智能风险识别、朔源、预测乃系统工程，须厘清复杂网络系统中各要素的因果关系，灵活运用关联分析法、干预法和反事实法进行科学的因果推断。

（2）基于区块链的风险治理机制创新

数智化时代下数据要素流通和数据资产确权是风险治理的关键。区块链技术具有去中心化、分散性、透明性和不可篡改等特征。在数据层面，区块链架构改变了传统资产流动中的“中心化”模式，赋能数据资产构成、确权与流通。在特定领域（例如金融和税务等），

区块链赋能其风险管理。基于供应链金融运营的全周期，区块链技术得以从数据的获取与甄别、交易流程的优化、风险管控与行为监管等层面进行广范围的赋能。进一步，通过区块链的标准化行业准则，完善信息和网络安全的监管制度，构建危机入侵监测、灾害恢复等管理制度，在风险发生的第一时间启动应急措施，将风险尽量降到最低。

在社会治理层面，有效利用区块链将使社会治理实现可程序化的算法治理，将区块链应用于风险治理中的数据安全风险（如数据质量和数据隐私安全等）、算法安全风险（如算法垄断和算法歧视等）、社会安全和文化风险（如信任变革、离群文化风险、开放文化风险和匿名文化风险等）。此外，构建区块链治理联盟是风险治理的新方式，基于联盟自治的区块链跨链机制是以链治链，通过构建多方共治的区块链联盟管理跨链网络，以解决区块链之间的数据共享、价值流通与业务协同问题，以此加强风险防范与管理。

4.2.4 小结

新一代数智技术应与各层面（国家、社会、社区和组织等层面）治理体系的深度融合，通过提供全景视图和实时分析，以及主动地、有针对性地识别和分析各类社会风险问题，提升各层面治理体系的系统性和统筹性，进而使得数字社会风险治理从以往以应急主导性的末端治理，向以智能预警与预判为主导的源头治理模式转变。通过大数据平台与人工智能技术，构建起一套能够自动实现数据挖掘、收集、分析、预测和预警的新型治理机制，并在数智化基础构建面向新型人工智能系统的智能仿真云、面向问题的复杂系统智能仿真语言、面向边缘计算的智能仿真计算机系统、跨媒体智能可视化技

术等。数智化下的风险治理体系具有高时效、高适应性，上下双向结合，进而强化数字社会风险治理决策的前瞻性和科学性。

4.3 面向企业风险智能分析的“人在回路”范式

当前复杂多变的环境下，企业往往生存和发展于各种风险和挑战中，提前对风险进行感知和预防是提高企业竞争力和维持可持续发展的重要手段。同时，企业间的连接越来越紧密，当一家企业发生风险时，风险很容易向其他正常企业蔓延，导致在关联公司之间引起连锁反应，严重时可能引发系统性金融风险。因此，企业风险智能分析受到越来越多专家和学者的重视。

随着人工智能技术的发展，机器学习、深度学习等模型被逐渐应用到企业风险分析上。以机器学习为代表的人工智能技术在特定领域任务已经表现出强大的信息处理和数据挖掘能力，能够有效地弥补人类分析决策中有限理性的局限。人工智能技术所具备的计算和感知能力，是模型与数据结合并以大量计算为特征所形成的机器智能，其在存储、搜索、感知和模式识别等严格定义的问题上性能表现优异，但在高级认知和复杂决策方面与人类智慧依然相去甚远。企业风险智能分析是一项高风险性的复杂管理任务，本身具有不确定性、非线性及动态性等内禀属性，仅依靠人工智能技术已经不足以应对风险智能识别和复杂管理决策中的挑战。为此，弥合风险智能分析任务的不确定性和机器模型能力的确定性，人类决策的鲁棒性和机器模型的脆弱性，数据分析的综合性和机器能力的片面性，是企业风险智能识别的未来发展新方向。

因此，本文构建了面向企业风险智能分析的“人在回路”新范式，引入人的智慧与决策，强调人、机器及数据之间的优势互补，

有机协作。同时，本文还分析了人类智慧与机器智能在“人在回路”新范式中具备的能力和发挥的作用，最后归纳出了人机协作的四种有效模式。

4.3.1 相关工作

（1）“人在回路”范式的相关研究

“人在回路”这一框架被提出用以提高系统的人机耦合性能，建立科学的人机协作机制。该框架尝试从更高维度设计混合智能管理系统，将人类真正融入系统构建、部署、决策的全流程中，并从机器智能的能力构建，以及引入人类智慧的人机协作模式构建两个层面切实推进人类智慧和机器智能协同合作，从而提高机器辅助决策分析的适用性和可靠性，解决面向更高层次的处理复杂问题的能力。

当前学术界对“人在回路”学习框架的相关研究主要是从以下三个方面展开，一是机器智能的能力构建，主要从性能、可解释性和可用性三个角度进行研究；二是人机协作模式的构建，主要从人类角色、交互接口和反馈方式三个角度进行研究；三是人类智慧和机器智能的任务分配，以财务规划、新的人机混合范式来预测强对流天气中的紧急决策、胸片诊断和皮肤癌诊断。

上述研究仍存在一些局限性。例如，当前研究主要针对机器学习任务流程中的某一环节提出人的介入，相关工作呈现碎片化，尚未有相关工作讨论全流程人类智慧的深度融合。另外，面向单一机器学习任务的基础能力构建，尚未有相关工作面向复杂决策系统提出灵活协作的“人在回路”范式。同时，在管理科学领域，面向企业风险智能分析的人机回路的协作范式也研究较少。

（2）基于人工智能技术的企业风险分析方法

早期对企业风险的预测与评估主要通过传统的专家判别法和数理统计模型，如 5P 原则、LAPP 原则、Z-score、Zeta 和 Logistic 回归模型等。随着人工智能技术的发展，机器学习模型逐渐被应用到企业风险预测上来。

基于机器学习的企业风险分析方法主要可分为两类，一类是基于企业自身属性特征进行建模。用于建模的数据仅和企业自身有关，比如企业的基本信息、财务指标、股权结构和年度报告等数据。常用算法主要包括支持向量机（SVM）、集成学习算法、决策树和神经网络模型等。近些年，为解决企业风险预测的动态性问题，循环神经网络及其改进模型开始广泛应用到风险预测中。另一类是基于企业的关联特征属性建模。由于企业及其相关主体（企业和金融机构等）间的依赖性日益增强，通过交易、担保、借贷等方式形成了错综复杂的关联关系，研究人员开始尝试以网络或图结构的形式表示这些企业关系，通过图表示学习等深度学习方法对这些企业关联关系进行表征，预测出目标企业的风险概率。

然而，企业风险的产生和发展是一个动态的过程，具有典型的非线性复杂系统的特征，其多层次和多重反馈特性使得企业风险识别变得异常复杂，动态性及关联性特点也增大了人们对风险预见规律的认知难度。单纯依赖端到端的机器学习模型得出的结论很难验证和可信，同时也很难通过统一的模型为企业给出一个确定的方案。

因此，本文构建了面向企业风险智能分析的“人在回路”新范式，将人类智慧引入企业风险智能分析这一高风险性、高不确定性的复杂管理任务中，通过人类智慧和机器智能的有机协作，实现对

企业风险精准、安全、稳定、动态的混合智能决策。

4.3.2 “人在回路”基本范式设计

“人在回路”新范式考虑以机器智能模块化的方式重新构建处理流程框架，如图 4-1 所示，从模型输入、中间结果、决策流程和优化目标等角度引入人类智慧，形成多模块和松耦合的学习框架，人和机器的关系是以人为核心，人指导机器，人机相互协作。其中，机器智能的参与过程可以总结为：①人对实际问题进行问题分解，转化成机器可以处理的任务，并将不同的任务下达给机器；②机器作为实现工具，完成被分配的任务后，将结果返回到系统中，由人类或其他模块进行后续处理；③人通过介入、调节、修正和反馈等方式与机器进行合作，将人类的经验智慧与机器的计算能力相融合，给出最后的判断。

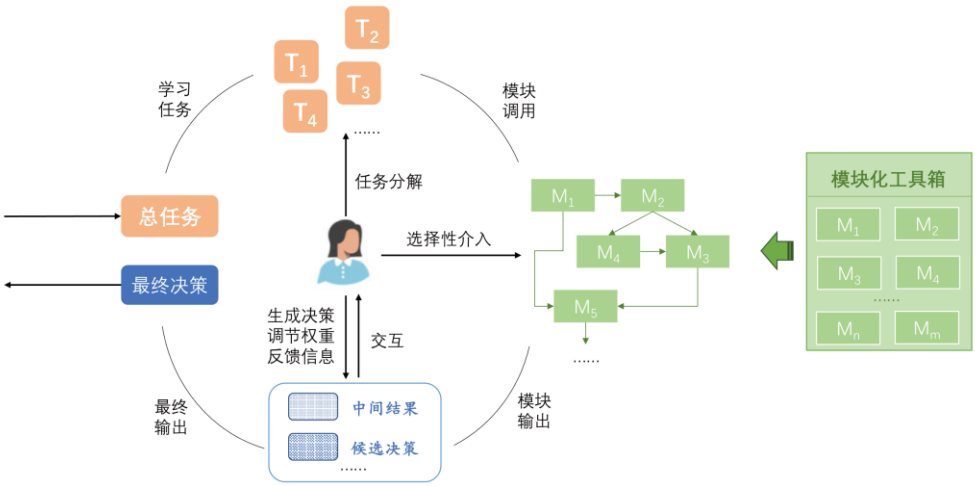

图 4-1 “人在回路”范式的流程框架

“人在回路”范式下的机器智能由多个结构简洁、目标明确的

模块组成，每个模块负责单一的任务，模块间保持低耦合度。与没有人类智慧介入的机器智能的主要区别在于：①模型输入。不限于任务开始时接受的数据，而是将人的倾向性具体化到模块中。②模块输出的中间结果。需要根据置信度评估来选择决定是否交由人来协助处理，人也可主动选择调节不同模块的权重来引导模型或取舍结果。③决策流程。机器智能的输出结果不再是决策结果，而是候选决策，以及相应的支撑信息，如决策可信度和相关案例分析等，最终让人来决定最终决策。④优化目标。不再是单一的精准率等，而是模块间的耦合度、人的参与程度和系统运行的综合绩效等。

此外，能够实现与人动态交互的机制是“人在回路”范式的首要条件。人从外部进行的动态交互方式是丰富的、多元化的，每个模块需要按照实际输入情况设计人介入的接口，同时以参数设置等方式让人来选择当前是否介入和介入程度等。另外，我们拟引入反馈回路，将这些输入再次反馈给模型，加大对人类反馈的关注，进而提升模型的可信度。

4.3.3 “人在回路”范式中人类智慧和机器智能的能力

（1）人类智慧引导的机器智能模块协作能力

从“人在回路”的流程设计来看，人类智慧与机器智能的协作可划分为两个阶段的回路：线下开发回路和线上预测回路。

在线下开发回路中，需要预先设计和开发多种机器学习模型，以及多个自动化分析模块，关键在于将知识工作划分为离散的任务，降低模块间的耦合度。人可以根据实际使用场景对机器智能的功能模块进行调研和设计，根据功能和结构进行拆解，降低模块间的耦合度，使得每个模块的职责相对独立，形成工具化、模块化的机器

智能工具箱，提升算法的复用性和可拓展性。

在线上预测回路中，为了完成实际任务，需要选择的合适的模块组织起来，借助不同模型或者模块提供的信息，从性能指标、可解释性、可用性多个角度使模块组合达到最优。

（2）基于多策略的高效数据增长能力

目前在机器学习中，利用开源数据进行分析的瓶颈之一是缺少大量的标注样本，且获得高质量标注样本的成本很高。“人在回路”的范式通过人的介入与协作，缓解机器智能面临的小样本和低资源的问题，实现机器智能的高效数据增长能力，降低数据成本。在线下开发阶段，利用主动学习策略让机器主动优先选择最有价值的未标注样本进行标注，以尽可能少的标注样本达到模型的预期性能，用少量的人类工作参与达到高效的数据增长。人的介入，除了实施具体的数据标注，还负责不同标注策略的选择及迭代次数的控制。同时，对于多模态数据，应用数据增强策略来实现数据的多元化增长，在不标注新样本的情况下以不同的方式进行数据补全和扩充。

（3）基于迁移学习的应用任务拓展能力

在“人在回路”范式下，我们更希望将一个经过精心设计、训练完备且性能尚可的机器学习模型价值最大化，尽量避免由新任务带来的模型重构的时间成本和人力成本。因此，本文设计的“人在回路”范式，利用迁移学习的方法来提升模块的应用任务拓展能力。

以企业风险级别预测模块为例，假设我们已根据有标记的原领域$D_s = \{x_i, y_i\}_{i=1}^{n}$构建并训练了一个用于预测企业风险级别的机器学习模型 Ma，然而随着时间的推移，受市场环境、相关政策变化等因素的影响，领域专家对企业风险等级的划分标准可能进行了更改，

即产生新的目标领域$D_t = \{x_i, y_i\}_{n+1}^{n+m}$，此时数据分布和标签均发生了变化，$P(x_s) \neq P(x_t)$。基于迁移学习的思想，我们无需从头开始重新构建一个新的机器学习模型来应对这种变化，而是可以基于预先训练好的模型，借助D_s的知识，更快速地获得一个可用于新的目标域D_t下的企业风险等级预测任务的模型。迁移学习的思想保留了在企业风险预见场景中已经构建的模型的部分结构，将已有模型拓展应用到其他的相关任务中，增强了已有模型的应用价值，同时帮助机器和人更快速地完成新模块的实现。

4.3.4 基于“人在回路”的企业风险智能分析框架设计

我们将“人在回路”基本范式具化到企业风险智能分析这一任务中，得到图 4-2 所示的框架。

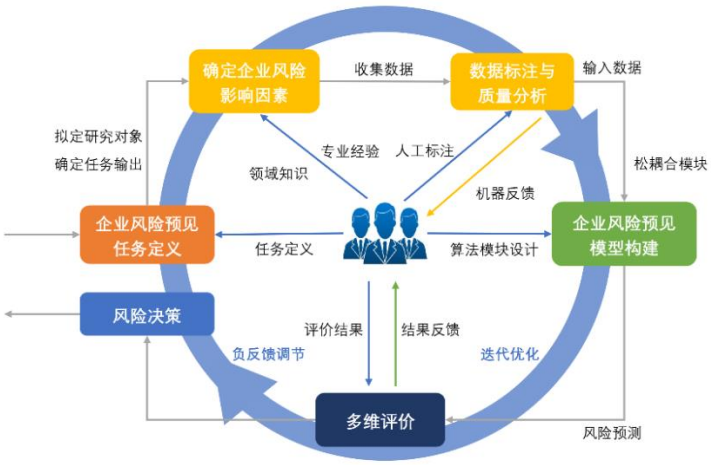

图 4-2 基于“人在回路”的企业风险智能分析框架

首先，人类根据企业风险智能分析场景的实际情况进行任务定义，拟定研究对象的企业类别、行业等，并确定任务的具体输出，如风险的类别、风险的等级等。其次，人类基于领域知识和专业经

验确定企业风险影响因素，指挥机器从合适的来源获取相应的开源数据，并在机器的辅助下完成数据标注和质量的管控。之后，人类根据企业风险智能分析任务的需要，构建多个低耦合的机器学习功能模块，并对每一个模块进行训练。在实际风险分析任务场景中，研究人员对功能模块灵活选择与组合后进行线上预测，模型将基于预测结果从风险预测准确度、资源成本消耗度、模块组合灵活度、人机协作适应度多维指标进行评估，并根据人类的反馈予以修改和提升。最终，经过迭代优化的机器智能模块作为工具，辅助人类做出合适的企业风险决策。

4.3.5 “人在回路”范式的人机协作模式

我们从“人”的角度出发，考虑了人机混合智能系统中人类智慧参与的多种差异，如介入时机、交互内容、扮演角色的不同等，归纳出组织型、监督型、决策型和合作型四种人机协作模式，在不同任务和场景中实现高度灵活的智能融合。

（1）组织型人机协作模式

机器在处理任务时具有单一性和特定性，更倾向于在小任务上达到“专”和“精”。而人擅长综合分析处理，因此，在图 4-3 所示的组织型人机协作模式中，人作为组织者，将项目分解成为多个特定的、具体的学习任务，交由不同的机器处理，从不同维度进行知识挖掘，将人的综合性和机器的单一性结合起来。而多种机器模型，则作为随需调用的工具包，支撑完成每个子任务。这种由人来完成任务设计与任务分配，将机器模型工具化、组件化的人机协作模式，是“人在回路”的机器学习范式的核心思想。

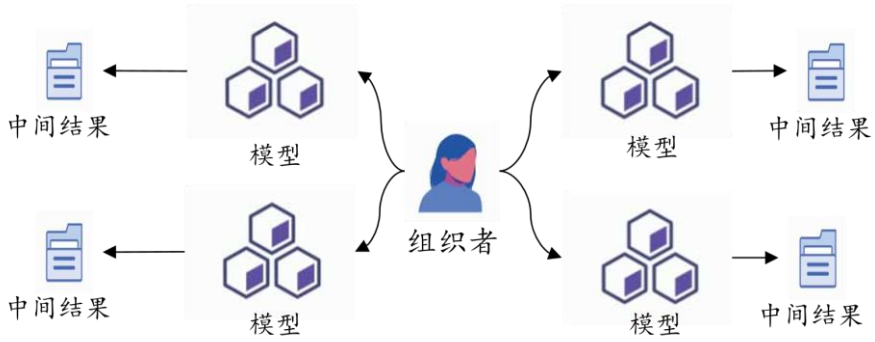

图 4-3 组织型人机协作模式

（2）监督型人机协作模式

机器智能的另一个局限性在于其对数据的敏感性，即轻微的噪声扰动也可能会干扰到结果的正确性。因此，监督型人机协作模式是将人引入进来作为监督者，在关键决策点给出专业判断。如图 4-4 所示，该模式主要应用于两个阶段：①在线下开发回路中，人以数据标注、数据修正、数据反馈和结果分析等方式与机器协作，这个过程本身也是一个回路，旨在使用较少量的数据，使模型达到尚佳的性能。②在线上预测回路中，则面临的是信息量过大的问题，此时机器是处理海量信息的主力，人的介入主要是辅助筛选、甄别机器把握不准的信息，防止重要信息的遗漏。模型通过设置可疑性和敏感性指标，将超过阈值的结果反馈给人，由人来监督判断后，再返回给机器。

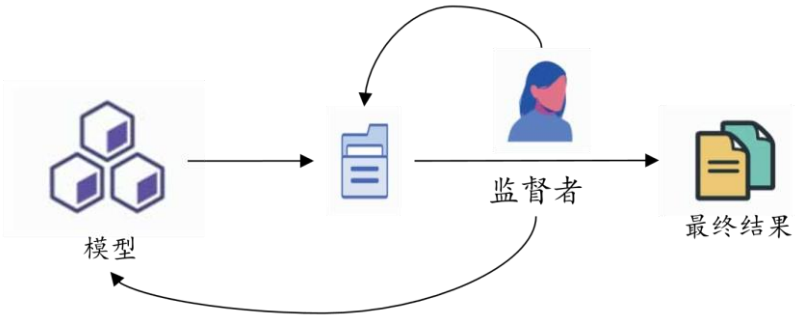

图 4-4 监督型人机协作模式

（3）决策型人机协作模式

人类常采用动态决策的策略，即根据情况的不断发展作出当前最合理的决策。而机器给出的结果是确定的，无法动态调整，我们将之称为机器对决策结果的脆弱性。实际上并不需要完全由机器作出决策，只需要将机器学习模型工具化，让它给出可选的决策候选集，最终的判断还是由人来把控，从而实现决策的动态性。在图 4-5 展示的决策型人机协作模式。首先，判别出全流程中的关键决策点，确定需要人介入的位置。关键的决策点可能是中间节点，也可能是末位节点，目的在于通过人机协作来提升中间结果的可信度；然后，根据机器给出的置信度、以及动态环境的变化情况等，由人决定是否介入决策。当需要人的介入时，机器通过给出候选决策集、相似案例提示等信息支持，以生成最终决策。整个过程需要形成一套标准化的分析流程。

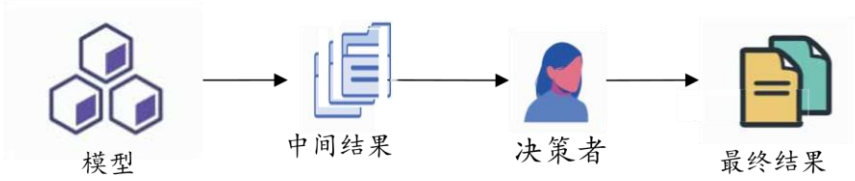

图 4-5 决策型人机协作模式

（4）合作型人机协作模式

机器学习，尤其是深度学习模型，在决策过程中的可解释性很难得到保障，造成人与机器合作的障碍。此时，如果人完全依赖于机器，容易引发自动化偏差的问题。

合作型人机协作模式考虑了人的自主权和纠错能力，保证决策的合理性，同时，需要把人在决策时的衡量因素传递给模型，提升模型在可解释性的方面的表现。如图 4-所示，在交互过程中，人以提供经验和知识的方式介入到机器学习的流程中来，这些知识可以首先由人整理，生成大规模的知识库或领域先验知识，然后通过学习或者互动的方式集成到模型中来；然后由人评估机器的表现，并生成反馈，以反馈的方式与机器连接，在迭代中不断改进。在不同的实践环境中，根据不同的任务分工，四种模式可以单独或是共同使用。这种灵活的按需组合使用的人机协作机制为人类提供自主权来指导、监督和反驳机器，可以增加算法的透明度，提升人的把控和对算法的信赖，防止由于算法的错误或偏差造成决策失误，实现人机共融和人机制约。

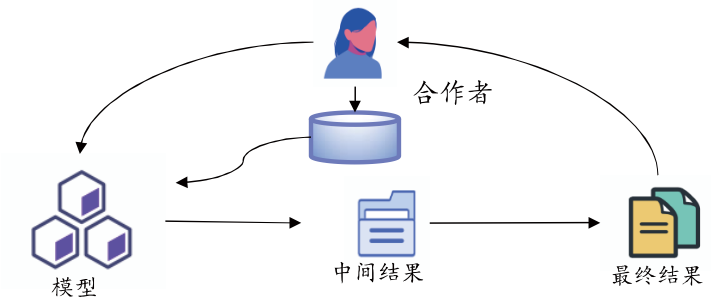

图 4-6 合作型人机协作模式

4.3.6 小结

尽管当今人工智能技术发展的如火如荼，然而其蕴含的机器智能在解决复杂决策任务时存在诸多问题和局限，如对数据质量的高度依赖，预测结果缺乏可解释性等，在企业风险管理领域的实践中仍很大局限性。可见，在面对高级认知和复杂决策特点的企业风险智能分析任务时，人类智慧不可缺失。本文分析了“人在回路”的基本范式以及人类智慧和机器智能在该范式下发挥的模块协作能力、高效数据增长能力以及任务拓展能力。基于此基本范式，本文创建了面向企业风险智能分析任务的“人在回路”创新范式，使风险预见、感知和分析决策的关键机制更多的回归到专家知识和智能技术的自然融合。最后，本文还归纳出了“人在回路”范式的组织型、监督型、决策型和合作型四种人机协作模式，具有深刻的理论意义和广阔的应用前景。

4.4 “共建共治共享”的数字社会风险治理制度创新

从物理世界到数字世界，数字化过程本身就会带来原先没有的或原先有但现在发生变化的新型风险。数字化安全是一种全新类型的安全范畴，不同于网络空间安全，数字化安全是一项系统性的、

专门针对数字化转型中普遍存在的总体性安全。物理世界安全、数字化安全、网络空间安全才能构成完整的安全范畴。国家层面的数字化安全体系包括：数字化技术应用中的总体技术安全评估、数字化技术应用中的应用场景风险评估、数字化技术应用中的配套风险控制机制设计，以及数字技术应用中的常态化监管体系建立、相关的标准、制度与法律体系建设等。目前，讨论较多的是物理世界安全和网络空间安全，但是数字化安全的研究还很少，而对于数字社会风险的治理，数字化安全必然是其治理的目标之一，需要在未来加强研究与实践。

随着信息技术在各行业各领域的普遍应用，整个社会运行对大数据的依赖程度越来越高。大数据为完善社会风险治理制度带来的巨大价值不言而喻，然而其引致的数字社会风险问题也引起各领域学者的重视和思考。数字社会兴起重塑人们的数字社会风险观，数字社会中的硬件、软件、算法、文本、符号等风险要素有别于传统社会风险要素，数字社会风险新现象、新形式、新内容、新载体对传统的行政法规、部门规章、司法解释、政策文件和规范性文件提出新要求。数字社会中政府、社会组织、网络运营者、新旧媒体、网民等多元主体参与风险治理的逻辑、依据、模式、路径、激励机制等亟需理论界和实践界在制度创新、体制创新、机制创新等方面深入探讨。数字社会的风险治理也离不开基于大数据思维的方法和制度创新，需要厘清多元社会主体之间的逻辑关系、权责关系，建立基于大数据手段的数字社会风险治理制度。

4.4.1 “共建共治共享”的数字社会风险治理制度框架

社会风险治理是社会治理的重要组成部分和关键核心问题，自

然也需要按照多元化逻辑进行治理制度创新。党的十九届六中全会《中共中央关于党的百年奋斗重大成就和历史经验的决议》中强调，要“建设共建共治共享的社会治理制度，建设人人有责、人人尽责、人人享有的社会治理共同体”。共建共治共享式社会治理制度（见图4-7）是国家治理方针的重要组成，也是大数据赋能数字社会风险治理的破题思路。大数据思维带来的社会风险治理方法和工具的革新已经出现，其为完善社会治理制度带来的巨大价值不言而喻，然而如何将之用于数字社会的风险治理仍需探讨。数字社会风险治理始终受到数据结构不全面、数据信息不对称、多元主体不协调和区域发展不平衡等问题的困扰，迫切需要基于大数据的思维和工具，融合多源数据建设数字风险治理大数据平台，引导多元主体共同参与数字社会风险治理，统筹城乡数字风险治理协同发展，切实实现共建共治共享式的数字社会风险治理。

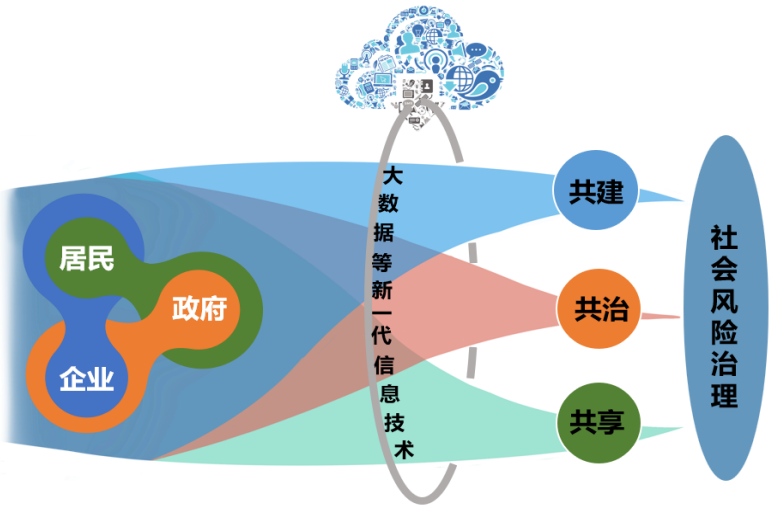

图 4-7 “共建共治共享”的数字社会风险治理制度

4.4.2 共建：融合多源数据建设数字风险治理大数据平台

政府要做好社会风险治理，要善于利用以大数据为代表的新一代信息技术。面对数字技术全方位渗透带来的新挑战，借助数字化技术加强数字社会风险治理，提高数字社会风险治理水平成为各级组织的共识。从数字社会生态体系治理的角度出发，数字社会的风险治理需要系统性谋划、综合性布置、体系性推进，融合源于智能传感、自动识别、自动定位、移动互联、区块链、大数据和人工智能等数字化技术的多样数据，建设基于现代信息技术的社会风险治理数据平台，进而合理用于数字社会的风险治理。

（1）多源数据融合建设社会风险治理数据平台

数字社会风险治理是现代化社会治理的新内容，政府要做好社会治理，也需要善于利用以大数据为代表的新一代信息技术。基于大数据的算法技术依赖于丰富的结构化数据，愈是庞大的样本、丰富的维度，大数据所能实现的工作和输出的结果则愈是多样。在数字社会风险治理中，政府可以依托在长期实际工作中所积累的基础数据，对数字信息采集过程、数字信息传输过程、数字信息整合共享过程、数字信息资源应用过程的风险评估和量化建模，实现对数字社会风险的精细化治理，然而数据来源的可靠已逐渐成为制约面向应用的算法开发的重要因素。

构建面向数字社会风险治理的社会风险治理数据平台，需要融合政企民的等多方数据。对于政府，需要将不同科层部门的数据进行融合、统一监管，在便于精准化供给公共服务的同时，着力保护居民隐私。一方面，通过一系列政府公共服务 APP 的设计、开发和投入使用，将涉及个人信息的相关业务进行专业化管理；另一方面，

利用身份证号等唯一可标识属性对居民基础数据进行链接匹配，实现政府端居民数据的融合，将数据采集、数据存储、数据分析和数据挖掘等环节集成，破除“数据孤岛”“信息烟囱”，从顶层设计推动数字政府的安全落地。其次，实现政企间数据融合，探索有效帮助多个机构在满足用户隐私保护、数据安全和政府法规的要求下，实现多源数据融合，并进行数据使用和建模的模式。在政府部门的领导下，共建社会风险治理全景式大数据，提升社会风险治理领域的数据数量和质量，增强大数据社会风险治理领域模型的训练效果，实现数字社会的风险治理。

（2）创新大数据方法实现广泛数据协同

民生数据关乎居民隐私需要严格的管理控制，而依托于大数据的应用技术需要海量数据用于建模训练，如何在保障居民隐私的前提下应用大数据等算法技术是数字社会风险治理的重要议题。联邦学习、差分隐私和区块链等算法技术创新为此提供了解决方案，其在跨机构之间的个人身份确认、企业经营监管和智慧城市建设等众多场景中均有广阔的应用前景。现如今，这些算法技术共同推动着传统数据流通模式和流程的变革，许多技术产品已能满足通用性需求，有望成为全社会数据流通网络的支撑型基础设施，在切实保障民众隐私的基础上，合理利用数字化改革过程中积累的海量运营管理数据和用户内容数据。

共建社会风险治理数据平台离不开数据共享，需要通过方法创新来确保共建平台的各环节安全高效。数据采集过程中，基于密码学的纵向联邦学习技术是实现隐私保护的一个有效途径，尤其适用于需要采集其他机构具有的高度敏感信息的情况。在多方参与的算

法训练过程中，难免需要将数据多方转手，也意味着相关数据存在着泄露风险，隐私计算中的差分隐私技术便适用于这一场景，通过引入噪声对源数据或者统计信息进行干扰，防止原始数据的泄露。区块链作为协作计算中的可信第三方，提供了可靠的传输通道和验证方法，同时由于采用了分布式数据存储方式，区块链还有极强的反篡改性。联邦学习、差分隐私、区块链等基于大数据方法的创新，为实现更广泛的数据协同提供了技术基础，为安全释放大数据的潜在价值提供了实现路径。

4.4.3 共治：引导多元主体共同参与数字社会风险治理

在传统的社会风险治理模式中，政府是治理主体的核心，虽然社会组织、居民等主体起到了一定的作用，但主体多元化的治理理论与实践尚需深入探索。随着互联网技术的普遍使用，企业、居民逐渐参与到社会治理中，成为社会治理共同体的重要组成。实现高水平数字风险治理就需要让多元主体深度参与到社会风险治理中，搭建多主体广泛、深入参与平台，打通多元主体协同治理渠道，实现政府领导、企业参与、居民自治的分布式协作治理机制。

（1）大数据赋能社会风险治理中的政府统筹领导机制

在多元共治的数字社会风险治理机制中，政府要担任好统筹领导角色，协调多方资源、协同多方力量。在多元共治的过程中，由于企业的逐利特征和当前社会组织自治能力的不足，政府需要充分发挥统筹领导作用，制定面向数字社会风险防治的多元治理格局的宏观框架，搭建起政府联系企业、居民的平台。通过以大数据为代表的新一代信息技术进一步对企业和居民赋权，使治理权力由政府垄断转换为多元共享，形成多元治理的社会结构，在推动形成数字

社会风险治理的新格局方面发挥着重要作用。

在数字社会风险治理中，由于治理主体的多元化带来的利益主体多元化，社会风险治理面临着前所未有的复杂局面，构建多元关系协调、权益公平分配的体制和机制具有全局性、根本性作用。面向利益导向型的企业主体，政府需要充分发挥企业自身的能力与特质，同时也需要利用法规等治理工具加以约束。面向发展尚不成熟的社会组织等主体，政府需要加以引导和培育，逐渐加强居民参与治理的意愿与能力。因此，政府在这种复杂的治理关系中必须发挥统筹领导作用，需要在完成自身供给职能的基础上，充分利用以大数据为代表的新一代信息技术搭建多元参与的互动平台；也就是说，以政府为统筹领导的“一核多元”的形态，政府作用发挥如何，直接关系社会风险治理整体效能。

（2）大数据赋能社会风险治理中的企业参与治理机制

相较于政府，企业对于数字社会的构成、变革更为敏锐，具有极强的创造力和社会活跃性。作为数字社会的主要建设者，引导企业参与对整个数字社会风险治理体系的构建至关重要。针对数字社会风险治理的企业参与治理，亟待解决政府与企业的权责划分，厘清政府和企业的权责边界，充分利用政府刚性的行政管理优势、企业专业的服务供给水平，以期将两者在各自边界内实现优势最大化，提高社会风险治理的效率与水平。

权责划分的首要是梳理权配置现状、发掘责权不配位现象，将实施管理和资源调配下放至企业，增强政府的统筹协调能力和工作抓手。在进行政企间的权责划分时需要明确以下三点：一是适当放权，政府相关部门的综合治理权退后一步；二是保留社会风险治理

过程中的行政权力、执法权力，对数字社会风险治理中企业全过程行为的监督权力和治理成效的考核权力，企业主体不应涉及执法权；三相应企业在参与数字社会风险治理的进程中，其管理权、治理权、运营权的侧重点在于借助自身业务能力和独有资源帮助提高数字社会风险治理的效率。

（3）大数据赋能社会风险治理中的居民参与自治机制

针对数字社会风险治理的公众参与治理逻辑，应系统梳理公众参与社会风险治理的参与方式及参与深度的特征，进而分析不同深度层级中公众参与动机以及激励机制，并研究如何促使居民向深度参与转化。通过探索居民通过自治手段来实现社会风险化解于基层，实现自我管理、自我服务、自我教育、自我监督，从源头化解风险。

居民是“政府领导、企业参与、居民自治的分布式协作治理机制”中最基础的治理主体，需要深入挖掘居民自我服务与自我管理中参与动机、参与方式的区别与联系，明确居民参与数字社会风险治理的内涵与外延。居民自我服务与自我管理是在以往居民诉求表达基础上，更深层次、更能体现公众参与理念的参与方式。对于居民来说，其参与数字社会风险治理的过程中，所感知的公共价值和私有价值也会逐渐发生改变，价值感知的变化也将对居民参与风险治理的动机，以及参与深度的转化产生影响，自然也将丰富数字社会风险治理的逻辑内涵。

在此背景下，政府、市场和居民的关系发生了改变，政府包揽逐渐变成政府统筹，传统的全能政府将转向“一核多元”的多主体治理模式，在多元社会治理主体共同参与的情境下，数字社会的风险治理体系变得更加复杂但也更加有效。

4.4.4 共享：统筹城乡数字风险治理协同发展

数字时代下，以智能手机为代表的新型交互媒介正改变着社会沟通形式，重构了人们的工作方式和社交行为。数字交互为个人和社会的合作交往提供新的途径，在给人们的生产和生活带来无限便利的同时，也催生了数字鸿沟，加剧了社会分化。让边缘和少数群体共享社会发展成果，是中国特色社会主义的本质要求。党的十九大明确要“建立健全城乡融合发展体制机制和政策体系”，想要打破数字鸿沟、让人人享有数字社会的便捷，就需要协同城乡，实现不同区域的协调发展。

公共资源本身具有逐利和集聚的趋向，自然会选择发展条件更为成熟的城市地区，数字资源也不例外，互联网企业往往优先布局发展较为成熟、消费能力较强的中大型城市，这也使得乡村地区面临较多不利影响。因此亟需快速清除各种阻碍城乡资源共享的制度藩篱和体制障碍，建立数字社会风险治理信息共建共享网络，开发智能、自动和精准的数字社会风险识别预警系统，逐步建立起涵盖主体、制度、硬件、软件、平台和内容等数字社会风险治理体系，提升数字社会风险治理精细化水平。具体说来，当前需要利用大数据手段，更高效地促进城乡间数字服务和数字网络等资源要素的合理配置与流动，需要在政府统筹下建立平衡调节机制，通过政策工具引导企业将数字服务落在欠发展的城镇乡村。根据实际需要开发相应软件平台，建立有利于资源要素向乡村配置的策略和方案，既要把更多的数字资源投向农村，同时也要确保农村资源在流出过程中发挥更大价值。最终达到城乡资源双向互动和优势互补的协调关系，让数字社会的发展跨越数字鸿沟，让各种各样的社会群体都能

享有数字发展红利。

4.4.5 小结

数字社会风险是数字时代的特殊产物，伴随着数字社会的发展而演化。数字社会风险的治理，也需要在传统治理经验的基础上，结合大数据的思维范式加以研究。数字社会风险治理的成效关乎数字社会的发展态势，共建共治共享的数字社会风险治理制度为解决数据垄断、信息孤岛、隐私泄露、数字鸿沟等各式各样的数字社会风险提供了解决路径。总的来看，数智时代背景下政府与其他社会主体之间的互动增强，共建共治共享式社会治理制度成为国家治理的政策方针，数据处理和信息交互能力的提升为构建数字社会风险治理共同体提供了极大助益，借助大数据和人工智能等新一代信息技术来实现数字社会风险治理多元化新格局成为必然趋势，对于坚持和完善中国特色社会主义制度，维护国家安全、社会安定、人民安宁，具有重要意义。

参考文献

[1] 蔡昌,赵艳艳,李梦娟.区块链赋能数据资产确权与税收治理[J].税务研究,2021,(7):90-97.

[2] 曾彩霞,尤建新.大数据垄断对相关市场竞争的挑战与规制:基于文献的研究[J].中国价格监管与反垄断,2017(6):8-15.

[3] 曾雄.数据垄断的竞争分析路径[C]新时代大数据法治峰会——大数据,新增长点,新动能,新秩序.2017.

[4] 陈兵.“数据垄断”:从表象到本相[J].社会科学辑刊,2021(2):129-136.

[5] 陈国青,曾大军,卫强,张明月,郭迅华.大数据环境下的决策范式转变与使能创新[J].管理世界,2020,36(2):95-105.

[6] 陈思进.元宇宙是下一张互联网?投资元宇宙或10年难赢利[J].法人,2021,(12):18-20.

[7] 陈月华,杨绍亮,李亚光,陈发强.智慧城市安全风险评估模型构建与对策研究[J].电子政务,2020,(5):91-100.

[8] 程亮.新基建背景下的智慧城市建设[J].中华建设,2020,(10):130-131.

[9] 崔岳,黄华,张明星.对抗样本在自动驾驶领域应用现状[J].合作经济与科技,2019,(7):76-79.

[10] 丁文文,王帅,李娟娟,袁勇,欧阳丽炜,王飞跃.去中心化自治组织:发展现状、分析框架与未来趋势[J].智能科学与技术学报,2019,(2):202-213.

[11] 方亚南,齐佳音.数字金融安全与监管.北京:经济管理出版社,2021.

[12] 黄匡时.计算人口学的学科范式、理论基础与技术方法[J].北京工业大学学报(社会科学版),2021,21(3):16-27.

[13] 黄欣荣.大数据时代的还原论与整体论及其融合.系统科学学报,2021,29(3):8-12.

[14] 黄杨森,王义保.网络化、智能化、数字化:公共安全管理科技供给创新[J].宁夏社会科学,2019 ,(1):114-121.

[15] 姜涛.人工智能数据安全风险和有效治理措施[J].法制博览(名家讲坛、经典杂文),2021(22):155-156.

[16] 卡斯特. 网络社会的崛起[M]. 夏铸九等译. 北京：社会科学文献出版社，2001.

[17] 李伯虎,柴旭东,张霖,李潭,卿杜政,林廷宇,刘阳.面向新型人工智能系统的建模与仿真技术初步研究[J].系统仿真学报，2018，30(2)：349-362.

[18] 李三希等.数字经济的博弈论基础性科学问题[J].中国科学基金，2021，35(5)：782-800.

[19] 梁玉成,政光景.算法社会转型理论探析[J].社会发展研究，2021,8(3)：21-43+242.

[20] 林义,刘斌.国家治理现代化视域下我国多层次社会保障制度的创新探索[J].经济体制改革，2021，(6)：20-25.

[21] 刘恩强,刘增良.一种多媒体社交网络安全风险评估方法[J].计算机应用与软件，2015，30(6)：267-268.

[22] 刘晓洁.央行数字货币面临的风险挑战及应对策略[J].人民论坛，2020，(23)：98-99.

[23] 刘晓曼.新基建背景下工业互联网安全形势分析与发展建议[J].保密科学技术，2020，(12)：51-55.

[24] 刘子涵.元宇宙:人类数字化生存的高级形态[J].新阅读，2021，(9)：78-79.

[25] 龙卿吉,魏钰,胡坚波,孔祥涛,汪玉凯,周艳.统筹发展和安全 提升数字化治理效能[J].人民论坛,2020，(33)：46-47.

[26] 龙卫球.数字化时代安全可信的法治保障与新型监管要求[J].传媒，2021，(18)：19-22.

[27] 陆学艺.当代中国社会结构与社会建设[J].红旗文稿. 2010，(18)：40-44.

[28] 罗戎,周庆山.我国数字内容产品消费模式的实证研究[J].情报理论与实践，2015，38(10)：67-72.

[29] 罗双玲,丁雨楠.基于区块链的数字内容治理:考察与思考[J]，2021，7(2)：16.

[30] 吕鹏.智能社会治理的核心逻辑与实现路径[J].国家治理，2021，(42)：28-32.

[31] 孟小峰,余艳.人工智能时代社会计算与社会智能展望[J].中国人工智能学会通讯，2021，11(8)：4-7+21.

[32] 孟小峰,朱敏杰,刘立新,等.数据垄断与其治理模式研究[J].信息安全研究，2019，5(9)：789-797.

[33] 米秀明,黄静."新基建"背景下网络安全技术创新发展研究[J].保密科学技术，2021，(7)：38:43.

[34] 牛喜堃.数据垄断的反垄断法规制[J].经济法论丛，2018(2)：25.

[35] 彭彪.传播新技术的社会风险及其治理[D],武汉大学，2009.

[36] 彭兰.算法社会的“囚徒”风险[J].全球传媒学刊，2021，8(1)：3-18.

[37] 彭磊.新基建时代如何保障工业互联网数据安全[J].中国工业和信息化,2021,(8):38:44.

[38] 蒲清平,向往.元宇宙及其对人类社会的影响与变革[J/OL].重庆大学学报(社会科学版)：2022，(1)：1-12

[39] 邱泽奇等.从数字鸿沟到红利差异——互联网资本的视角[J].中国社会科学，2016，(10)：93-115.

[40] 日立东大实验室.社会 5.0：以人为中心的超级智能社会[N].北京：机械工业出版社.2020.

[41] 沈湘平.元宇宙:人类存在状况的最新征候[J].阅江学刊，2022，14(1)：44-52+172

[42] 时振涛.基于演化 Agent 的人工社会系统建模方法及其应用研究[D].兰州理工大学，2013.

[43] 宋保振.“数字弱势群体”权利及其法治化保障[J].法律科学，2020，(6)：53-64.

[44] 宋华,杨雨东,陶锋.区块链在企业融资中的应用:文献综述与知识框架[J].南开管理评论，2021，1-9.

[45] 孙玮,李梦颖.扫码:可编程城市的数字沟通力[J].福建师范大学学报，2021，(6)：132-143.

[46] 谭九生,范晓韵.算法“黑箱”的成因、风险及其治理[J].湖南科技大学学报(社会科学版)，2020，23(6)：92-99.

[47] 谭昆乐,邓智平.区块链技术赋能社会治理的风险与前景.社会治理，2021，(9)：50-57.

[48] 唐明圣,毛新军,周会平.基于Agent的人工社会建模语言研究.计算机研究与发展，2015，52(5)：1036-1049.

[49] 唐思慧.大数据时代信息公平的保障研究——基于权利的视角[M].北京:中国政法大学出版，2017.

[50] 王飞跃,蒋正华,戴汝为.人口问题与人工社会方法:人工人口系统的设想与应用[J].复杂系统与复杂性科学 ,2005，2(1)：1-9.

[51] 王飞跃,平行系统方法与复杂系统的管理和控制.控制与决策，2004，19(5)，485-489

[52] 王飞跃.基于社会计算和平行系统的动态网民群体研究[J].上海理工大学学报，2011，33(1)：8-17.

[53] 王贵松.数字化管理与安全法的人性之维[J].浙江社会科学，2012，(2)：39-41+156.

[54] 王国成.基于 Agent 真实行为揭示社会经济复杂之谜——集成建模与计算实验的实现途径[J].中国社会科学院研究生院学报，2012 (05)：35-41.

[55] 王少泉.数字时代“信息茧房”现象的生成机理与优化途径[J].重庆科技学院学报，2018，(6)：12-16.

[56] 王思斌.社会学教程[N].北京:北京大学出版社,2010.

[57] 王勇.论数字社会的政权结构失衡及其补正[J].学术交流，2021 ,(6)：67-76.

[58] 王雨薇,国世平.中央银行数字货币面临的挑战及风险防范研究[J].云南财经大学学报，2020，36(2)：12-18.

[59] 危红波.我国数字社会风险治理责任分配[J].学术交流，2021，(10)：130-143.

[60] 吴靖,应武.走向数字社会主义:工业化视角下的马克思主义传播技术批判[J].全球传媒学刊，2021，8(3)：2-22.

[61] 吴善东.数字普惠金融的风险问题、监管挑战及发展建议[J].技术经济与管理研究，2019，(1)：66-69.

[62] 吴忠泽.新基建新技术引领智能交通产业高质量发展[J].中国科技产业，2021，(1)：6-9.

[63] 徐顽强,徐玉婷,兰兰.数字社会中非政府组织参与政府治理的研究综述[j].电子政务，2012，(9)：2-8.

[64] 徐一帆,吕建伟,史跃东,狄鹏.基于贝叶斯学习的复杂系统研制风险演化分析[J].系统工程理论与实践，2019，39(6)：1580-1590.

[65] 杨建辉.对"数据垄断"的几点思考[J].中国证券期货，2017(7)：2.

[66] 尹丽英,张超.中国智慧城市理论研究综述与实践进展[J].电子政务，2019，(1)：111-121.

[67] 袁勇,王飞跃.平行区块链:概念、方法与内涵解析.自动化学报，2017，43(10)：1703-1712.

[68] 袁勇,王飞跃.区块链技术发展现状与展望.自动化学报，2016，(4)：481-494.

[69] 张国清.分配正义与社会应得[J].中国社会科学，2015，(5)：21-39.

[70] 张莉.新基建将成为产业数字化转型加速发展的新基石[J].中国对外贸易，2020，(9)：24-25.

[71] [1]张龙辉,肖克.城市智能治理中的算法失灵及消解策略[J].电子政务，2022(07)：98-112.

[72] 张元好,曾珍香.城市信息化文献综述——从信息港、数字城市到智慧城市[J].情报科学,2015,33(6):131-137.

[73] 赵越强.公共和私有部门数字货币的发展趋势、或有风险与监管考量[J].经济学家,2020 ,(8)：110-119.

[74] 周尚君.数字社会对权力机制的重新构造[J].华东政法大学学报,2021 ,(5)：17-26.

[75] 周翔,刘欣.数据垄断的困境与隐忧[J].人民论坛，2013(15)：20-21.

[76] 邹静,张宇.数字金融的研究现状、热点与前沿——基于 Cite Space 的可视化分析[J].产业经济评论，2021，(5)：133-146.

[77] Bayatbabolghani F, Blanton M. Secure Multi-Party Computation [J]. Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, 2013.

[78] Builder C H , Bankes S C .Artificial Societies: A Concept for Basic Research on the Societal Impacts of Information Technology[J]. 1991.

[79] Burrell, J., & Fourcade, M. The Society of Algorithms[J]. Annual Review of Sociology, 2021, 47:213-37.

[80] Chen Y, Ding S, XU Z, et al. Blockchain-based medical records secure storage and medical service framework[J]. Journal of Medical Systems, 2019, 43(1):1-9.

[81] Szegedy C , Zaremba W , Sutskever I ,et al. Intriguing properties of neural networks[C].In 2nd International Conference on Learning Representations, 2014.

[82] Fan Z, Fang H, Zhou Z, et al. Improving Fairness for Data Valuation in Federated Learning [J]. ArXiv, 2021, abs/2109.09046.

[83] Fan F , Xiong J , Li M ,et al. On interpretability of artificial neural networks: A survey[J]. IEEE Transactions on Radiation and Plasma Medical Sciences,2021.

[84] Grossman S J , Hart O D .The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration[J].Journal of Political Economy, 1986, 94(4):691-719.

[85] Hanseth,Ciborra. Risk, complexity and ICT.Northampton, MA: Elgar,2007.

[86] Hart O, Moore J. A Theory of Debt Based on the Inalienability of Human Capital [J]. Social Science Electronic Publishing, 1994, 109(4):841-79.

[87] Oliver H , Andrei S , Vishny R W .The Proper Scope of Government: Theory and an Application to Prisons[J].Quarterly Journal of Economics, 1997, 112(4):1127-1161.

[88] Yogamani S .Adversarial Attacks on Multi-task Visual Perception for Autonomous Driving[J].Journal of Imaging Science and Technology, 2021.DOI:10.2352/J.ImagingSci.Technol.2021.

[89] Jam, V. D. A., Agm, V. D. J. The First-Level Digital Divide Shifts from Inequalities in Physical Access to Inequalities in Material Access[J]. New Media & Society, 2019, 21(2):354 - 375.

[90] Jan A.G.M. van Dijk. The Network Society: Social Aspects of New Media London: SAGE Publications,2006.

[91] Lu J , Sibai H , Fabry E .Adversarial Examples that Fool Detectors[J]. 2017.

[92] Kerber W .Digital Markets, Data, and Privacy: Competition Law, Consumer Law, and Data Protection[J].MAGKS Papers on Economics, 2016.

[93] Lazer, D. The rise of the social algorithm[J]. Science. 2015,348(6239):1090-1091.

[94] Lazer, D., Pentland, A., Watts, D. J., Aral, S., Athey, S., Contractor, N., & Wagner, C. Computational social science: Obstacles and opportunities[J]. Science.2020,369(6507):1060-1062.

[95] Olhede, S. C ,Wolfe,et al.The growing ubiquity of algorithms in society: implications, impacts and innovations[J].Philosophical transactions of the Royal Society. Mathematical, physical, and engineering sciences, 2018,376(2128):20170364.

[96] Peter Gomber, Jascha-Alexander Koch, Michael Siering. Digital Finance and TinTech: Current Research and Future Research Directions[J]. Journal of Business Economics.2017,(87):537-580.

[97] Price R J, Shanks G G. A semiotic information quality framework: development and comparative analysis [J]. Journal of Information Technology, 2005, 20:88-102.

[98] Chesney R , Citron D K .Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security[J].Social Science Electronic Publishing, 2018.

[99] Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J. F., Breazeal, C. & Wellman, M. Machine behaviour. Nature. 2019,568(7753):477-486.

[100] Scekic O, Miorandi D, Schiavinotto T, et al. SmartSociety -- A Platform for Collaborative People-Machine Computation[C] IEEE International Conference on Service-oriented Computing & Applications. IEEE, 2016.

[101] Shen J., Zhou J., Xie Y., Yu S., Xuan Q. Identity Inference on Blockchain Using Graph Neural Network. Blockchain and Trustworthy Systems. BlockSys[J]. Communications in Computer and Information Science, 2021:3-17.

[102] Shuai W, Yong Y, Wang X, et al. An Overview of Smart Contract: Architecture, Applications, and Future Trends[C],Proceedings of the 2018 IEEE Intelligent Vehicles Symposium.

[103] Wang F-Y, Xiao W, Li L, Li L. Steps toward Parallel Intelligence[J]. IEEE/CAA Journal of Automatica Sinica, 2016, 3(4):345-348.

[104] Wang F-Y. The Emergence of Intelligent Enterprises: From CPS to CPSS[J]. IEEE Intelligent Systems, 2010, 25(4):85-88.

[105] Wang X, Zheng X H, Zhang X Z, Zeng K, Wang F-Y. Analysis of cyber interactive behaviors using artificial community and computational experiments[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47(6):995-1006.

[106] Wirth N. The Development of Programs by Stepwise Refinement [J]. Comm. ACM, 1971,(14):221-227.

[107] Zhang Z Y. Digital Rights Management Ecosystem and its Usage Controls: A Survey [J].International Journal of Digital Content Technology & Its Applications, 2011, 5(3):55-272.

附录

主要作者简介（按姓氏拼音排序）

孟小峰 中国人民大学教授，CAAI 社会计算与社会智能专委会主任，ACM 中国 SIGSPATIAL 分会主席。主要研究领域包括数据智能、数据治理、社会计算与社会智能等交叉学科。发表学术论文 200 多篇，出版专著“网络与移动数据管理”3 部曲等。曾获省部级特等奖 1 次、二等奖 3 次，中国计算机学会“王选奖”一等奖、第三届北京市高校名师奖等。
齐佳音 中国人工智能学会社会计算与社会智能专委会副主任，上海对外经贸大学人工智能与变革管理研究院院长、教授，国家级人才计划入选者，七次入选爱思唯尔中国高被引学者，研究方向为前沿技术与决策创新。主持完成国家级的重大/重点科研项目，成果发表在 Information Systems Research、《管理科学学报》国内外重要学术期刊。
邓建高 河海大学商学院副教授，CAAI 社会计算与社会智能专委会委员，主要研究方向为情感计算、行为分析、网络舆情治理。
傅湘玲 北京邮电大学副教授，CAAI 社会计算与社会智能专委会委员，主要研究方向包括智慧金融、深度学习、自然语言处理以及文本挖掘等。
黄匡时

中国人口与发展研究中心研究员，中国人工智能学会社会计算与社会智能专委会副主任，第七届“春晖杯”中国留学人员创新创业大赛优胜奖，第七届中国人口科学优秀成果奖一等奖，在《人民日报》、《人口研究》、《中国人口科学》、Journal of Family Issues 等刊物发表论文数十篇，主要研究兴趣：数字和计算人口学、计算社会科学。
李瑾颉 上海师范大学，博士/讲师，主要研究方向为数字风险和算法治理。
李三希 中国人民大学经济学院教授，中国人民大学数字经济研究中心主任、数字经济文理交叉平台首席专家，主要研究方向为数字经济、信息经济学和产业组织理论，主持和完成国家自科基金多项，发表论文 30 余篇。
刘鲁宁 哈尔滨工业大学长聘教授，经管学院副院长，国家社科重大项目首席专家，中国人工智能学会（CAAI）社会计算与社会智能专业委员会委员，主要研究方向为大数据赋能基层社会治理、数字化赋能乡村振兴、人工智能赋能公共服务。
宋保振 山东大学副教授，“山东大学法学院大数据法治研究中心”研究员，主要研究方向为数字治理、数字法治。
吴超 浙江大学公共管理学院教授，CAAI 社会计算与社会智能专委会委员。

吴联仁 上海对外经贸大学副教授，主要研究方向为智能社会治理，主持和参与国家级科研项目 10 余项。
宣琦 浙江工业大学教授，浙江工业大学网络空间安全研究院院长，CAAI 社会计算与社会智能专委会委员，主要研究方向是人工智能应用及安全、社交网络数据挖掘、网络信息传播。
余艳 中国人民大学副教授，CAAI 社会计算与社会智能专委会委员、副秘书长，主要研究领域包括数字转型与创新、知识管理、智能社会。
朱宏淼 上海对外经贸大学副教授，主要研究方向为知识管理、复杂网络传播动力学分析及应用、公共卫生管理。