

国际数字生态指数的测算与分析*

乔天宇^① 张蕴洁^② 李铮^② 赵越^③ 邱泽奇^{**②④⑤}

①北京大学大数据分析与应用技术国家工程实验室 北京 100871

②北京大学社会学系 北京 100871

③北京大学工学院 北京 100871

④北京大学中国社会与发展研究中心 北京 100871

⑤北京大学数字治理研究中心 北京 100871

摘要: 数字化发展已经成为影响国际发展新格局的重要变量。如何对国际数字化发展格局作出判断,是亟待研究的时代性议题。从数字生态的视角出发,提出了刻画国际数字生态的基本框架,包括数字基础、数字能力、数字应用和数字规制四个核心维度。根据这一框架测算了国际数字生态指数,对各国数字化发展状况予以评估,并挖掘各国数字生态发展的不同结构模式。同时,通过研究数字化领域国家间互动关系网络,考察各国占据的结构位置和所受的结构约束,希望为我国制定相关发展策略提供参考。

关键词: 数字化发展; 数字生态; 数字治理; 国际数字生态; 数字生态指数

DOI: 10.16582/j.cnki.dzzw.2022.03.002

一、引言

数字技术变革已是当下驱动变迁和发展的主要动力。在数字技术创新的直接作用下,发展数字产业和传统产业的数字化已成为驱动经济发展的新动能,同时,数字技术也正在渗透到社会发展的方方面面。数字化也与世界政治、经济、安全等多方位格局密切相关,为应对数字化转型带来的机遇与挑战,各国(地区)已出台相关战略。例如,欧盟委员会发布《塑造欧洲数字化未来》(Shaping Europe's Digital Future)发展战略,明确提出要维护欧盟的数字主权,提升其在数字经济领域的竞争力^[1];新美国安全中心(CNAS)推出的《设计美国的数字发展战略》(Designing a U.S. Digital Development Strategy)也把数字化发展放在事关美国国家安全和全球战略的重要位置^[2]。

中国同样高度重视数字化发展,自党的十八大以来,党中央曾多次强调数字经济事关国家发展大局,要将其置于国家战略的高度。《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》也对中国未来的数字化转型发展的方式和路径进行了顶层设计^[3]。

当今世界正处于百年未有之大变局,数字化发展无疑是影响这一变局的关键变量。全球数字化发展中的机遇与挑战同在,竞争与合作并存。在这样一个新的国际竞技场上,中国目前处于一个什么样的位置?换言之,我们应如何对各国数字化发展状况进行科学评估,进而对数字化发展的国际格局予以刻画?目前,国内外各类机构已研发出一系列指数工具用于评估各国数字化发展状况^[4],这些发布机构根据其自身的利益需要、价值取向和研究重点所制定的指数工具各具特色,都可作为研究各国数字化发展状况的重要参考。但它们在反映数字化发展国际格局时仍存在一些不足。首先,测度内容上有偏重,多以对数字经济的测度为主,但仅针对数字经济开展评价难免有失偏颇。其次,既有的绝大多数指数仅以国家(或地区)为单位,对其发展程度进行排序,但排序只是刻画格局的一种方式,难以反映更为重要的

*基金项目:中国科学院学部工作局“中美欧数字治理格局研判关键问题研究”(项目编号:2021-ZW07-B-014)。

**通讯作者

收稿日期:2022-03-03

修回日期:2022-03-05

国家间互动关系。

国际视野下的数字生态研究将涉及环境、生态系统、群落生态三个方面的内涵,将影响数字化发展的制度环境、数字化发展中涉及的多重要素及其关联,以及国家间关系等予以综合考量。从数字生态视角出发能更好地刻画各国数字化发展及与此相关的国际格局。本文从以下几个角度对国际数字生态予以研究:第一,构建一套刻画国际数字生态的基本框架,并系统开展数据收集工作;第二,测算国际数字生态指数,并对指数测算结果展开相应分析;第三,利用国家间关系网络刻画各国在国际发展格局中的结构位置,例如生态位特征,并对国家间的关系性质给予判断。

二、刻画国际数字生态的基本框架

我们从数字基础、数字能力、数字应用和数字规制这四个维度对各国的数字化发展情况予以研究(参见图1)。其中,数字基础、数字能力和数字应用与中国数字生态指数关注的三个主要维度相同^[5]。数字规制是在考察国际数字生态时新增加的维度,其反映了影响一国数字化发展的制度环境,体现了生态概念的环境内涵。

首先,我们将数字基础细化为基础设施和数据资源两个维度。其中基础设施进一步考察能源和硬件,能源包括电力供给量、供给的覆盖性和稳定性、新能源等内容,硬件考察数字通信设备拥有状况。数据资源关注数据规模和数据开放,分别反映国家的数据资源存量和开放性,前者主要考察数据中心和网民数量等,后者进一步分为准备、实施和成果三个维度。

其次,数字能力体现数字人才和技术创新状况。前者考察各国的数字人才规模、结构和流动。后者包含成

果和开发者两个维度,成果由一国的研究论文、专利申请、在线创新力和创新成果转化状况等内容体现,开发者由一国的开发者年龄、相对经验、学习新技能的频率、薪资水平和工作环境等内容体现。

再次,为体现主要数字应用场景,我们考察了数字政府、数字经济和数字社会三个维度。数字政府包括在线服务、电子参与、开放政府数据和地方在线服务等内容。数字经济以数字产业化和产业数字化为核心内容,同时对数字贸易予以考察,包括数字产业增加值规模、数字企业竞争力、产业数字化规模、产业数字化转型水平以及数字货物与服务贸易等状况。数字社会则聚焦接入性、数字健康和数字教育三个方面:接入性反映各国居民对互联网和ICT的接入和使用情况;数字健康反映在医疗健康领域中数字技术的应用状况,如医疗机构或患者对远程医疗和移动医疗等的使用;数字教育反映数字技术在教育领域的应用情况,如学校中学生与教室的人均上网设备、学生在课堂学习的ICT内容等。

最后,数字规制包含数字商务、数字流通和数字风控三个维度。数字商务关注各国在电子商务和数据税方面的立法情况;数字流通从法规和政策的角度对各国公共数据、私人数据以及数据的跨境流通的潜在能力进行衡量;数字风控考察个人信息保护法规和数据安全法规两方面的内容,对各国应对数字化进程中潜在风险的能力予以评估。

以上四个维度,共同构成刻画国际数字生态的基本框架。本研究将以这一框架为基础,测算国际数字生态指数。指数的指标体系与上述框架相对应(详见表1)。

三、国际数字生态研究的数据来源

本研究针对41个国家收集相关数据,这些国家包括中国、27个欧盟国家、英国、美国,以及俄罗斯、澳大利亚、加拿大、日本等各大洲或地区的主要国家。为契合研究目的,我们在数据收集上,一是坚持“多源”原

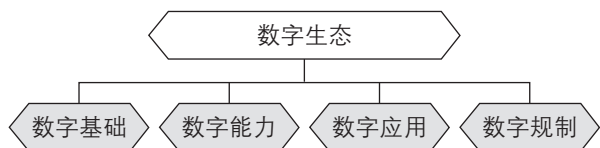


图1 刻画国际数字生态的基本框架

表1 国际数字生态指数指标体系设置

一级指标	二级指标	三级指标
数字基础	基础设施	能源
		硬件
	数据资源	数据规模
		数据开放
数字能力	数字人才	人才数量
		人才结构
		人才流动
	技术创新	成果
		开发者
数字应用	数字政府	在线服务
		电子参与
		开放政府数据
		地方在线服务
	数字社会	接入性
		数字健康
		数字教育
	数字经济	数字产业化
		产业数字化
		数字贸易
数字规制	数字商务	电子商务法规
		数字税法规
	数字流通	公共数据开放法规
		私人数据开放法规
		数据跨境法规
	数字风控	个人信息保护法规
		网络安全法规

则，广泛收集来自各类组织机构发布的能够反映相关内容的的数据；二是关注能直接或间接反映国家间相互作用的数据，这为利用关系网络刻画国家间数字生态奠定基础。当有多个类似数据源存在的情况下，研究优先选择那些由权威机构发布、能够覆盖更多国家、年份最新的数据源，保证数据具有权威性、全面性和时效性等特征。课题组共收集到了来自十余个机构的28个数据源，并将其综合运用在国际数字生态指数测算和对国家间关系的判断上。这些数据大体上可以划分为以下五种类型：

第一，统计数据。这类数据主要来自一些权威机构的统计部门以及专门的行业组织，包括如联合国、世界银行、欧盟、国际电信联盟（ITU）、世界贸易组织（WTO）、经济合作与发展组织（OECD）、全球移动通信系统协会（GSMA）等。在考察基础设施、数字社会中的接入性、数字经济中的数字经济增加值和数字贸易规模等内容均会涉及此类数据。在利用国家间关系刻画数字生态时，数字商品与服务贸易活动是国家间相互作用的重要体现，联合国国际贸易数据库（UN Comtrade Database）对国家间贸易互动有详尽的统计。

第二，既有指数型数据。这类数据是经过一定测算程序得到的结果，多以标准化得分形式呈现，能反映各国在特定方面的发展水平以及相对排名状况。体现数据开放、数字政府、数字经济中的产业数字化等内容的数据均属于此类。其中，数据开放分别引用了万维网基金会公布的“开放数据晴雨表”全球报告^[6]和开放数据观察组织发布的“开放数据盘点”年度报告^[7]中的结果；数字政府引用了联合国“电子政务调查报告”中公布的在线服务、电子参与、开放政府数据和地方在线服务四个子指数；数字经济的产业数字化内容中包含了欧盟委员会测算的国际数字经济社会指数和戴尔科技公布的2020年数字化转型指数。

第三，微观调查数据。除统计数据和既有指数型数据外，本研究还收集到了一些与研究框架相关的微观调查数据。比如针对技术创新维度中的开发者情况，我们选用了全球最大的开发者在线技术问答平台Stack Overflow每年所做的开发者调查数据；针对各国数字教育情况，经合组织（OECD）所做的国际学生评价项目（The Programme for International Student Assessment，简称PISA）是颇具国际影响力的教育主题调查项目，我们从中选取了反映学校使用ICT和接入互联网状况、学生在课堂学习的ICT技能等内容的数据。与此前两类数据不同，微观数据除了能够汇总得到对国家层面属性的

测量外,我们还可以利用其开展更多样的研究工作。

第四,文本型数据。对数字规制的研究主要依据文本型数据。我们检索了目标国家相关官方法律及管理条例文本,由编码员根据给定框架进行编码。编码框架是在世界银行的“全球数据规制诊断”(Global Data Regulation Diagnostic)^[8]评估框架上修改得到的^{注1}。

第五,来自互联网平台的数据。还有部分研究使用的数据来自互联网平台,比如衡量数字能力的论文状况数据来自Web of Science,反映数字专利状况的来自佰腾网的专利大数据^{注2},数字人才的部分数据引自清华大学经济管理学院互联网发展与治理研究中心与领英(LinkedIn)合作发布的《全球数字人才发展年度报告(2020)》^[9]。该类型数据具备覆盖国别较广、数据条目与指标设置中三四级指标高度契合的特征,为指数编制提供了有力的支持。我们还可以使用以上部分数据进行国家间的关系研究,例如,基于专利和论文的引用状况数据,可对技术创新方面的国家间关系加以刻画。

以上数据构成了我们进一步开展国际数字生态研究、构建国际数字生态指数,以及判断各国在数字化发展领域上互动关系的基础。

四、国际数字生态指数的测算方法

(一) 数据预处理

在对国际数字生态指数进行测算之前,我们通过以下三种方式对数据进行了预处理。

1. 缺失值插补

针对部分国家缺失最新数据的情况,首先采用冷卡插补法,收集该国在缺失指标上的既往数据进行填充;如果该国既往数据仍然缺失,则使用平均值插补法。

2. 部分数据对数化处理

由于一些反映对象规模总量的数据(如网民数量)的绝对数值大、各国之间的差异大,导致这部分数据呈现偏态分布。本研究对以上数据进行对数化处理,以令其更接近正态分布的情况,便于后续的指数计算。

3. 标准化

在计算指标权重和指标聚合之前,采取“最小-最大”(Min-Max)标准化方法,将除数据规制^{注3}部分外的底层数据统一转化到[10,100]区间,作为对四级指标的测度,公式如下所示:

$$X_j = \frac{X_j - \min_i X_i}{\max_i X_i - \min_i X_i} \times 90 + 10$$

(二) 指标权重计算方法

我们主要使用加权聚合的方式对国际数字生态的总指数得分进行测算。在一级指标以下的加权聚合中,主要使用自编码器法(Autoencoder)和熵值法进行指标权重计算。最终的总指数得分通过对数字基础、数字能力、数字应用和数字规制这四个一级指标得分等权加总得到。

1. 自编码器法

自编码器法是一种能对数据实现降维的无监督的机器学习算法,其基本原理为:将输入的高维度数据利用一个神经网络编码器得到一组新的低维度数据,再利用一个神经网络解码器输出一组高维度数据,力图重构原始输入(参见图2)。通过最小化输入数据与输出数据之间的重构误差来训练模型,进而能得到图中神经网络各连边上的权重,中间得到的编码可以看作是输入数据的主要特征。研究对数字基础、数字能力和数字应用三个一级指

注1:由于“全球数据规制诊断”项目中包含了我们关注的部分国家,因此这里实际上的编码工作仅针对那些未包含在该项目中的国家进行。

注2:江苏佰腾科技有限公司是国内一家专业从事知识产权服务的高新技术企业,研究团队与该公司合作获取了与数字技术创新专利状况数据。

注3:数据规制数据由于是经过专业人员打分,故不对分布进行调整。

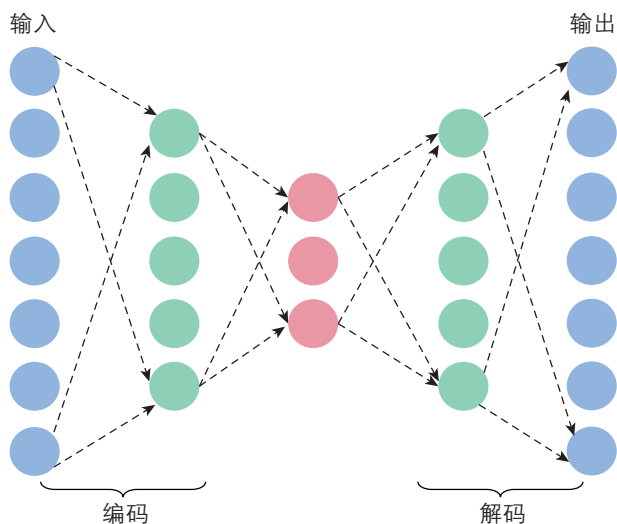


图2 自编码器示意

标使用自编码器法计算权重和得分。

该算法模型可以运用到对指标权重的计算中来。根据国际数字生态指数指标体系设置(参见表1), 我们可以将具体测量指标作为自编码器中的输入数据, 将三级指标、二级指标和一级指标作为神经网络的隐藏层, 根据指标间层级关系结构建立神经网络中的连边。这样, 由模型训练得到神经网络中连边上的权重即可作为指标聚合时使用的权重, 由自编码器得到的数据降维结果(即图1中的“编码”部分), 即可作为对应一级指标的得分。

自编码器作为机器学习领域的经典算法, 过去多用于自然语言处理和图像处理, 这里创造性地将其运用在对指标权重的计算之中, 据我们文献调研, 这在指数构建的研究中属首例。

2. 熵值法

与自编码器法不同, 熵值法是一种自动计算指标权重的常用方法, 它依靠数据分布的离散程度来确定指标权重的大小。在信息论中, 熵是对概率分布不确定性的一种度量, 对于随机变量 X , 其概率分布为 $p(x)$, 定义熵如下:

$$H(X) = E[-\log p(x)]$$

具体而言, 在计算指标聚合的权重时, 需将经归一化处理的测量指标转换为一个离散概率分布:

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}}$$

其中, x_{ij} 为个案 i (本研究中为国家)在指标 j 上的水平, 它是指标原始取值经标准化处理后的结果。 p_{ij} 是对它们进行归一化得到的结果, 进而我们能将其看作 $[0, 1]$ 区间上的一个离散分布。如果该指标的分布比较集中, 即意味着其区分性不强, 该概率分布的信息量较大, 不确定性较小, 熵比较小; 反之, 如果该指标的分布较为分散, 则其区分性较大, 不确定性很大, 熵比较大。根据熵的这一特性, 我们计算指标分布的熵值来判断此指标的离散程度 e_j :

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n p_{ij} \ln p_{ij}$$

更进一步, 再将各个指标的熵值转化为权重, 指标的区分度越高的, 则该指标的权重更大:

$$W_j = \frac{1 - e_j}{\sum_{j=1}^m (1 - e_j)}$$

在本研究中, 我们使用熵值法确定了数字规制部分的各层级指标权重。

3. 以等权方式生成国际数字生态总指数

在使用自编码器法和熵值法完成对数字基础、数字能力、数字应用和数字规制这四个维度下的指标权重计算, 形成四个一级指标得分之后, 以等权的方式, 即每个一级指标的权重均为0.25, 生成国际数字生态的总指数, 以此来综合判断各国数字生态的发展现状。

五、数字生态的国家排名与发展现状

利用上述方法, 我们首先测算了41国的数字生态指数, 表2显示了各国排名。中国的数字生态发展排在第7位, 排名前六的国家分别是英国、美国、德国、法国、荷兰和瑞典。

从一级指标上看, 中国在数字基础和数字应用方面有较明显优势, 排名均位于前五, 数字能力方面位居中

表2 41个国家的国际数字生态指数测算结果及国家排名

排名	国家名称	排名	国家名称	排名	国家名称
1	英国	15	意大利	29	巴西
2	美国	16	丹麦	30	捷克
3	德国	17	墨西哥	31	立陶宛
4	法国	18	爱沙尼亚	32	希腊
5	荷兰	19	奥地利	33	印度
6	瑞典	20	葡萄牙	34	匈牙利
7	中国	21	罗马尼亚	35	克罗地亚
8	西班牙	22	斯洛文尼亚	36	马耳他
9	加拿大	23	斯洛伐克	37	以色列
10	韩国	24	俄罗斯联邦	38	日本
11	澳大利亚	25	乌克兰	39	塞浦路斯
12	波兰	26	比利时	40	拉脱维亚
13	芬兰	27	卢森堡	41	伊朗
14	爱尔兰	28	保加利亚	—	—

上游，数字规制方面的排名则相对偏后。

在数字基础方面，中国位居第四，美国、英国和加拿大分别位居前三，从得分上看，美国的领先优势较为明显（参见图3）。

在数字能力方面，中国位居第十，排在中国前面的国家包括美国、英国、德国、以色列、瑞典、芬兰、加拿大、韩国和俄罗斯，这说明中国在数字能力建设上还有很大的赶超空间（参见图4）。

在数字应用方面，中国仅落后于美国，位居第二，其余排名前五的国家还包括英国、韩国及荷兰（参见图5）。若对数字应用下设的二级指标得分做进一步细致的比较分析，能够发现中国在数字经济方面的表现更为突出，得分仅次于美国；在数字政府方面的排名也居于

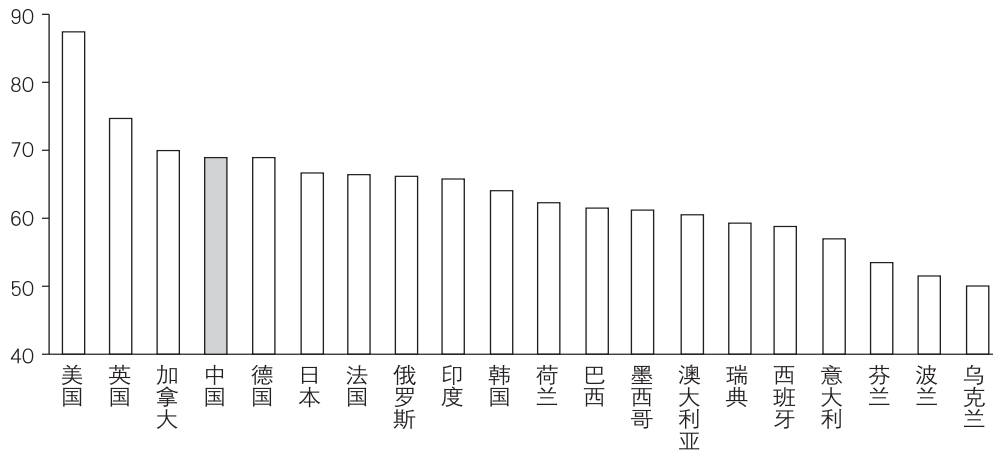


图3 数字基础得分排在前20的国家

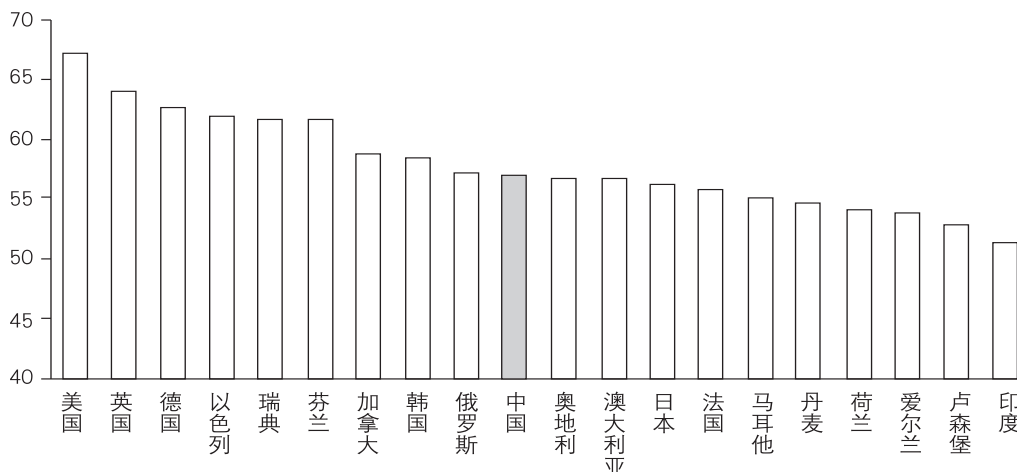


图4 数字能力得分排在前20的国家

乔天宇 张蕴洁 李铮 赵越 邱泽奇 · 国际数字生态指数的测算与分析

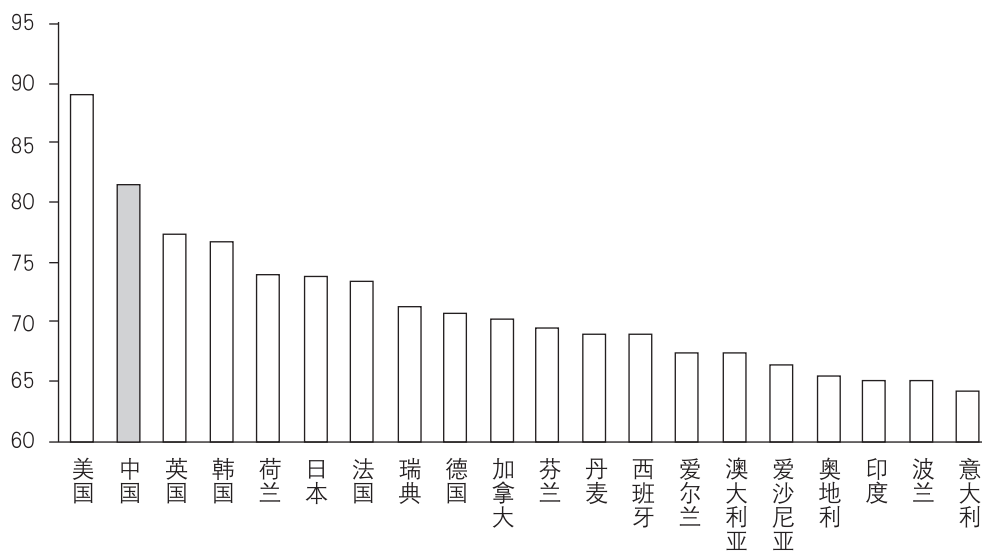


图5 数字应用得分排在前20的国家

上游；但在数字社会方面，中国居第15名，其中数字教育仍有很大发展空间。

在数字规制方面，排名前20的国家中有19个是欧洲国家（参见图6）。欧洲各国在文化传统上重视保护个人隐私，对数字规制的制定起步较早，规制数量总体偏多，约束性也偏强。另外，由于欧盟层面对数据保护有整体性立法，各国在实践中也主要以贯彻欧盟立法为主，因此在数字规制上表现出极大的同质性。中国在数字规制上排第33位，美国则更靠后，排在了第37位，这

在一定程度上反映出中美两国的数字规制与欧盟相比完备性较低。但数字规制的完备也并不必然会促进其他数字生态要素的发展，对此需要具体分析，将在后文中予以讨论。

我们采用雷达图，进一步展示了中、美、英、德、法、俄六国在二级指标上的表现（参见图7）。从雷达图形状来看，英、德、法三国的的发展状况总体上比较相似，尤其数字商务、数字流通和数字风控等方面的规制完备，但这三国在数字经济方面较中、美有一定差距；

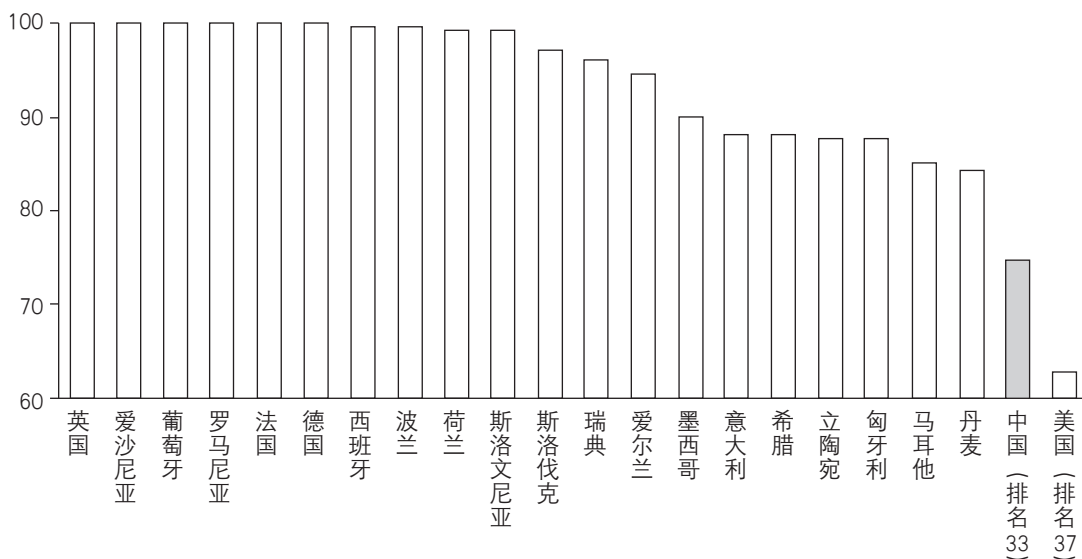


图6 数字规制得分排在前20的国家及中国、美国

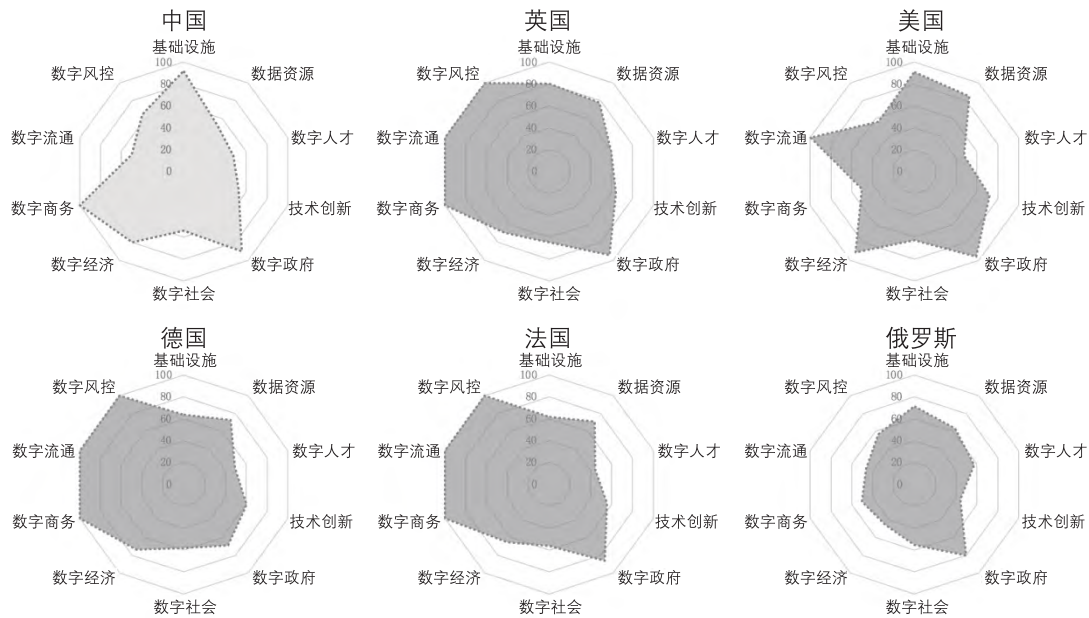


图7 中、美、英、德、法、俄的二级指标得分雷达图

美国在数据资源、数字经济和数字政府上都具备很高的得分，而在数字商务和数字风控上的得分较低；俄罗斯的数字生态发展在多个维度上落后于图中列出的其他国家。中国在基础设施和数字经济上的得分较高，仅落后于美国，较欧洲各国具备优势；在数据资源开放上，中国有明显短板。

总的来看，国际数字生态指数的计算结果反映出中国在数字化发展上具有一定优势，尤其体现在数字基础和数字应用方面。利用数字生态指数，我们除了能从多个维度对各国数字化发展特点作系统评价、识别国家间差异外，还有必要通过分析各要素间的关联性，揭示数字生态发展的结构模式。同样，各国在结构模式上也可能存在差异。

六、数字生态发展的结构模式

对国际数字生态发展结构模式的分析，主要围绕基础设施、数据资源、数字人才、技术创新、数字政府、数字社会、数字经济、数字商务、数字流通和数字风控这10个二级指标的测算结果展开。首先分析这10个二级

指标之间的相关关系，图7呈现了它们的得分分布、相关系数矩阵以及统计检验结果。

从全部41个国家来看，首先，数字应用的发展与数字基础息息相关。数字应用的三个二级指标与数字基础的两个二级指标得到的六对相关关系均显著为正（参见图8），这说明它们之间是相互促进、共同发展的。其次，数字应用的三个二级指标与数字能力下的技术创新也均呈现出显著的正相关。但数字能力下的数字人才指标，只与数字政府和数字经济呈现出显著的正相关性，与数字社会的相关性较弱。再次，从数字应用下的三个二级指标间关系来看，数字政府与数字社会显著相关（相关系数为0.62， $p < 0.001$ ），但数字经济却与数字政府、数字社会之间的相关性较弱，且统计上不显著（相关系数分别为0.23和0.20， p 值分别为0.143和0.215），这说明数字经济与数字政府、数字社会不一定是同步发展的。最后，数字规制下设的三个二级指标之间关联紧密，但它们与其他二级指标间的相关性均较弱，且多数呈现负向关系。其中，数字商务与数据资源、数字人才、数字经济的负相关，主要是由于绝大部

乔天宇 张蕴洁 李铮 赵越 邱泽奇·国际数字生态指数的测算与分析

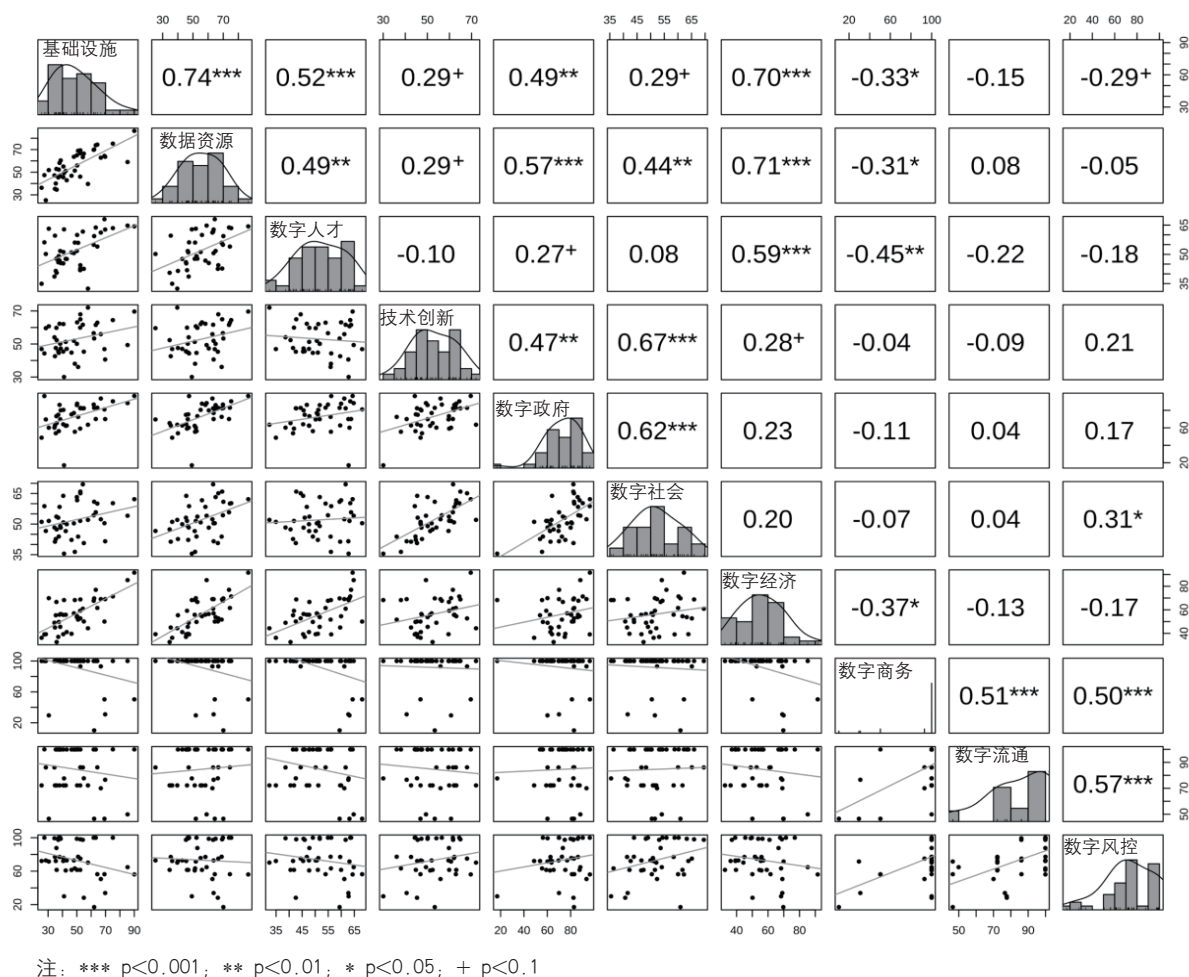


图8 国际数字生态指数中十个二级指标之间的相关系数矩阵

分国家的数字商务得分较高且差异性较小,受美国等少数国家的极端案例影响所致^{注4},这一负向关系并不意味着数字商务法规抑制了以上要素的发展。

我们这里对数字生态各要素所作的关联性分析,类似于对生态系统要素间耦合状况的考察。但是以上分析假定41个国家共享同一种数字生态的结构模式,并没有划分出不同的数字生态类型,进而无法反映不同的数字生态结构模式。但实际上,各国在数字生态发展上可能会表现出不同的耦合状况,即存在差异性的结构模式。为揭示这种差异,我们利用关系性类别分析的方法

(relational class analysis,简称RCA)^[10-11]进行了探索。RCA同时考察变量和国家之间的关系,将变量间逻辑关系体现出一定相似性的国家聚合在一起,进而揭示数字生态要素间关联的不同结构模式。我们将其看作是不同类型生态系统耦合状况的体现。

经RCA模型计算,研究考察的41国可划分为两种数字生态发展的结构模式。我们以关系网络图的形式展现了这两种结构模式(参见图9、图10)。图9、图10中节点表示指数的二级指标,若两个二级指标之间的相关系数在0.1的水平下统计显著,则在表示这两个二级指标

注4: 这些国家包括美国、日本、俄罗斯、印度和以色列。整体上看,它们在基础设施、数据资源、数字人才和数字经济等方面存在一定发展优势,但数字商务规制的发展较不完备。

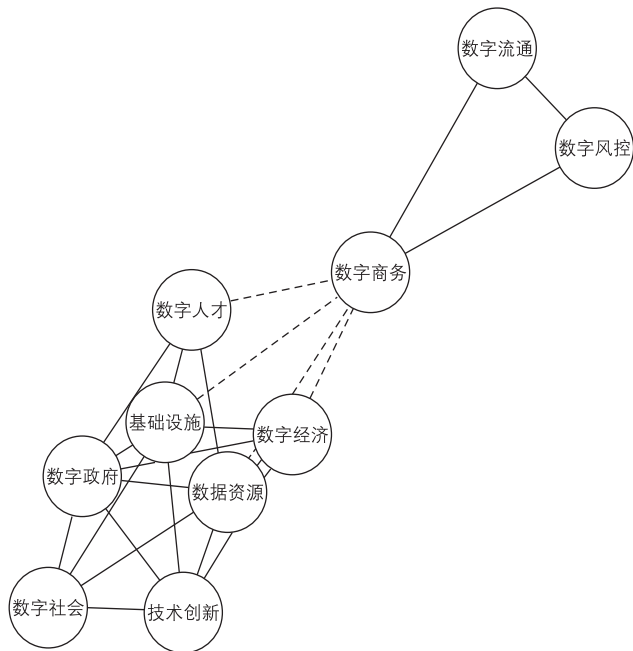


图9 紧耦合的数字生态发展结构模式

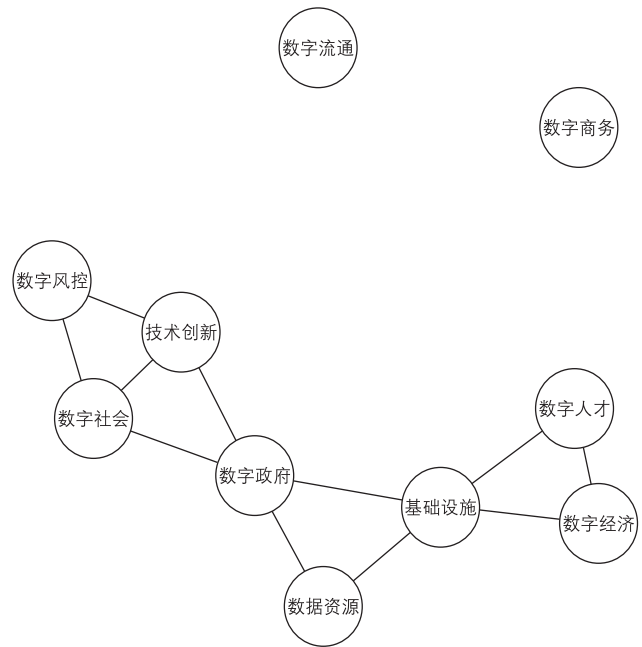


图10 松耦合的数字生态发展结构模式

的节点之间连一条边，实线表示二者呈正相关，虚线表示二者呈负相关。

这两幅图直观反映出两种数字生态发展结构模式间的显著区别。对于第一种模式，数字基础、数字能力和数字应用下的二级指标紧密地耦合在一起，数字规制下的数字商务、数字流通与数字风控三个二级指标也相互关联；相比较而言，第二种模式中的各二级指标间的联系较为稀疏，数字规制下的三个二级指标间也互不关联。

在第一种紧耦合模式中，我们看到数字风控与技术创新和数字社会呈显著正向关系。数字风控本是对规制中约束性法规的考察，但这里的正向关系或许并非意味着数字风控对这两方面的发展具有促进性作用，而是具备较高技术创新和数字社会发展水平的国家有可能率先出台了数字风控的相关规制。

我们考察的41国都分别属于哪种数字生态发展结构模式，中国又属于其中哪种类型呢？其中，属于第一种模式的国家共有24个，美、英、德、法都属于这一模式，欧盟的其余绝大多数国家都被划入此类。属于第二

种模式的国家共有15个，包括韩国、加拿大、巴西、伊朗、墨西哥，以及意大利、奥地利、比利时等少数欧盟国家。中国属于第二种模式，这说明与美欧等多数发达国家相比，中国数字生态发展呈现出一种松散的结构，各要素间的耦合性并不紧密。在本文的最后一部分中，我们将对此具有的潜在意义予以讨论。另外，俄罗斯和以色列未被划分进这两种模式中的任意一种。

七、利用国家间关系网络刻画数字生态

如果说此前的分析还主要是从生态系统的内涵下理解数字生态，本部分则尝试从群落生态的角度来刻画数字生态。群落生态学主要研究物种间的相互作用，关注诸如竞争、利用与合作共生等关系类型。在这里，我们将国家类比为物种，希望判断以上关系类型是否存在于国家间的数字化发展格局之中。因此，在数据收集时，我们尤其注重那些能够直接反映国家间互动或交换关系的内容，如数字产品与服务的国家间贸易、数字人才的国家间流动、数字技术专利的国家间引用等数据。在本部分中，我们将用一个示例来说明，如何利用这些国家

间关系网络数据来刻画国际数字生态。

这里使用的是移动电话这一具体数字技术产品的国家间贸易数据，数据来源于联合国国际贸易数据库(UN Comtrade Database)，在国际数字生态研究框架中，该数据属于数字经济维度，用于考察数字贸易。利用2020年各国出口移动电话的产值信息，构建了一个国家间移动电话贸易网络。我们将描述各国在该网络结构中所处的位置特征，并尝试利用网络结构提供的信息，对国家间的关系类型给出判断(参见图11)。

图11是对41个国家间移动电话贸易网络做的可视化呈现。其中，节点的颜色表示我们根据网络结构提供的信息对41国所处生态位划分的结果。很多从事社会网络研究的学者都认为可根据结构等价(structural equivalence)来衡量生态位，并结合生态学的“竞争排斥”原理，可认为处于同一生态位上的行动主体间存在潜在竞争关系。^[12-13]此处的划分结果显示，中国、美国、加拿大、巴西、日本、韩国等六个国家处在同一个生态位上，意味着在移动电话贸易领域，这些国家间可能存在着一一定程度的竞争。从图中还可看出，该网络具有较明显的“核心-边缘”结构特征。我们利用博尔加

蒂等^[14]发明的“核心-边缘”算法模型对这一观察做了严格的验证。离散模型发现，在这41个国家中，中美两国处于绝对的核心位置，其他39个国家都处于边缘，观测的网络结构与模型得到的理想“核心-边缘”结构之间有良好的拟合($p=0.008$)。

在对国家间移动电话贸易网络进行初步描绘后，我们利用美国社会学者伯特研究网络中结构洞的方法，对各国在该贸易网络中受到的结构性约束(structural constraint)予以测量。伯特^[12]认为，一个节点在网络中受到的结构性约束越低，它可能会跨越更多的结构洞，从而在网络中占据优势位置。表3包含各国基于上述方法计算的结构性约束指数和国家排名。总体来看，中国在该网络中受到的约束最低，由此可以推断，在移动电话贸易中，中国占据着相当的优势位置。在我们关注的欧美主要国家中，德国与中国类似，在该网络中受到的约束较低，而美国和英国受到的约束则相对较高，分别排在第10和第16位；41个国家中受约束最高的5个国家分别是墨西哥、加拿大、日本、巴西和澳大利亚。

进一步使用伯特提供的方法考察两两国家间彼此约束的状况，以处于核心位置的中美两国为例，分析它们

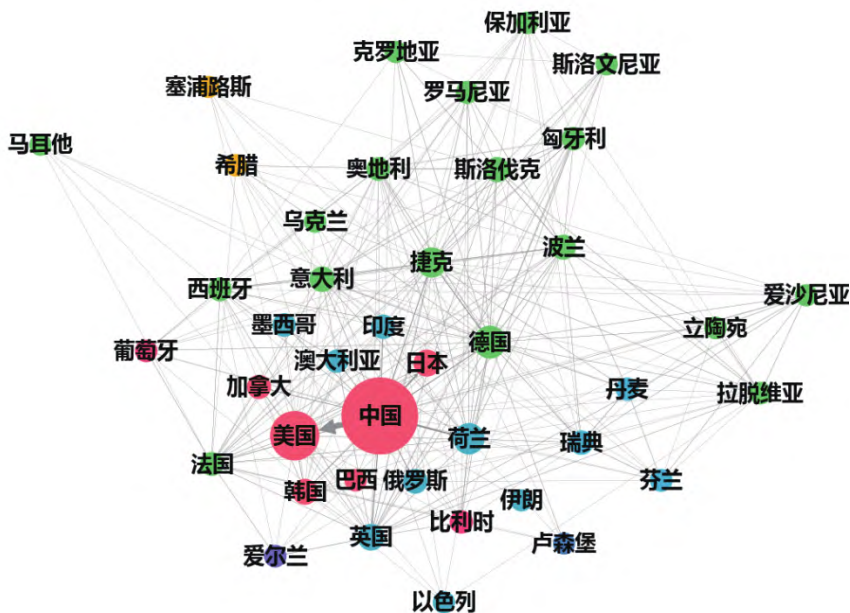


图11 41个国家间的移动电话贸易网络

表3 41个国家在移动电话贸易网络中受到的结构性约束程度

国家	结构性约束	排名	国家	结构性约束	排名
墨西哥	1.22	1	西班牙	0.46	22
加拿大	1.22	2	罗马尼亚	0.46	23
日本	1.20	3	意大利	0.46	24
巴西	1.18	4	塞浦路斯	0.44	25
澳大利亚	1.17	5	奥地利	0.44	26
俄罗斯	1.09	6	克罗地亚	0.43	27
韩国	1.05	7	保加利亚	0.43	28
伊朗	1.00	8	法国	0.42	29
以色列	0.98	9	捷克	0.42	30
美国	0.90	10	波兰	0.41	31
爱尔兰	0.85	11	希腊	0.40	32
印度	0.74	12	立陶宛	0.40	33
乌克兰	0.74	13	斯洛伐克	0.37	34
比利时	0.73	14	瑞士	0.35	35
丹麦	0.71	15	拉脱维亚	0.35	36
英国	0.69	16	卢森堡	0.35	37
荷兰	0.67	17	德国	0.35	38
芬兰	0.53	18	马耳他	0.32	39
斯洛文尼亚	0.48	19	爱沙尼亚	0.31	40
匈牙利	0.46	20	中国	0.26	41
葡萄牙	0.46	21	—	—	—

受到的来自其他国家的约束。研究发现，该网络中对中国构成约束最高的五个国家分别是美国、荷兰、日本、德国、英国，中国向这些国家出口（包括直接和间接）移动电话的贸易额相对其他国家更大；而对美国构成约束最高的五个国家分别是中国、韩国、加拿大、荷兰和日本，其中中国对美国构成的约束远远超过其他国家，

可以认为，美国的移动电话海外市场处在一种高度依赖中国的状态。图12中虚线和实线之间所夹的面积可表示一种结构洞信号，面积越大代表主体可以利用的结构洞越多^[12]。很明显，中国的结构洞信号要强于美国。图中的宽幅地带表示机会型关系，如图12（左）中，美国对于中国来说就属于这种类型，伯特认为在此类关系中，主体（此处指中国）“有充分的谈判余地，并因此能够控制局面”。高而窄的地带表示受到约束的关系，图12（右）中，中国对美国而言就属于典型的这类关系，美国的移动电话出口高度依赖中国，但同时并没有可供其利用的结构洞，美国在与中国的移动电话贸易互动上谈判余地很小，处在一种受制于中国的劣势位置。

这一研究示例展示了如何在国际数字生态的框架下，通过搭建国家间互动关系网络和根据网络结构信息，对特定国家在某一数字产品国际交易场域中所处的位置，以及特定国家间的关系性质给出判断。对中美两国而言，尽管二者均位于移动电话贸易网络的核心，但又由于它们同处在一个生态位上，相互间可能存在潜在竞争。不过相对美国来说，中国在该贸易网络中具有一定的相对优势，表现为受到的结构性约束较小，且机会空间较大。

八、总结与展望

数字化发展已经成为影响国际发展新格局的重要变量。在当前国际形势发生新变化的背景下，如何针对数

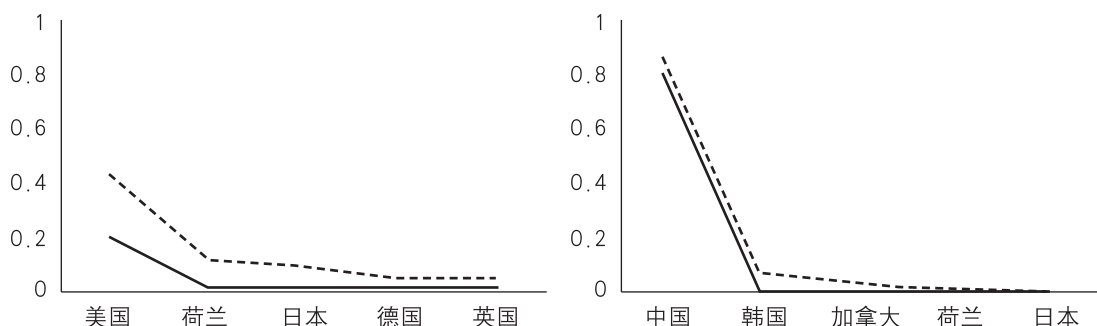


图12 中国（左）和美国（右）在移动电话贸易网络中的结构洞信号

字化发展格局给出刻画和判断,是亟待破题并回答的时代性议题。本文从数字生态的角度对此进行了初步研究,从生态视角上简要介绍了我们对国际数字生态的理解,并提出了刻画国际数字生态的基本框架。该框架由数字基础、数字能力、数字应用和数字规制四个维度构成。

在当前阶段,国际数字生态的研究包含以下两方面内容:第一,在上述基本框架指引下,收集相关数据,构建并测算了国际数字生态指数。根据测算结果,我们能得到有关数字化发展状况的国家间排名,这种研究方式与目前很多国内外机构开发的评估各国数字化发展的指数工具相类似,不同之处在于,本研究增加了对数字规制维度的考察。第二,我们通过研究数字化发展领域的国家间相互作用网络,从群落生态内涵的角度对国际数字生态加以刻画,并提供了一个基于国家间数字贸易网络的研究示例。根据网络结构提供的信息,研究能够考察各国在关系场域中所受结构性约束的状况,划分生态位,并基于此给出对国家间关系性质的判断。

从国际数字生态指数测算结果来看,中国的总指数得分在41个国家中位居第七,可以认为,中国整体上已经处在世界数字化发展的头部梯队中。但中国目前在数字能力上的排名仍较靠后,通过培养数字人才、发展数字技术创新,进而提升能力仍旧是第一要务。同时,对于中国而言,构成数字生态系统的各要素之间耦合程度较弱,联系相对稀疏,与美国和英、法、德等欧洲主要发达国家相比,在数字生态发展的结构模式上还有实质性的差别。

从目前世界上大国的数字生态发展状况来看,数字规制中的数字流通和数字风控与数据资源、数字人才、技术创新和数字经济中的一些指标呈现出了负相关性。这在一定程度上表明数字规制的约束强度越高,数字能力发展受到抑制的可能性就越大,数字经济的发展也越可能受到影响。以中美欧举例来看,欧盟国家颁布的数

字规制在约束性强度上最高,但欧洲各主要国家的数字产业化发展水平整体上不如美国和中国,数字经济尤其是电子商务、数字平台、共享经济等新兴业态的国际竞争能力已经落后,这一定程度上是由其约束过度的数字规制导致的。

数字生态的发展要在数字能力与数字规制之间寻求平衡、适合的路径。数字规制可能是柄“双刃剑”。一方面,数字规制可以从源头上促进数字资源的开放流通、便利交易;还可以通过明确各参与者的权利义务,从而促进交易的达成及信任的建构,对于治理数字化发展中可能产生各类问题也十分重要。但另一方面,就像不合适的生产关系会制约生产力的发展那样,约束不当、尤其是过度约束的数字规制也可能会影响数字生态的良性发展。

根据我们对国际数字生态的研究,中国目前在数字能力上的排名仍较靠后。通过培养数字人才、发展数字技术创新,进而提升能力仍旧是第一要务。从世界范围看,目前有关数字化发展的实践远远走到了理论前面。近十到二十年间,数字产业突飞猛进,数字技术创新层出不穷,但相关的理论研究才刚刚起步,我们还面对很多空白领域有待认知,很多争议有待讨论。鉴于数字规制与数字能力提升之间可能存在的负向关系,对于那些认识上尚未完备或当下仍存在争议和分歧的领域,可以考虑暂缓制定相应的规制,采取包容审慎监管的策略,允许试验性发展,给相应的规制留足“变通”运行空间,以避免约束性过高的数字规制不当抑制数字能力提升,最终影响到数字生态的健康良性发展。当下,发展数字能力是中国数字生态建设的首要目标,这是决定中国能否在世界百年未有之大变局的背景下,抓住新一轮科技革命带来的新发展机遇,在世界各大国的竞争博弈中最终胜出的关键所在。

当然,以上还都是国际数字生态研究的阶段性成果,相关研究工作都刚刚起步,难免存在局限和不足。

今后,我们将会在国际数字生态的理论框架下,继续丰富研究数据源,并利用已收集到的数据,开展更多的研究工作,尤其是从群落生态视角出发,利用国家间相互作用关系网络对国际数字生态进行刻画的研究。这类研究对于中国进一步明确数字化发展方向,在新发展格局下制定国家战略,推动国内国际双循环相互促进,都将具有重要的指导意义。

参考文献:

- [1]European Commission. Shaping Europe's Digital Future[R]. Brussels: European Commission, 2020: 2.
- [2]Siddharth M, Kristine L, Joshua F. Designing a U.S. Digital Development Strategy[R]. Washington DC: Center For A New American Security, 2020: 1.
- [3]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[S/OL]. (2021-03-13)[2021-12-10]. http://www.gov.cn/xinwen/2021-03/13/content_5592681.htm.
- [4]王娟,张一,黄晶,等.中国数字生态指数的测算与分析[J].电子政务,2022(03):4-16.
- [5]徐清源,单志广,马潮江.国内外数字经济测度指标体系研究综述[J].调研世界,2018(11):52-58.
- [6]The World Wide Web Foundation. Open Data Barometer 4th Edition — Global Report[R]. Washington DC: The World Wide Web Foundation, 2017: 6-9.
- [7]Open Data Watch. Open Data Inventory 2020/21[EB/OL]. (2021-02-01)[2022-01-10]. <https://odin.opendatawatch.com/Report/annualReport2020>.
- [8]Chen R. Mapping data governance legal frameworks around the world: Findings from the global data regulation diagnostic[R]. Washington DC: World Bank Policy Research Working Paper, 2021: 35-44.
- [9]清华大学经济管理学院互联网发展与治理研究中心,领英.全球数字人才发展年度报告(2020)[R].北京:清华大学,领英,2020:4-20.
- [10]Goldberg A. Mapping shared understandings using relational class analysis: The case of the cultural omnivore reexamined[J]. American Journal of Sociology, 2011, 116(05): 1397-1436.
- [11]Boutyline A. Improving the measurement of shared cultural schemas with correlational class analysis: Theory and method[J]. Sociological Science, 2017, 4(15): 353-393.
- [12]Burt R S. Structural holes: The social structure of competition[M]. Cambridge: Harvard University Press, 1992: 209.
- [13]奇达夫 M, 蔡文彬. 社会网络与组织[M]. 王凤彬, 朱超威, 译. 北京: 中国人民大学出版社, 2007: 71.
- [14]Borgatti S P, Everett M G. Models of core/periphery structures[J]. Social Networks, 2000, 21(04): 375-395.

作者简介:

乔天宇,北京大学大数据分析与应用技术国家工程实验室博士后,研究方向为技术社会学、组织社会学、计算社会学。

张蕴洁,北京大学社会学系博士研究生,研究方向为技术社会学、计算社会学。

李铮,北京大学社会学系硕士研究生,研究方向为技术社会学、计算社会学。

赵越,北京大学工学院博士研究生,研究方向为机器学习、数据价值、数字经济。

邱泽奇,北京大学社会学系教授,博士生导师,北京大学中国社会与发展研究中心主任,北京大学数字治理研究中心主任,研究方向为技术应用社会变迁、组织社会学、社会研究方法。