中国引领科学趋势；西方开辟新趋势

各国引领科学的方式各不相同：有的推动既有趋势，有的则与之背道而驰，而这会对技术竞争产生持久影响。

作者： Jeffrey W. Lockhart[1,2], Jamshid Sourati[2,3], Feng Shi[2,4], James Evans[4,5,6]*

[1] 社会学系，加州大学伯克利分校，CA 94720。[2] 芝加哥大学知识实验室，5735 South Ellis Avenue，芝加哥，IL 60637。[3] 德保罗大学计算机学院，芝加哥，IL 60614。[4] 芝加哥大学社会学系，1126 E 59th St，芝加哥，IL 60637。[5] 圣塔菲研究所，1399 Hyde Park Road，圣塔菲，NM，87501。

国家如何塑造科学前沿对技术竞争至关重要，但包括论文数量、引文和颠覆性指数在内的标准指标都是回顾性的，无法区分根本不同的领导策略。我们开发了两项前瞻性的基于模型的测度，并将其应用于自1990年以来的数千万篇文章。第一项将研究路径嵌入由概念和科学家构成的不断演化的超图中，以识别在新兴领域的领先地位——即预判科学群体走向的研究。第二项将以往研究中所借鉴的思想和学科的演化样本嵌入其中，以识别在令人惊讶的新方向上的领先地位，即当意外组合变得常规、科学围绕它们重组之时。大约十年前，中国成为新兴领域的全球领导者，远早于其在产出量上领先，这反映出其大规模察觉并放大初步共识的能力。美国和欧洲则呈现相反特征：涌现份额下降，但在具有前瞻性的研究上持续领先，尤其是跨越学科边界的研究。这些模式在不同数据库、归因方法和战略领域中均可复现，包括人工智能、生物技术、能源和半导体。国家引领科学的方式，或是解读格局，或是重塑格局，而两种策略所需的制度条件彼此存在张力。这些策略的分布有望在未来数十年塑造技术领导力的全球结构。

* 通讯作者。电子邮箱：jevans@uchicago.edu (J.E.)

科学技术发展作为全球性的集体现象展开。新的思想和工具作为公共产品不断积累并扩散（*1–3*），在一波波创造性破坏中取代先前的方法（*4–6*）。然而，科学资助不成比例地发生在国家内部（（*7–9*）），其理由是在较短时间尺度上对产业或军事优势的贡献。这种张力使人们将关注度集中于全球科学中的国家领导力，将其作为经济和军事实力的代理指标（*10–12*）。但包括论文数量、科学家人数和引文在内的标准指标所反映的是过去而非未来的领先地位。

我们提出两种互补的领先地位衡量方法，用以捕捉各国如何塑造科学发展轨迹（*13*，*14*）。第一种方法识别那些引领集体关注度转向新兴领域的研究。这类科学研究能够预见并推动全球科学界的势头。第二种方法识别那些偏离可预测发展轨迹的先见性进展，它们作出出人意料的贡献，并预示后续研究方向。这类科学研究转向意想不到的可能性，随后各学科领域也会跟进。

我们将这些衡量方法应用于自1990年以来数千万篇研究论文，发现十年前中国就已成为新兴科学领域的全球领军者，远早于其在出版物数量上取得领先，这表明中国采取了有意放大初步共识的战略。然而，中国、日本和韩国在先见性进展方面明显落后。美国和欧洲则呈现相反的趋势：它们在新兴科学发展轨迹中的份额不断下降，但在不可预测的突破方面始终保持领先地位。

为将这些维度转化为可操作的指标，我们评估各国在所有科学领域及关键战略领域（包括人工智能、生物技术、能源和半导体）中的贡献。我们首先通过将研究路径嵌入由思想与科学家构成、不断演变的超图中，识别各国在新兴领域的领先地位。在此前的研究中，该模型能够准确预测各个科学领域未来的发现及其发现者（*15*）。它揭示了科学概念之间的“构造运动”，并预测组成思想相互催化时新领域的涌现。我们识别出研究思想汇聚、相关研究的增长加速且关注度已相当高的新兴领域。由于科学家自然地处于这一嵌入空间中，我们可以根据各国科学家在最接近新兴领域核心的科学家群体中所占的比例，判断哪些国家有望引领这些领域。在新兴领域领先的国家顺应趋势，以预判科学群体的动向。

我们的第二种方法通过避开人群来识别领先地位。该方法将每种概念组合投射到不断演化的科学工作嵌入中，以评估其出现的概率。既有研究表明，出人意料的论文更有可能产生高影响力（*16*）。我们将这一方法扩展用于识别哪些国家能够产出具有先见性的论文——这类工作在发表时令人意外，却会在未来研究中成为常规。涌现体现的是在可预测将会汇聚的领域中的领先地位，而先见性体现的是将科学引向意外方向的领先地位（图 1）。

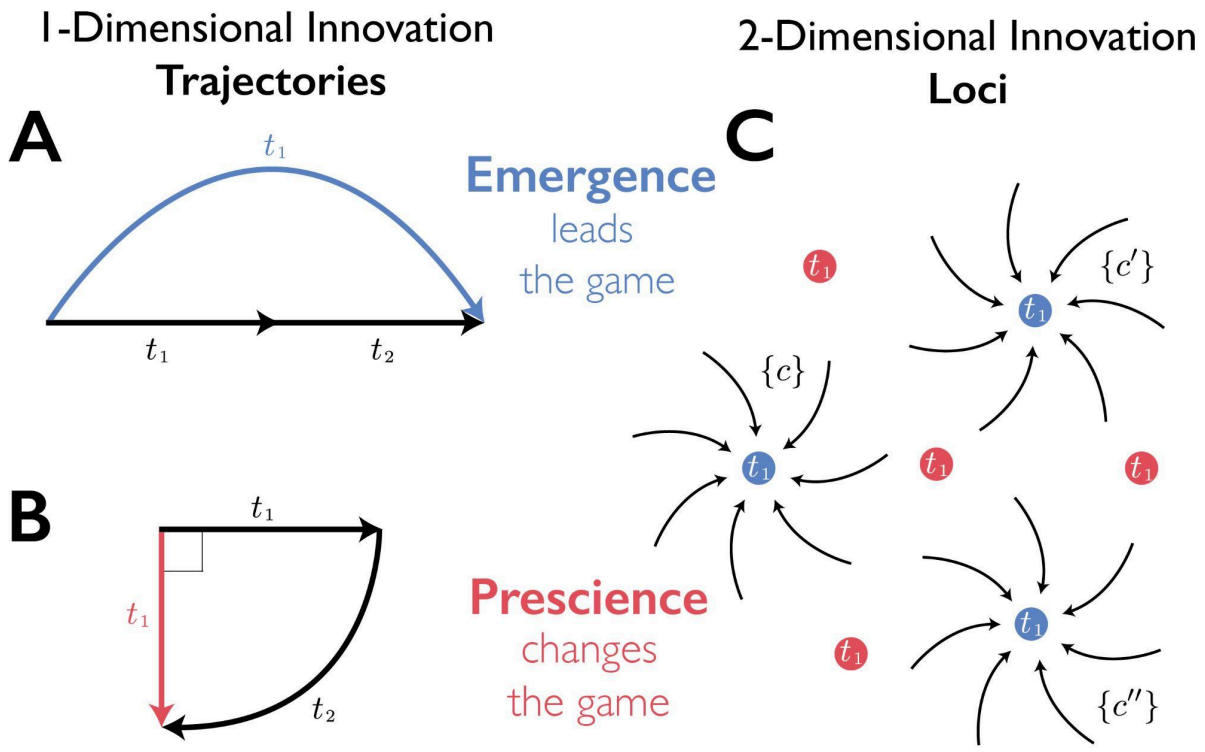


图1. 创新的涌现与先见性方法之间的概念对比。A-B，一维创新轨迹，其中黑色箭头表示创新在时间1和时间2的众数方向。时间1的创新涌现方法通过跳向正在接近的科学前沿，引领时间2的发展轨迹（**A**）。时间1的创新先见性方法则通过偏离原有轨迹来改变发展轨迹，而科学前沿改变方向，以与先见性创新相汇合（**B**）。**C**，时间1的二维创新轨迹点集。概念集合{c}、{c'}和{c"}分别在时间1趋于收敛。创新涌现方法预见这些收敛，而先见性方法则偏离这些收敛。在时间2，这些集合中的概念改变方向，以接近时间1的预见性贡献。

我们将这些面向未来的指标与既有指标进行对比：生产力（出版物数量）、关注度（引文）（*17*）以及颠覆（*18*，*19*）。对于颠覆，我们采用基于引用的颠覆性指数，用于量化一篇论文是否在引用网络中超越了先前的工作（*18–21*）。该指数反映了学界是通过引用前人的成果，将目标工作视为科学“大厦中的一块砖”，还是通过忽略前人的成果，将其视为新的基础。不同于从前瞻性角度考察那些引领新进展的工作，涌现关注的是未来可能出现的突破；颠覆则通过回溯性地识别下游对工作成果的认可，找出科学关注度发生转向的节点。

我们通过在两个数据库（Web of Science 和 OpenAlex；图 S4）中重复分析、改变论文—国家归属方法（图 S7），并在排除焦点国家后重新计算模型，以确认这些结果具有全球而非仅仅局部的重要性，从而验证了结果的稳健性。创新领域在全球范围内构成一个共享的知识共同体，而非局限于个别国家。各国在不同领域的投入虽不均衡，但没有任何国家的进展只局限在本国境内。一国的开创性工作会推动全球发展。

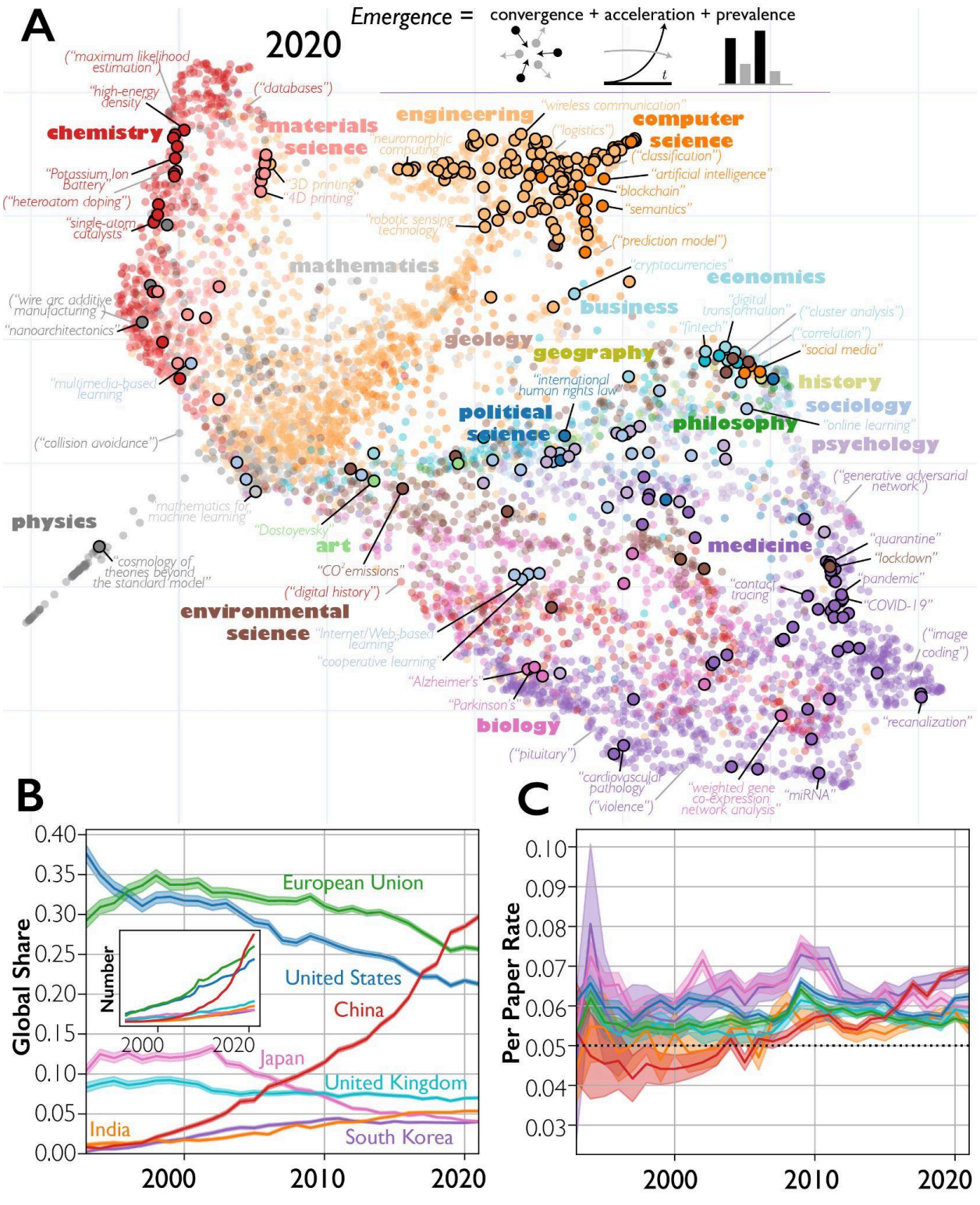
结果

涌现

在我们的模型中，当科学概念之间的余弦距离接近 0.1 时，它们在新研究中共同出现的似然会增加 800% 以上。这些涌现式关注的迸发界定了发现格局的转变。我们通过识别迅速汇聚的簇来捕捉这些变化。就其延续了发现格局中此前的演变轨迹而言，这些簇代表着预期的发展（见方法）。位于这些迸发区域中心的科学家，是最持续、最积极地参与在此发表发现与发明的群体。位于可能在不久的将来迸发的汇聚区域中的科学家，则有望引领这些领域的进展。

位于新兴研究领域核心区域附近的科学家，往往来自中国机构。与美国科学家相比，中国科学家涉足的涌现领域近乎多出三倍，且每个领域的科学家人数多出四倍。鉴于中国学术队伍迅速壮大，中国作者在新兴领域中的占比也在上升，这不足为奇。尽管如此，这一趋势出现的时间之早仍令人瞩目。中国作者人数在2016年接近美国、在2019年接近欧盟，但中国作者在新兴领域中的占比早在5年前就已超过这些地区。

图2. 涌现式创新。**A**，对2020年发表的论文进行随机游走，这些论文通过概念和作者相互连接；再将随机游走嵌入100维空间，并将概念投影到二维UMAP空间。对这些随机游走进行嵌入，可以准确模拟现有科学家在未来科学中最有可能建立的概念联系。每个概念点根据标注的学科领域着色（例如，政治学）；带黑色边框的点表示涌现领域，其定义依据右上角插图所示的标准：（1）邻近概念的汇聚；（2）包含该概念的论文数量加速增长；（3）包含该概念的论文数量达到阈值。部分涌现领域以不带括号的形式标注（例如，人工智能、新冠肺炎），邻近的非涌现领域则以带括号的形式标注（例如，分类、垂体）。**B**，计算1990年至2023年涌现式科学的全球占比，并将其归于在概念区域本身附近10位科学家中至少有一位、且其研究包含该概念的国家。图中仅纳入占比最高的七个持有份额的国家或地区，但所有国家或地区均列于图S4–S7中。美国和欧盟的全球涌现占比下降，而中国的占比大幅上升并取得领先地位。插图对比了全球占比与涌现式论文数量，后者在所有国家均有所增长。**C**，单篇论文涌现率显示，这种方法长期以来一直是日本和韩国创新战略的特点；但中国的单篇论文涌现领先地位在2016年超过美国，并于2018年超过任何其他国家或地区（其他计算方法和稳健性检验见图S4–S7）。


A
Emergence = convergence + acceleration + prevalence
2020
("maximum likelihood estimation")
"high-energy density"
("databases")
chemistry
materials science
engineering
"wireless communication"
("logistics")
computer science
"neuromorphic computing"
("classification")
"artificial intelligence"
"Potassium Ion Battery"
("heteroatom doping")
"3D printing"
"4D printing"
"blockchain"
"semantics"
"single-atom catalysts"
"robotic sensing technology"
("prediction model")
mathematics
"cryptocurrencies"
economics
("wire arc additive manufacturing")
"nanoarchitectonics"
business
"digital transformation"
("cluster analysis")
"fintech"
("correlation")
geology
geography
"social media"
"multimedia-based learning"
"international human rights law"
history
political science
philosophy
sociology
"online learning"
("collision avoidance")
psychology
("generative adversarial network")
"mathematics for machine learning"
physics
"Dostoyevsky"
art
medicine
"quarantine"
"lockdown"
"cosmology of theories beyond the standard model"
"CO2 emissions"
("digital history")
"contact tracing"
"pandemic"
"COVID-19"
environmental science
"Internet/Web-based learning"
("image coding")
"cooperative learning"
"Alzheimer's"
"Parkinson's"
biology
"recanalization"
("pituitary")
"cardiovascular pathology"
("violence")
"weighted gene co-expression network analysis"
"miRNA"
B
Global Share
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
European Union
United States
China
Japan
United Kingdom
India
South Korea
Number
2000
2020
2000
2010
2020
C
Per Paper Rate
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
2000
2010
2020

先见性

在我们区分可预测与不可预测的近期技术科学发现与发明的动态模型中，我们可以识别先见性论文，即那些“超前于时代”的论文。这些研究成果在发表时高度不可预测，但随着研究分布转向并围绕它们重组，后来显得高度可预测。实际上，未来研究追随它们的引领。

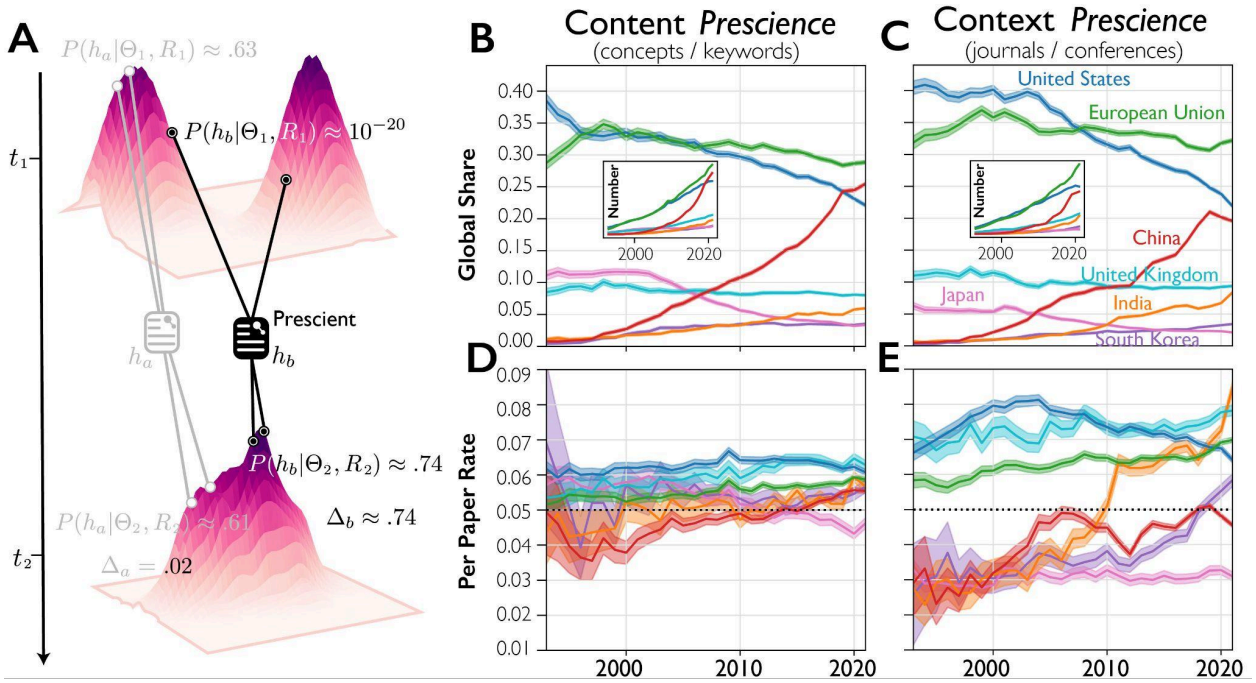


图3. 先见性创新。A，概念示意图，说明内容预见性与情境预见性如何计算及其所代表的现象。论文h_a（表示为超图）将内容流形（Θ_1）中邻近的概念加以组合，这种邻近性取决于它们在当代论文中的频繁组合，以及当代科学家因其在文献中的出现频率（R_1 及其与频率峰值的接近程度）而对其可见性。因此，h_a 在t_1 发表时并不令人意外（$p = .63$），在t_2 仍显而易见（$p = .61$）。相比之下，论文h_b 组合了可见性较低（R_1）且在（Θ_1）中距离更远的概念，从而形成一篇非常令人意外的论文（$p = 10^{-20}$）。然而，在第2期，科学的其余部分已经发生变化，使得h_b 所涉及的概念变得不再令人意外，如今频繁出现并被常规组合（即峰值已经合并）。就其内容而言，h_a 在首次发表时显而易见，而h_b 具有先见性，预示或促成了科学博弈的变化。同一图示也可重新诠释以说明情境预见性：h_a 组合了来自明显相关发表平台的思想，而h_b 则从此前很少（即便有过）被一起使用的发表平台汲取思想。**B-C**，1990至2023年先见性科学的全球占比根据各国在先见性排名前10%的论文中的份额计算并归功于相应国家，插图追踪此类论文的数量。中国先见性论文的份额正在增长，在内容预见性方面领先世界，但在情境预见性方面落后。**D-E**，先见性科学的单篇论文比率揭示了不同的情况。顶尖亚洲国家在单篇论文内容先见性上略低于世界其他地区，表明其组合非常规思想的似然较低，但中国已从亚洲群体的底部升至顶部。这些国家在情境预见性上远低于美国、英国和欧洲，表明其教育与研究致力于从传统学科（例如物理学）或预期组合（例如化学+ 材料科学）中汲取思想，其中印度显示出最显著的近期增长（参见图S4–S7，了解跨数据库、国家排除、阈值和作者归属方法的稳健性检验）。

如果说涌现体现的是科技科学发现与发明中由可察觉的潮流和趋势所带来的可预测发展，那么先见性体现的则是难以预测的偏离

这些趋势、预示着新趋势的出现。在对科技科学领域进行具有战略意义的关注度和资源投入时，前者更有把握，后者风险更高。然而，如果这些投入取得成功，后者会重塑科技科学格局，而前者则会按可预测的方式延续既有趋势。

图3A展示了科技注意力景观。在时间零点将左右两个峰值中不太可能相互关联的概念联系起来的新论文，会在这些峰值彼此靠近、并提高沿该传统开展的未来工作的可预测性的程度上，于时间一点变得具有先见性。图3B和图3C显示了所选国家在全部先见性论文中所占的份额。请注意，我们计算了两种不同的先见性重组。第一种是情境预见性，衡量一篇论文参考文献列表中期刊组合的新颖性，代表该研究所借鉴的学科多样性。第二种是内容预见性，衡量描述该研究实质方面的作者提供的关键词组合的新颖性。请注意，若一篇论文的作者来自多个国家，则该论文被平等地归属于所有作者所在的国家，因此各份额之和并不等于100。在这些分面中，内嵌图追踪了每个国家先见性出版物的总数，以提醒读者所有国家都在扩大其研究组合，没有任何国家处于衰退。图3D和图3E显示了每个国家每篇论文的先见性比率。这使得人们能够在剔除总出版物数量之后，更直接地比较每个国家的研究特征。

科学进步的整体景观

将涌现、先见性和颠覆与产出率结合起来，可以勾勒出全球科学进步的整体景观。这些指标的趋势与传统指标（图4）相似，但关键在于，各项指标发生变化的时间并不一致。中国论文的产出量迅速增长，但其引文份额却落后多年。这在一定程度上是由机制因素造成的——中国可供引用的既有研究较少。不过，篇均引用次数显示，直到2010年，中国论文在主要产出国中仍处于末位；2016年达到与欧盟持平的水平，2021年则与美国持平。这与前沿化学研究受到引文折扣的研究发现一致（*22*）。

比较涌现与先见性时，也会出现类似的模式。涌现——观察趋势并以可预测的方式延伸趋势——相对直接，但并非毫不费力。在2010年–2020年间，中国在新兴研究的比率上从末位跃居首位，反映出其对前沿研发的投入，尤其是在关键技术领域（图S8）。然而，中国在先见性方面仍落后：先于新领域出现的激进突破。与其他国家相比，一些国家（美国、英国、欧盟、印度）的国家研究文化和资助激励可能更鼓励跨学科研究。声望也可能发挥作用：前所未有的研究若源自高地位国家和机构，就更有可能引领趋势——这或许能解释美国和英国在先见性方面的领先地位，却无法解释印度的情况。

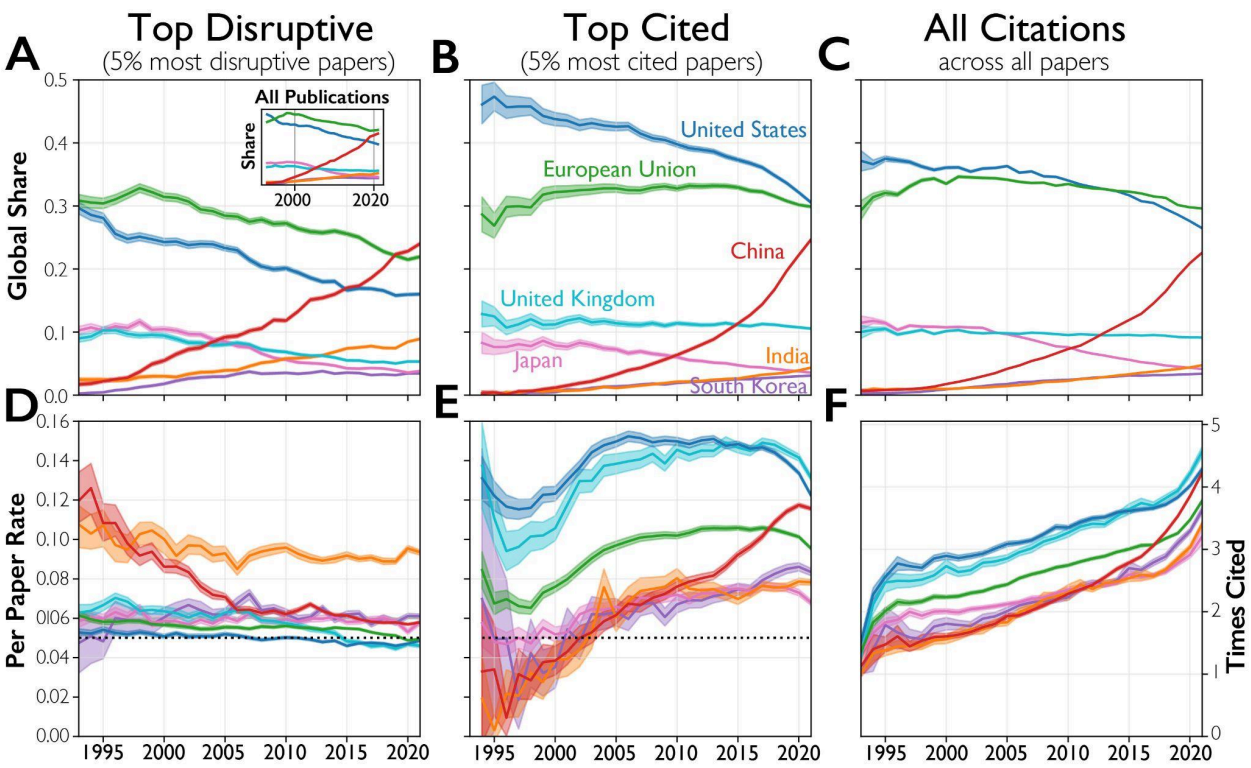


图4。颠覆、引用与发表。A-C，全球论文中颠覆性最强的前5%论文、被引次数最多的前5%论文以及所有引文的全球占比，并以插图显示出版物的全球占比。出版物总量与颠覆性最强论文的占比走势高度一致：美国和欧盟相对于中国的崛起持续下降，中国于2018年在颠覆性最强论文占比方面领先全球（**A**）。尽管直到最近，美国的出版物占比低于欧洲，但其顶尖论文和总体引用占比一直更高。中国在顶尖论文和总体引用方面正迅速接近美国和欧盟（**B-C**）。**D-F**，颠覆性最强的前5%论文、被引次数最多的前5%论文以及总体引用的单篇论文比率。我们看到，中国颠覆性下降与引用上升的趋势呈反向关系，这与印度颠覆性占比保持稳定且引用不断上升的情况形成对比。美国和英国在单篇论文引用方面领先，但中国于2017年在颠覆性最强论文和总体引用的单篇论文引用方面超过欧洲，目前几乎与美国和英国持平。

讨论

我们采用两种互补的方法来评估各国在科学进步中的角色。首先，我们将研究路径嵌入由科学家和概念构成的超图中，以识别各国在新兴领域的领先地位。其次，我们嵌入论文，以捕捉概念组合出现的（不）可能性，从而识别具有前瞻性的研究。我们将这些方法与既有指标进行比较，包括产出量、引用和引用图上的颠覆性指数，以识别哪些研究被认为催生了新的方向，哪些则是在既有方向上继续发展。我们在所有领域以及战略领域中应用这些方法，包括人工智能、能源、生物技术和半导体（图S8）。

这些互补的衡量指标表明，各国在科学领域并非只是领先或落后——而是以不同方式领先。中国的主导策略是发现并强化初步共识：在已显现出趋同趋势的领域大举投资，并大规模配置人才。这绝非轻而易举，也不只是简单模仿。要预判哪些新兴领域会产生重大影响，需要高超的判断力；要迅速落实，则需要制度层面的

协调能力，而鲜有体系能够与之匹敌。由此，中国塑造全球科学发展轨迹的方式并非改变其方向，而是加快其进程——通过最有效地研判科学格局，率先抵达前沿。

美国和欧洲追求的是一种相辅相成却又各有不同的领先地位。它们在具有前瞻性的研究方面持续保持优势——这类研究偏离可预测的发展轨迹，并预示未来方向——反映出其制度文化重视跨越学科界限、容忍失败，并赋予令人意外的成果以声望。这一点在情境预见性方面尤为明显：借鉴看似不相关的思想传统并加以组合，仍是西方一项显著优势。差距并非单由产出量造成；即使按篇均计算，差距依然存在，这表明差异源于研究文化，而不只是规模。

这些模式对科学政策和资助战略有直接影响。希望在新兴领域取得领先地位的国家，可受益于集中开展前沿扫描、迅速调动资源以及组建大型协作团队——这些正是中国现行体系的优势所在。希望推动先见性进展的国家，则需要作出不同的制度承诺：持续支持好奇心驱动型研究，开展跨学科培养与招聘，采用能够容忍高风险研究及其不确定时间安排的资助机制，并制定不会因成果出人意料而予以惩罚的评估标准。1945年，Vannevar Bush在向美国总统富兰克林·D·罗斯福提交的建立探索型美国国家科学基金会的提案中，倡导了许多与第二种做法相关的要素（*10*）。两种战略并非互斥，但它们会在各国国家研究体系内争夺关注度与资源；偏重其中一种，可能会削弱维持另一种所需的条件。

这种张力在数据中清晰可见。中国不断上升的涌现占比，与每篇论文先见性相对持平或仅有适度提升相伴随，尤其是在情境预见性方面。相反，美国和欧洲在新兴领域已让出阵地，尽管它们仍保持先见性领先地位，这表明那种能够产生意外突破的去中心化、由研究者驱动的模式，在某个有前景的方向变得清晰可辨时，可能不太善于迅速集中力量。用统计检验理论的语言来说，涌现政策会将第一类错误降至最低，从而使它所提出的方向中有更多能够成功；而无导向的先见性策略则必然会更有效地将第二类错误降至最低，因为最佳可能方向中会有更多被提出。任一体系能否同时培育这两种能力，仍是一个悬而未决的问题，并对技术竞争具有重大影响。

最后，这些发现提醒我们，不应依赖任何单一指标来衡量科学领导力。论文数量和引文反映的是过去的投入。颠覆性指数能够捕捉学术共同体的反响，却无法体现未来的发展方向。涌现性和先见性提供了前瞻性评估，但二者衡量的是不同方面——前者追踪把握科学格局的能力，后者衡量重塑科学格局的能力。一个国家若在产出量和新兴领域方面占据主导地位，却在具有前瞻性的研究上落后，就能很好地利用已知机遇，却难以创造新的机遇。反过来，情况也可能如此。全面评估科学竞争力需要综合运用所有这些指标，而它们所揭示的战略选择，归根结底关乎一个国家将建设怎样的科学未来。方法

Data

我们的建模使用 Web of Science（WoS）数据库。分析主要聚焦于1990—2023年，因为意外性与预见性分析所需的作者生成的关键词数据在1990年之前较为稀缺（见图 S1）。基于作者的分析始于2008年——这一年 WoS 的索引和作者元数据关联有了显著改善——并截止于我们拥有完整数据的最近一年。我们排除了综述论文和非英语论文。我们还使用 OpenAlex*23*（一个与 WoS 类似的开放获取数据库）重复了这些分析；该数据库对发表渠道的覆盖范围更广，尤其涵盖了发表领先人工智能研究的会议。OpenAlex 重复分析的结果详见补充信息，并总体上证实了 WoS 分析所得的发现模式。

涌现建模

一个科学领域由一个中心关键词及其在训练后的嵌入空间中的 24 个最近邻构成，因此每个领域包含 25 个关键词。如果某个领域在某一年具备以下特征，我们便将其定义为正在涌现：在此前几年中，该领域中心关键词获得的集体关注度快速增长（频率增长），使其在当年的提及次数大幅增加（流行度）；同时，在近期时间范围内，邻近关键词之间的成对语义相似度呈上升趋势（关键词汇聚）。嵌入空间采用专家感知词嵌入方法（*15*）构建，该方法在文献的超图上运行 Deepwalk 算法（*24*），为每个关键词生成一个嵌入向量。该网络中的超节点由作者和关键词构成，超边则是出版物；同一篇论文中的一组关键词和作者彼此相连。给定关键词的最近邻是在生成的嵌入空间中确定的。某个关键词在给定学科领域中的流行度，简单来说就是作者在分析年份内于该领域文献中使用该关键词的次数，再除以该领域所有关键词的总次数。频率增长率通过将形式为 $ae^{bt} + c$ 的时间指数函数拟合到同一学科领域在指定的 5 年时间窗口内计算出的关键词年度频率来求得，拟合函数的系数 b 则用作频率增长率指标。关键词汇聚用于量化该领域中的邻近关键词在嵌入空间中彼此靠拢的速度，其依据是过去三年中成对距离变化速度的平均值。平均两两距离大幅缩短，意味着汇聚速度较快，也就是说，该领域的关键词正在更快地聚集起来，形成这个领域。需要注意的是，某个关键词的流行度和频率增长取决于我们关注的学科领域，而收敛评分则独立于学科领域计算，因为嵌入空间是基于所有可用出版物统一计算的。

要将某个领域标记为新兴领域，我们要求其在全部三项标准上都得分较高。此外，为了排除中心关键词提及频率呈次指数增长的领域，我们通过R^2 量化指数拟合的质量，并将其作为第四项指标。我们根据各领域在这四项指标上的表现分别进行排序，得到四个独立排名。这些

独立排名随后通过对各领域的名次取平均值进行合并，从而为每个领域得出一个综合得分。最终得分较高的领域，其中心关键词很可能处于一组正在汇聚的关键词的核心位置，并呈现显著的频率增长和流行度。我们选取最终得分最高的前 1% 的领域，将其界定为科技科学发现与发明中的新兴领域。为考虑不同科学学科的规模差异，我们针对每个学科分别进行了上述分析。我们将现代书目数据库（如 OpenAlex）采用的 19 个领域视为主要科学学科。

为了判断一篇论文是否应被标记为新兴论文，我们首先计算了每篇论文与新兴领域之间的距离。为此，我们仅选取至少包含一个关键词的论文，并使用 Web of Science 的学科标签来划分学科。随后，我们考察了同一组学科中所有已被标记为新兴领域的领域，并计算论文与这些新兴领域之间的距离。为量化距离，我们用关键词嵌入的质心表示论文和领域，并计算余弦距离。一篇论文与一组新兴领域之间的最终距离，定义为该论文的质心嵌入与其他领域的质心嵌入之间的最小距离。新兴论文定义为与新兴领域距离最小的论文中排名前 5 个百分位的论文。

建模先见性

新的研究内容组合可能会令人意外，因为它此前从未成功过，尽管有人曾考虑并尝试过这种组合（*25–28*）；也可能是因为它从未被设想过——这代表了一种在学科对话中难以企及的思想组合。我们的模型阐述了一个简单的思维过程：科学家和发明家会将彼此结合的事物组合起来，这些事物（1）在科学或技术上相近；（2）在认知上显著。虽然形式简单，但与文献中的模型乃至先进的深度学习模型相比，该模型在捕捉新颖性和影响力方面更为有效（*16*）。具体而言，我们将关键词或来源形成组合的似然建模为以下因素的函数：（1）它们在潜在嵌入空间中的接近程度；（2）它们通过以往使用频率对科学家的显著性。具体而言，每个节点 *i*都与一个嵌入（或潜向量）相关联，该嵌入将节点映射到潜在空间 θ_i
通过优化观测到的论文和专利的似然而构建。θ_{id}潜向量的每个条目表示节点*i* 属于潜在维度*d*的概率。这些潜向量所基于的维度可自然地还原科学领域；参见（*16*）。超边*h*（一篇论文）中内容的互补性，被建模为内容节点（即，

关键词）加载到相同的维度上，。我们通过作者生成的$\sum_d \prod_{i \in h} \theta_{id}$关键词以及根据分类推断的关键词来测量这些维度，发现结果相似。语境模型也是如此，其中语境是被引用为思想来源的期刊或会议。我们为每个节点关联一个潜在标量，以反映其认知显著性，或科学家对其的接触程度*i* r_i，并测量其在网络中的整体连接度。随后我们对

组合的倾向性（ ）——即我们预期其实际出现在论文和λ_h h专利中的可能性——建模为节点在 中的邻近程度与其显著性的乘积：
h

$$\lambda_h = \sum_d \prod_{i \in h} \theta_{id} \times \prod_{i \in h} r_i \quad (1)$$

接着，将实现组合h 的出版物或专利数量建模为均值为该值的泊松随机变量。最后，超图的似然是λ_h G观察到每一种可能组合的似然的乘积。

在生物医学、物理科学和发明领域，基于往年数据，该模型能够在超过 95% 的情况下正确区分最终发表的内容组合与随机组合（生物医学：曲线下面积=0.98；物理学：曲线下面积=0.97；发明：曲线下面积=0.95）（*16*）。对未来组合的成功分类表明，我们根据已发表研究中频繁出现的内容与来源组合构建的嵌入空间，描绘了一个潜在知识空间。研究人员在产生新论文和专利时，往往会在这一空间中进行局部探索。这与先前关于科学研究中的惯性和保守搜索的发现一致（*25*，*29–32*）。

有了对科学与技术发展预期的衡量，我们将组合h的新颖性评估为其组合方式的不可能性或惊异度（*33*）：

$$novelty(h) = -\log \sum_d \prod_{i \in h} \theta_{id}. \quad (2)$$

由于我们分别对内容和来源进行建模——在一组模型中，一篇论文是内容（例如关键词）的组合；在另一组模型中，它是来源（例如引用期刊）的组合——我们也分别测量其内容组合和来源组合的新颖性。

我们通过衡量一篇论文发表时与两年后的惊异度下降幅度，来评估其先见性，从而识别出最初令人意外、但因领先地位或运气而逐渐成为常态的研究成果。由于上文所述的超图嵌入模型是动态的，且底层嵌入空间会随年份变化，我们为每篇论文计算两个惊异度分数：一个使用论文发表当年的嵌入，另一个使用两年后的嵌入。这样，我们便能通过两个惊异度分数之差来衡量论文的先见性：最初令人意外、后来却不再那么令人意外的论文，正是具有先见性或“超前于时代”的论文；它们在发表时不同寻常，却与未来的研究成果相符。与新兴领域的定义一样，如果一篇论文在其学科领域中位于前5个百分位数，我们就将其定义为具有先见性。补充信息中展示了采用其他阈值的结果。

衡量颠覆

为了捕捉科学界对论文创新性的认可程度，我们基于引用网络计算科学出版物的颠覆指数（*18*、*19*、*21*）。颠覆指数D用于表征焦点论文在多大程度上颠覆或巩固了既有科学成果，其计算方式如下：

$$D = \frac{n_f - n_b}{n_f + n_b + n_r}$$

其中，n_f 是后续论文中仅引用焦点论文而不引用其任何参考文献的论文数量，n_b 是后续论文中同时引用焦点论文及其参考文献的论文数量，n_r 是后续论文中仅引用焦点论文参考文献的论文数量。颠覆指数介于 -1 和 1 之间，用以描绘焦点论文颠覆或推动科学发展的程度。如果一项研究完全盖过先前研究，吸引所有后续关注都聚焦于自身，而其思想源头则被遗忘，那么该指标等于 1。如果一项研究总是与其思想源头一同被引用，反之亦然，那么可以认为该研究彰显了先前研究的重要性，此时该指标等于 -1。例如，IBM 于1982年获得的扫描隧道显微镜美国专利，以及 Cetus 于1987年获得的“核酸序列扩增方法”（即聚合酶链式反应，PCR）美国专利，均与诺贝尔奖有关，且颠覆性得分均为 1；而亚马逊成功获得的在线购物“一键式”专利得分则远低于 1，并且推动了先前的电子商务创新（*18*）。颠覆指数越高，意味着焦点论文越有可能颠覆现有思想，其创新性也越受科学界认可；指数越低，则表明焦点论文更倾向于发展或巩固既有知识。为确保比较公平，我们仅计算同时具有参考文献和引用论文的论文的颠覆性得分。与现有文献（*18*、*19*、*21*）类似，我们采用 CD_5 指数，即使用五年时间窗口的颠覆指数；计算 n_f、n_b 和 n_r 时，仅纳入在焦点论文发表后五年内发表的论文。我们将学科领域内颠覆指数位居前 5 个百分位数的论文定义为颠覆性论文，以便与其他分析保持一致。

图

图显示了年度比例。在适当情况下，我们计算了 95%置信区间，并以线条周围的阴影区域表示。

数据可用性

Web of Science 数据可从科睿唯安获取。OpenAlex 数据可在此处公开获取：https://openalex.org

致谢

我们感谢美国国家科学基金会的支持（资助项目：J.E. 获 2404109 号，J.E. 和 J.L. 获 2241237 号）、美国空军科学研究办公室的支持（FA9550-19-1-0354 和

FA9550-15-1-0162 号资助 J.E.），以及美国国防高级研究计划局的“科学创造力的建模与测量”资助项目（J.E.）。我们感谢关键技术分析国家网络的成员，以及该网络接触过的政策界成员所提供的启发和反馈。

作者贡献

J.L.、J.S. 和 J.E. 共同构思并设计了本研究，解读研究结果，并起草和修订了稿件。F.S. 提供了概念、建模和软件方面的支持。J.L. 整理了数据，J.L. 和 J.S. 进行了数据分析。J.L. 和 J.E. 设计并制作了图。

参考文献

1. P. A. David, D. Foray，知识社会的经济基础。*Pol. Futur. Educ.* **1**，20–49（2003）。2. J. E. Stiglitz，《知识作为全球公共产品》，载于全球公共产品（牛津大学出版社纽约，1999），第308–325页。3. J. Mokyr，雅典娜的馈赠（普林斯顿大学出版社，新泽西州普林斯顿，2011）。4. J. A. Schumpeter，《资本主义、社会主义与民主》（1942）。《经济文献杂志》 **20**，1463（1976）。5. U. Akcigit, J. Van Reenen，创造性破坏经济学：*Aghion* 和 *Howitt* 主题的新研究（哈佛大学出版社，2023）。6. W. Sombart，《战争与资本主义》（战争与资本主义）。慕尼黑和莱比锡：*Duncker & Humblot* 出版社（1913）。7. R. R. Nelson, N. Rosenberg，技术创新与国家体系。国家创新体系：比较分析 **322**（1993）。8. C. Freeman，“国家创新体系”的历史视角。*Cambridge J. Econ.* **19**，5 24 1995–（）。9. 经济合作与发展组织，《弗拉斯卡蒂手册*2015*》（经济合作与发展组织，2018）。10. V. Bush，《科学：无尽的前沿：Vannevar Bush 致总统的报告》。科学研究与发展办公室主任，美国政府印刷局，华盛顿（1945年）。11. P. Kennedy，《大国的兴衰》（Vintage Books 出版社，纽约，1987年）。12. D. C. Mowery, N. Rosenberg，《技术与经济增长的追求》（剑桥大学出版社，英格兰剑桥，1989）。13. F. Shi, J. G. Foster, J. A. Evans，编织科学之网：科学演进结构的动态网络模型。*Soc. Networks* **43**，73–85（2015）。14. J. Sourati, J. Evans，借助人类与外星人工智能加速科学，*arXiv*

*[*计算机科学-人工智能*]* （2021）。http://arxiv.org/abs/2104.05188。

15. J. Sourati, J. Evans, 借助人类感知人工智能加速科学研究，*arXiv [*计算机科学-人工智能*]* (2023年)。~~http://arxiv.org/abs/2306.01495~~。16. F. Shi, J. Evans，研究内容与背景的出人意料的组合与影响力相关

nd

远缘学科的科学局外人由此涌现。《自然·通讯》 **14**（2023年）。17. B. Hofstra, V. V. Kulkarni, S. Munoz-Najar Galvez, B. He, D. Jurafsky, D. A. McFarland，《科学中的多样性-创新悖论》。美国国家科学院院刊 **117**，9284–9291（2020年）。18. R. J. Funk, J. Owen-Smith，《技术变革的动态网络测度》。《管理科学》 **63**，791–817（2016年）。19. L. Wu, D. Wang, J. A. Evans，大型团队推动科学与技术发展，小型团队则颠覆科学与技术

y.

自然 **566**，378–382（2019）。20. Y. Lin, J. A. Evans, L. Wu，科学的新方向源于断裂与分歧。*J. Informetr.* **16**，101234（2022年）。21. M. Park, E. Leahey, R. J. Funk，论文和专利的颠覆性正随时间减弱。自然 **613**，138–144（2023年）。22. S. Qiu, C. Steinwender, P. Azoulay，谁站在中国（科学）巨人的肩膀上？来自化学的证据（2022年）。https://doi.org/10.3386/w30772。23. J. Priem, H. Piwowar, R. Orr，OpenAlex：一个完全开放的学术著作、作者、发表平台、机构和概念索引。arXiv [预印本] （2022年）。https://doi.org/10.48550/arXiv.2205.01833。24. B. Perozzi, R. Al-Rfou, S. Skiena，“DeepWalk：社会表征的在线学习”，载于第*20*届*ACM SIGKDD*知识发现与*D*国际会议论文集

ata

挖掘（美国计算机协会，纽约州纽约市，2014 年）*KDD '14*，第 701–71 页

0.

25. A. Rzhetsky、J. G. Foster、I. T. Foster、J. A. Evans，通过选择实验加速集体发现。 美国国家科学院院刊 **112**，14569–14574（2015）。26. J. G. Foster、A. Rzhetsky、J. A. Evans，科学家研究策略中的传统与创新。 [预印本] （2015）。https://doi.org/10.1177/0003122415601618。27. H. Rheingold，思想工具：下一场计算机革命背后的人物与思想（Simon & Schuster Trade，1985）。28. D. Simandan，感到惊讶并令自己惊讶：个人与社会变迁的地理学

e.

《进步的人文地理学》 **44**，99–118（2020年）。29. M. Callon、A. Rip、J. Law，绘制科学与技术的动态：科学社会学

e

现实世界中的（施普林格，1986）。30. R. Guimerà、A. Díaz-Guilera、F. Vega-Redondo、A. Cabrales、A. Arenas，《拥塞条件下局部搜索的最优网络拓扑》。[预印本] （2002）。https://doi.org/10.1103/physrevlett.89.248701。

31. T. E. Stuart、J. M. Podolny，《局部搜索与技术能力的演化》。战略 *c*
Manage. J. **17**, 21–38 (2007). 32. T. Jia, D. Wang, B. K. Szymanski，《研究兴趣演化模式的量化》。*Nature Human Behaviour* **1**，1–7（2017）。33. T. M. Cover、J. A. Thomas，《信息论基础》（John Wiley & Sons，2012）。

补充信息

材料与方法

1.1 Web of Science 数据

主要分析使用的是科睿唯安提供的 Web of Science（WoS）核心合集数据副本。本文报告的分析聚焦于2000—2023年。我们将分析范围限定为英语研究论文和综述论文。

根据作者的机构隶属关系为论文确定所属国家。在主要分析中，如果一篇论文的作者来自多个国家，我们会将全额计入每个国家。对于 2008年以前索引的论文，WoS 提供作者名单及其机构隶属关系列表，但并未将二者关联起来。2008年以后，WoS 会将机构隶属关系与特定作者关联，因此我们可以评估根据第一作者、末位作者或通讯作者所在国家为论文确定所属国家所产生的影响。结果在定性上相似。

1.2 排除非英语来源

如果对科学文献的国际分析仅限于一种语言，就会产生偏差。尤其是中国拥有强大的中文学术出版业，而这部分并未纳入我们的分析。我们将范围限定为英语，部分原因是技术限制：目前根本没有能够全面、多语言索引学术出版物的数据来源；即便有，也难以跨语言关联实质性主题和领域。不过，如今全球绝大多数学术出版物都使用英语（视学科领域而定，占 80% 至 95%），而随着英语在学术写作中的主导地位和激励作用日益增强，这一比例还在不断上升。(*1*, *2*) 此外，正如本文所做的那样，研究中国如何在英语科学文献中崛起，本身就很有意义，尽管这并不能完整呈现中国开展的所有研究。

1.3 新兴领域

1.3 超图 Deepwalk

Deepwalk 算法（*3*）包括在网络上执行随机游走采样，然后基于生成的节点序列训练词嵌入模型，从而学习节点的向量表示。这一表示学习过程是为常规图设计的，其中每条边连接一对节点。超图在结构上的不同之处在于，（超）边连接多个（超）节点。这里，我们沿用专家感知嵌入方法（4），通过执行广义随机游走将 DeepWalk 算法扩展到超图。与标准随机游走不同，广义随机游走会在随机选择下一个邻近节点之前，增加一个超边采样步骤。我们为每个随机游走序列生成了 20 个节点，并通过将专家感知 DeepWalk 算法中的非均匀采样参数 设为 1，使作者节点与非作者节点的数量保持平衡

。对于每个分析年份，我们构建了一个由过去 5 年间产生的文献构成的超图。每年生成的随机游走序列数量设为该时间窗口内文献涉及的关键词数量。图 S1 展示了科学词汇随时间的增长情况：每年的关键词数量从 1990 年的不足 250,000 个增至 2020年 的 750,000 多个，而平滑了逐年波动的 5 年超图窗口中，独特关键词数量超过了 150 万个。

1.3.2 关键词涌现评分

基于频率的评分，即频率增长和流行度，是根据各年份摘要中关键词的提及次数计算的。收敛评分的计算方式不同，取决于关键词训练得到的嵌入。对于任意给定的嵌入为$E_{w,t}$ 的关键词 w，在年份t，我们考察其对应的领域，即年份t中与 w 最近的 25 个关键词。随后，我们考察该领域内所有关键词之间的两两距离在过去三年中的变化情况。具体而言，对于所有满足 w_i、w_j属于与中心关键词 w 对应领域的词对 (w_i , w_j)，我们有：

$$\Delta_t(w_i, w_j) = d(E_{wi,t} , E_{wj,t}) - d(E_{wi,t-1} , E_{wj,t-1})$$

$$\Delta_{t-1}(w_i, w_j) = d(E_{wi,t-1} , E_{wj,t-1}) - d(E_{wi,t-2} , E_{wj,t-2})$$

$$\Delta_{t-2}(w_i, w_j) = d(E_{wi,t-2} , E_{wj,t-2}) - d(E_{wi,t-3} , E_{wj,t-3})$$

其中，d 是一种距离度量（在我们的案例中为余弦距离）。对于新兴领域，我们预期嵌入集合会收缩，或者等价地，距离差 Δ's 为负，因为两两距离会逐年减小。换言之，w_i 和 w_j 在年份t的语义距离会比在年份t–1 更近，因此 $d(E_{wi,t} , E_{wj,t}) \le d(E_{wi,t-1} , E_{wj,t-1})$。距离差 Δ's 越负，对应领域的收敛程度就越高。因此，我们将上述定义的 Δ's 在过去三年间、对该领域的所有关键词取平均值，作为收敛评分。

图 S2A 显示，直到 21 世纪初，医学的新兴领域数量最多；此后工程学超过医学，并迅速增长。我们针对每个主题分别进行了涌现分析，因此某个关键词可能在一个学科领域中界定了一个新兴领域，而在另一个学科领域中却不存在或不属于新兴领域。不同主题间重叠的新兴领域在早期较少，随后逐渐增加。图 S2B–C 显示了 2010年和 2020年各学科领域之间的两两重叠情况：最大的重叠从工程学—化学（2010年）转向工程学—计算机科学（2020年），反映出计算方法的中心地位日益提升。2010年，最大的重叠出现在“工程学”和“化学”之间；而到 2020年，共享的新兴领域数量最多的则是“工程学”和“计算机科学”。

1.4 预见性建模

1.4.1 衰退领域的论文

先见性涉及那些在发表后其所属领域发生概率大幅上升的论文，我们也可以考察相反的情况：发生概率降低的论文，即处于衰退领域的论文。我们观察到与先见性论文相似的模式：中国学者所占比例偏高，而美国和欧洲学者所占比例偏低。换言之，中国研究的方差更大。

1.4.2 预见性的含义

我们的预见性模型能够识别出那些“超前于时代”的论文：这些论文将关键词或来源组合在一起，而这些组合在焦点论文发表时出乎意料，后来却成为常态。重要的是，我们采用论文发表后 2 年的时间窗口。出于必要，我们使用了这个相对较短的时间窗口，因为我们的目的是识别当前科学领导力的状况；如果采用更长的时间窗口，我们就只能识别历史上的先见性论文。时间窗口的长度会影响预见性得分所衡量的内容，以及它们不衡量的内容。具体而言，我们的预见性得分识别出那些超前于时代*2* 年的论文。

1.4.3 预见性与意外性

预见性和意外性体现了新颖性的不同方面。意外性衡量论文中的组合在发表时有多么不同寻常；预见性则衡量这些组合后来是否成为常态。图 S3 展示了这些指标与引文影响力之间的关系。无论是内容（关键词）还是背景（参考文献）指标，与意外性相比，预见性与论文跻身被引次数最高的 10% 论文之列的似然相关性更强。这一模式在最高百分位数处尤为明显：预见性高于第 95 百分位数的论文（这是主要分析中“先见性”论文的阈值，以垂直虚线标示）比仅仅具有意外性的论文更有可能获得高引用。这表明，科学影响力并非源于与众不同本身，而是源于预见学科领域将要发展的方向。

1.5 设计稳健性

1.5.1 数据库的影响（Web of Science 与 OpenAlex）

为检验稳健性，我们在另一个数据集 OpenAlex（Priem 等，2022年）上重复了所有 WoS 分析。总体而言，OpenAlex 的规模更大，对出版渠道的质量和类型限制更少，因此涵盖了更多学位论文、报告、资助项目和会议论文集。值得注意的是，OA 没有作者自选关键词数据，因此我们没有使用关键词，而是采用 OA 将出版物自动分类为 6.5 万个“概念”的分类结果，其中许多概念与传统的作者关键词相似。

使用该数据得出的结果在定性上相似。下文展示了各国使用细粒度概念（第 3 级及以下，详细程度大致相当于 Web of Science 作者关键词）计算的预见率。

图 S4 比较了 Web of Science（上行）和 OpenAlex（下行）在全部四项创新指标上的结果。定性模式是一致的：中国的份额在所有指标上随时间上升，在新兴领域和内容预见性上超过美国和欧盟，而在情境预见性和颠覆性论文上仍然较低。

使用 OpenAlex 概念得到的结果在定性上相似，但与使用关键词的结果略有不同。在各自子领域中写作的作者，其视角和动机不同于对所有已发表成果进行分类的科学计量数据库。其次，关键词是在发表之前选定的，而概念则是事后应用的，并受益于事后视角。我们的预见性建模专门旨在捕捉那些在当时的学科领域内并不明显、但在事后视角下变得显而易见的创新。因此，使用诸如概念这类基于事后视角的变量，可能会削弱估计。最后，OpenAlex 收录的出版物范围比 Web of Science 更广，包括地位较低的期刊以及其他类型的出版物（学位论文、图书、白皮书等）。抽样框的差异也可能解释两个数据库之间的适度差异。

1.5.2 排除本国论文的影响

所呈现的主要结果基于一个涵盖所有论文的模型。然而，美国、欧盟和中国等主要科研产出国占全球科学文献的相当大份额，因此一个国家的创新和趋势可能不会传播到世界其他地区。我们通过拟合其他模型来检验这一点：这次只使用没有相应地区作者的论文进行模型训练，然后评估从外部世界的视角来看，这些论文会呈现出怎样的面貌。

图 S5 展示了从模型训练中排除各主要科研产出国（美国、中国、欧盟）论文后的结果。实线表示全样本结果；淡色线表示排除论文后的结果。无论排除哪个国家，新兴领域和语境前瞻性指标基本保持不变，这表明这些形式的创新会在国际间扩散。然而，内容预见性呈现出国家特异性：从外部世界的视角来看，各国在语言趋势方面似乎都较少具有先见性（淡色线低于实线）。这表明，在描述新科学工作方面，各国存在显著的国家文化差异。

1.5.3 改变“顶尖”论文阈值的影响

我们的主要结果将各项指标排名前5%的论文定义为创新性工作，但这一阈值是任意设定的。图 S6 展示了三个阈值下的结果：前 1%（第 99 百分位）、前 5%（第 95 百分位数）和前 10%（第90百分位）。新兴领域和内容预见性的结果在不同阈值下均保持稳定。对于颠覆性论文，美国和欧盟的趋势稳健，但阈值越高，中国所占比例越低，这表明中国的颠覆性论文更多集中在第90至第95百分位，而非前1%。情境预见性

则呈现相反的模式：中国论文集中在第90至第95百分位，而美国和欧盟的论文则明显更多地出现在第99百分位。

1.5.4 作者归属的影响

将论文归属到国家并不简单：论文通常有多位来自不同国家的作者，有些作者还拥有多个国家的机构。图S7展示了采用五种归属策略的结果：仅按通讯作者、仅按第一作者、仅按末位作者、按任一作者归属（所有国家均获全额计入，与我们的主要分析一致），以及一致归属（仅纳入所有作者来自同一国家的论文）。中国在新兴领域份额上升、西方主导地位下降的定性模式，在所有归属方法下都保持一致，表明这些结果并非国际合作模式造成的假象。

1.6 主题领域

图 S8 展示了美国国家科学基金会和美国国会认定具有战略重要性的研究领域的结果，包括生物医学技术/基因组学/合成生物学、人工智能/机器学习/自主系统、先进通信、灾害预防、数据管理/安全以及能源。上方各行显示创新论文的份额；下方各行显示相对于出版物数量的创新率（虚线标示 5% 基线）。结果显示出领域特异性差异：中国在多数领域的新兴领域中处于领先地位，而美国在前瞻性指标方面仍具优势，尤其是在人工智能/机器学习和生物医学领域。规模较小的学科领域中，置信区间（阴影区域）更宽。

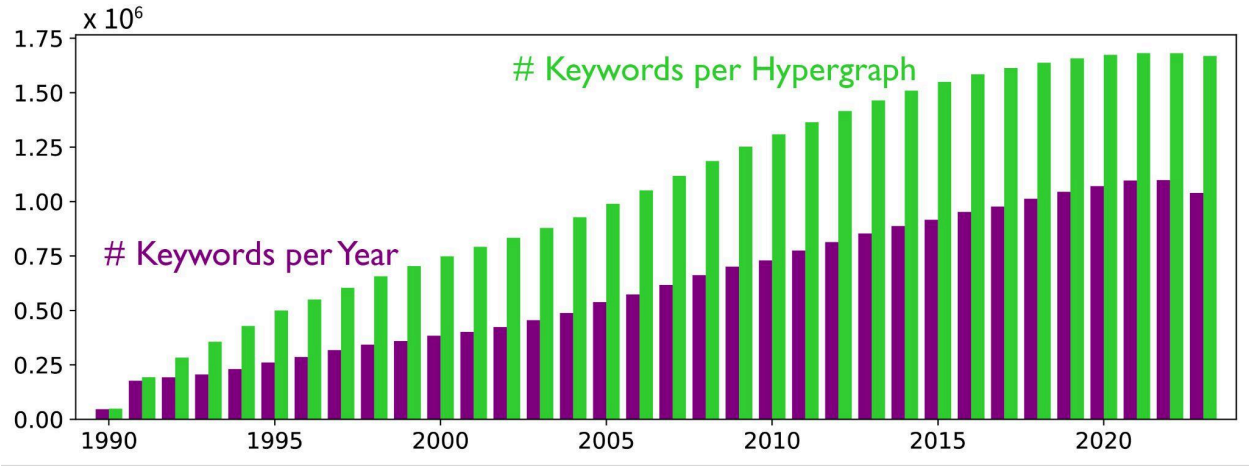


图 S1。 超图嵌入分析捕捉到的科学词汇增长。按单个年份（紫色柱）以及基于滚动的 5 年发表时间窗口构建的超图（绿色柱）统计科学文献中出现的唯一关键词数量，时间范围为 1990–2023年。5 年超图窗口平滑了逐年变化，同时保留了足够的时间上下文，以进行嵌入收敛分析。到 2020年，词汇总量已超过 150 万个关键词，反映出科学研究的迅速扩张与日益专业化。

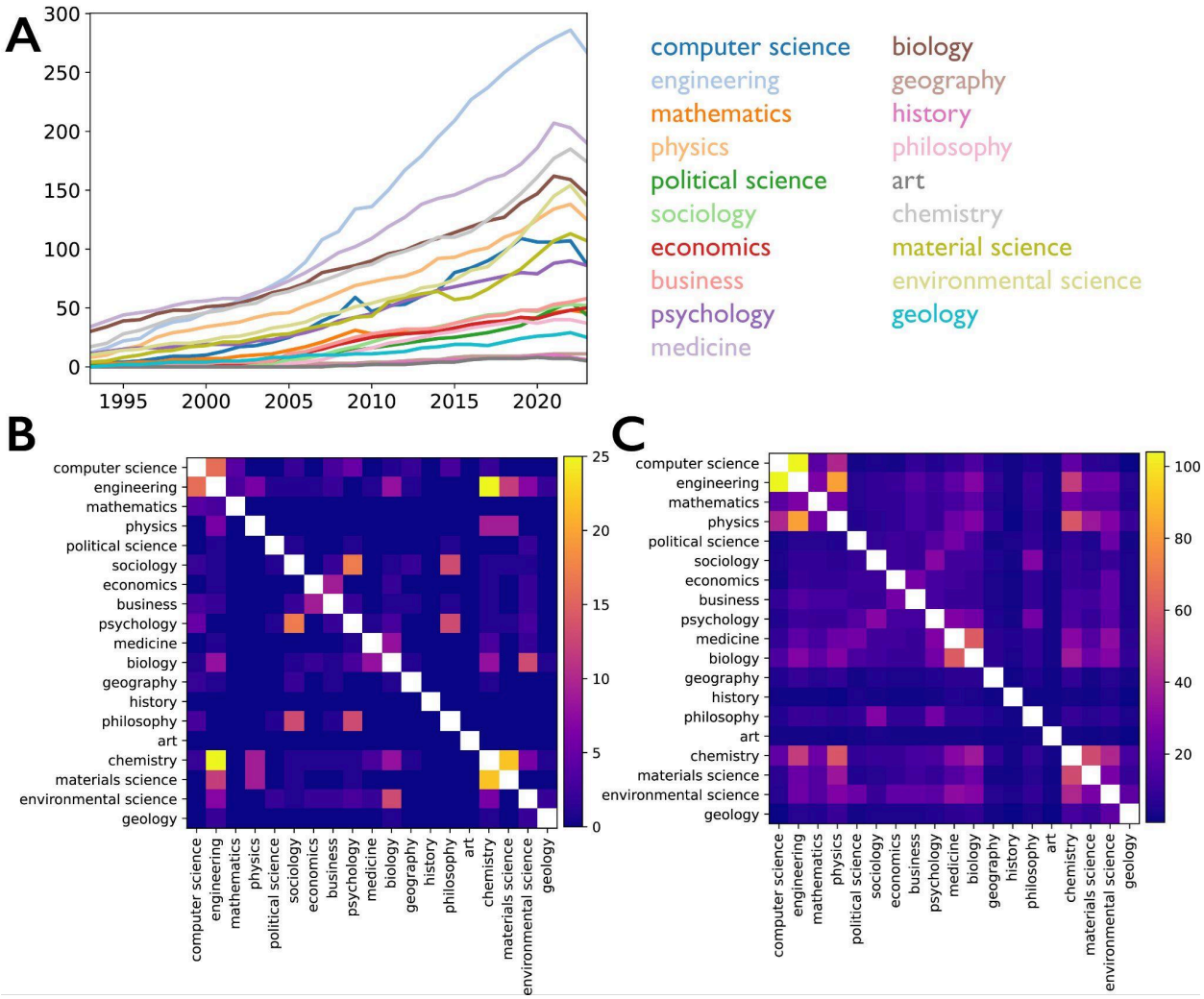


图 S2. 各科学学科的新兴研究领域。 **A**，1993–2025年，每年在18个科学子领域中识别出的新兴领域数量。直到21世纪初，医学（深蓝色）一直是新兴领域产出的主要学科；此后，工程学（浅绿色）超过医学并迅速增长。近年来，计算机科学（紫色）领域的增长呈加速趋势，其中包括机器学习和生成式人工智能（见图 2）。**B–C**，热图展示了2010年（**B**）和2020年（**C**）各学科两两之间重叠的新兴关键词数量，颜色越浅表示重叠程度越高。最大的跨学科重叠从2010年的工程学—化学转变为2020年的工程学—计算机科学，反映出计算方法在各领域中的核心地位日益凸显。

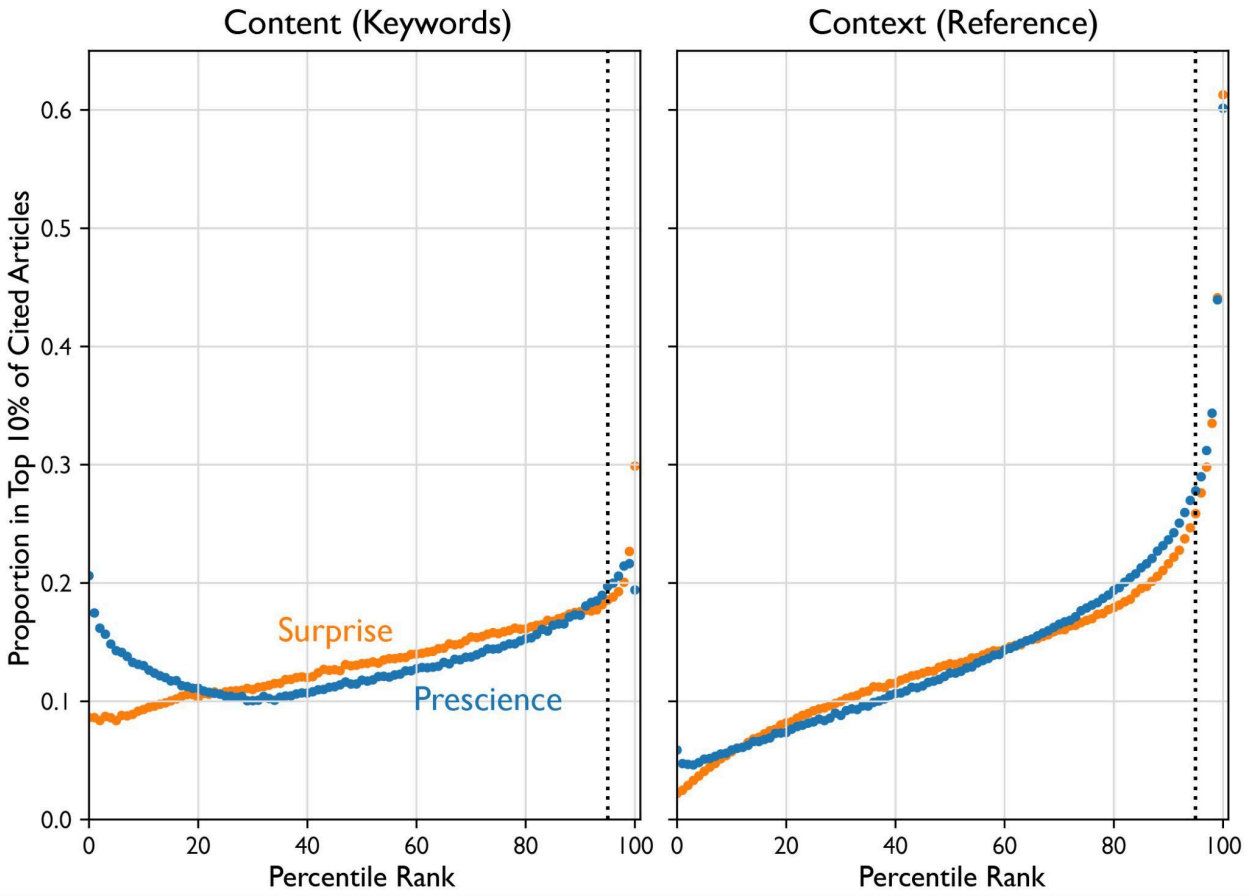


图 S3. 新颖性、先见性与引文影响力之间的关系。 按意外性（橙色）和先见性（蓝色）指标的百分位排名划分，引用次数位居前 10% 的论文所占比例。左图：内容（关键词）指标；右图：语境（参考文献）指标。垂直虚线表示主要分析中用于定义先见性论文的第95百分位阈值。与意外性相比，先见性与引文影响力呈现出更强的正相关关系，尤其是在最高百分位数处。这表明，融合了后来逐渐成为常规要素的内容的论文，比仅在发表时显得不同寻常的论文更具影响力。

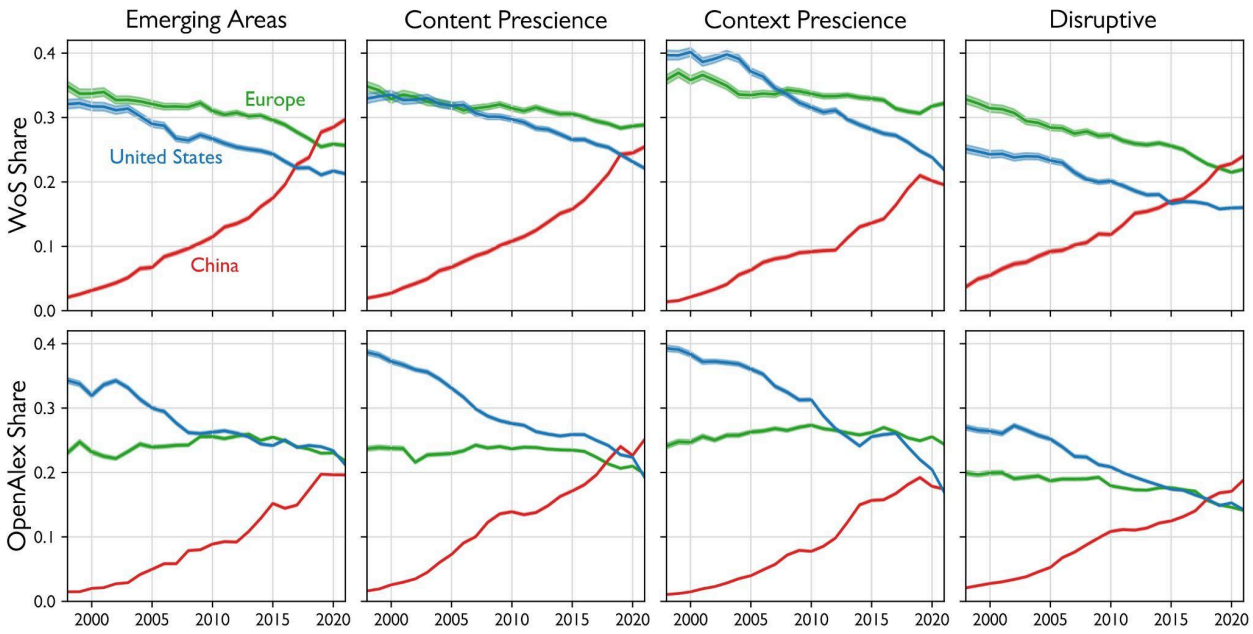


图 S4. 不同文献计量数据库中结果的稳健性。按国家划分的创新论文占比（中国为红色、美国为蓝色、欧洲为绿色），涵盖四项创新指标：新兴领域、内容预见性、情境预见性和颠覆性论文。上排：Web of Science（WoS）数据。下排：OpenAlex 数据。尽管覆盖范围存在差异——OpenAlex 纳入了 WoS 未收录的学位论文、报告和会议论文集——但不同数据库中的定性模式仍保持一致。随着时间推移，中国在所有指标中的占比均有所上升，并在新兴领域、内容预见性和颠覆性方面超过美国和欧盟，但在情境预见性方面仍处于最低水平。

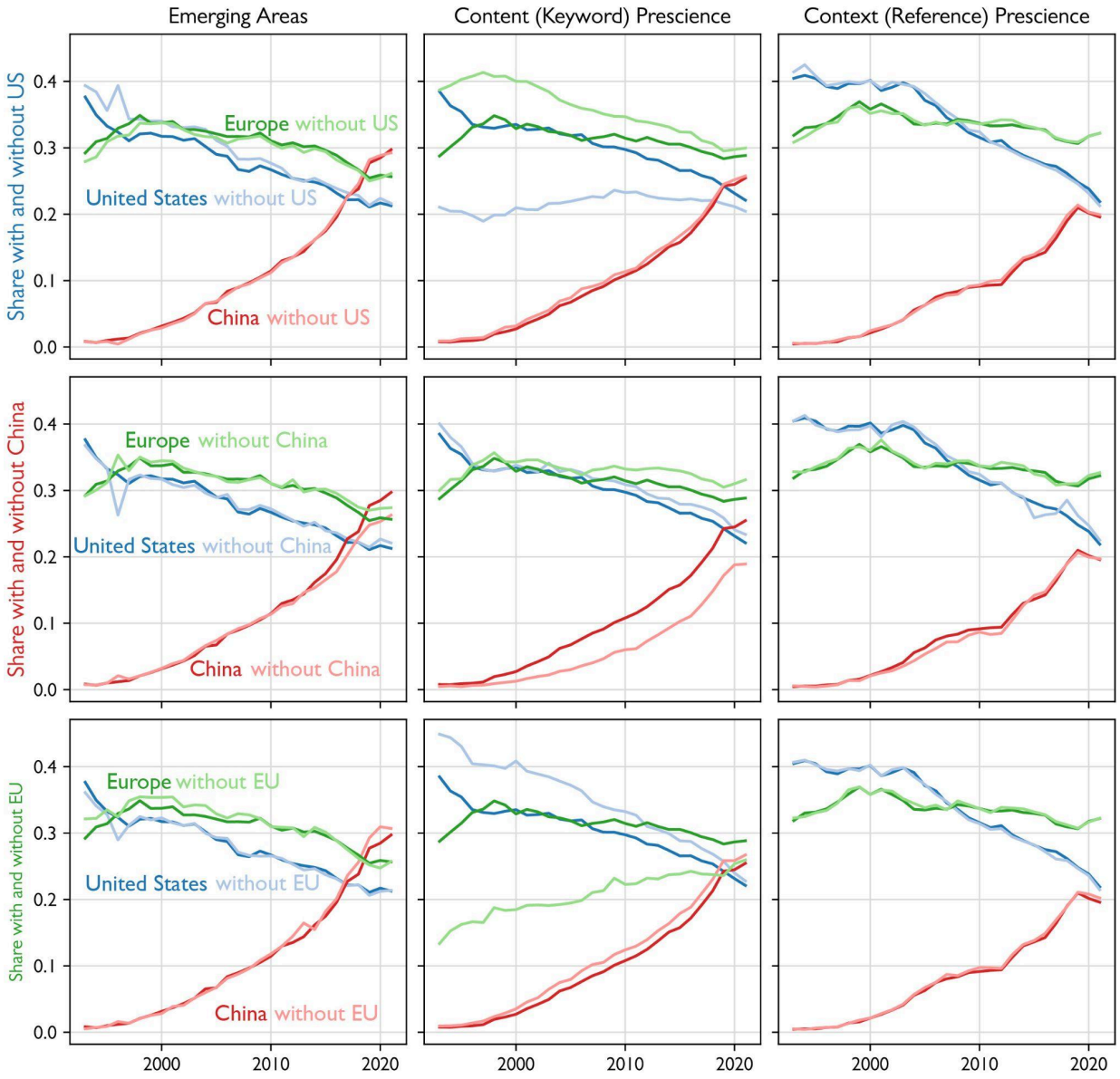


图 S5. 排除主要科研产出国后创新度量的稳定性。每一行展示了从模型训练中排除一个主要科研产出国的论文后所得的结果：美国（上）、中国（中）和欧盟（下）。实线表示全样本结果；淡色线表示排除指定地区后的结果。无论排除哪个国家或地区，新兴领域和语境预见性度量基本保持不变，这表明这些形式的创新会在国际间传播。然而，内容预见性表现出一定程度的国家特异性：从外部世界的视角看，各国在语言趋势方面似乎都不那么具有先见性，这表明描述科学工作的方式存在显著的国家文化差异。

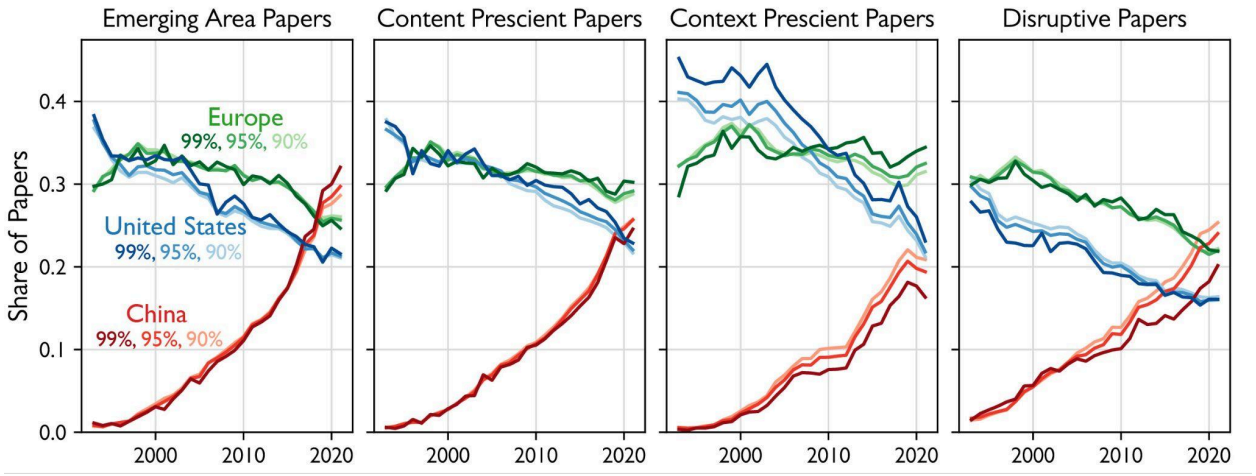
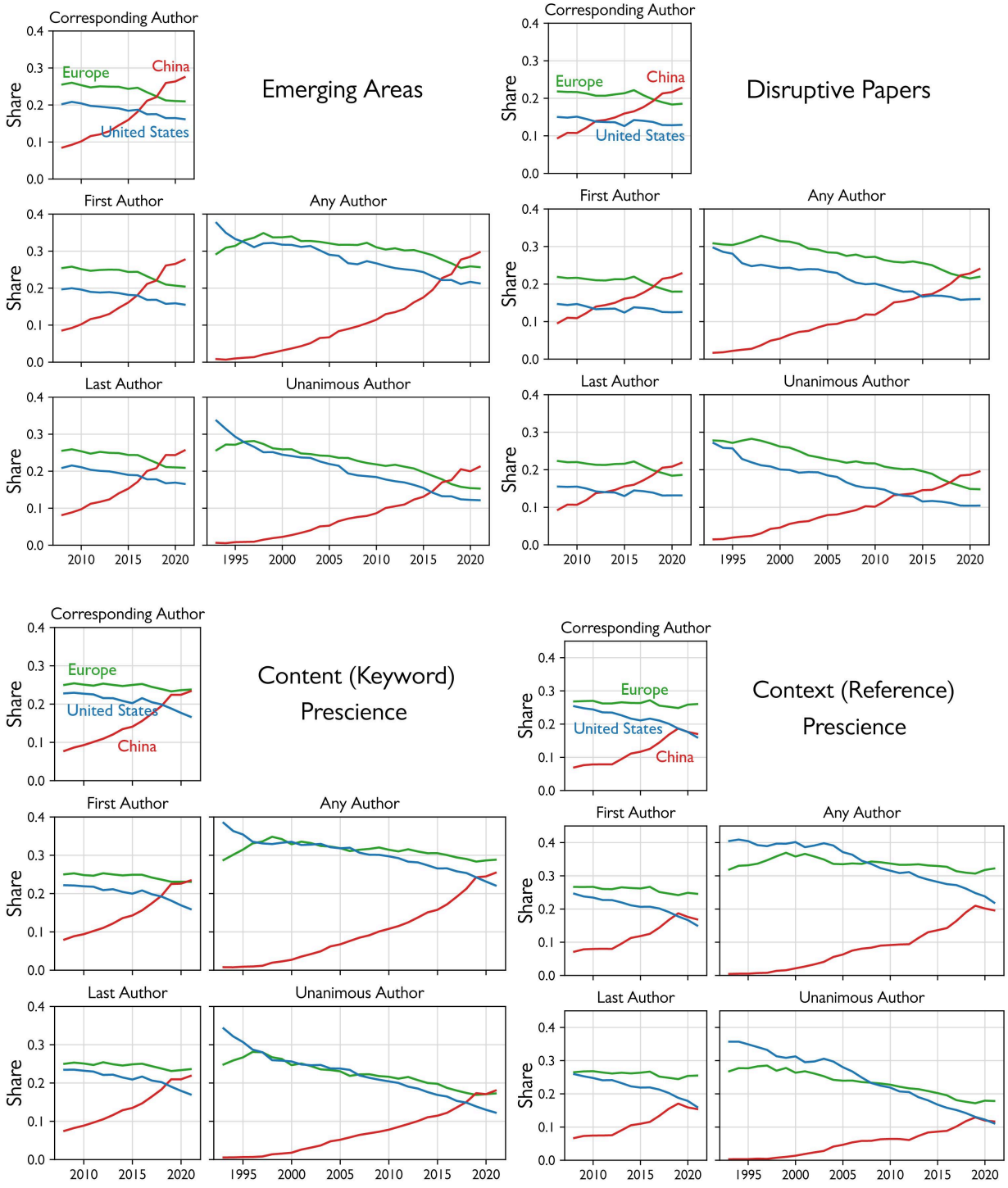
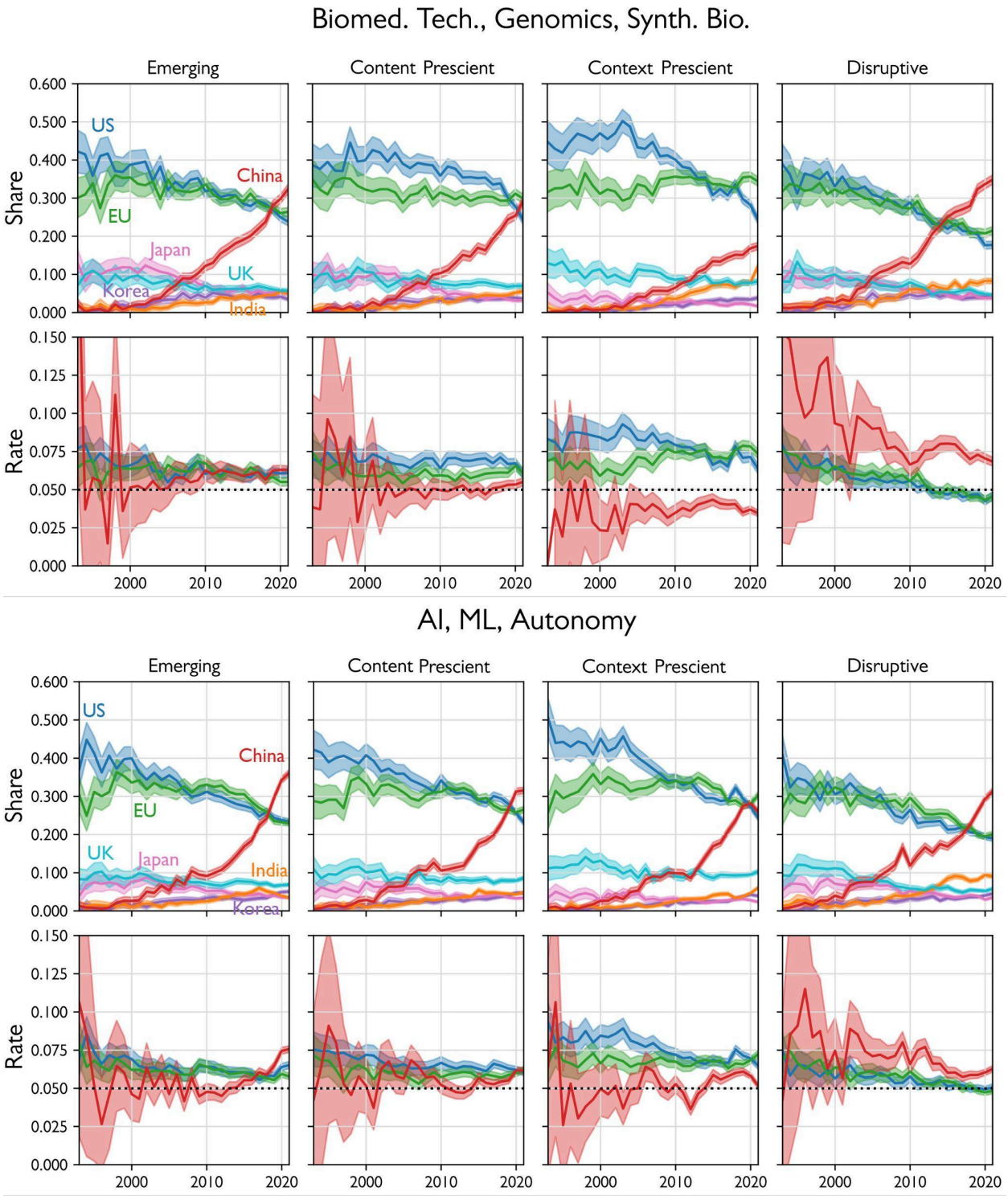


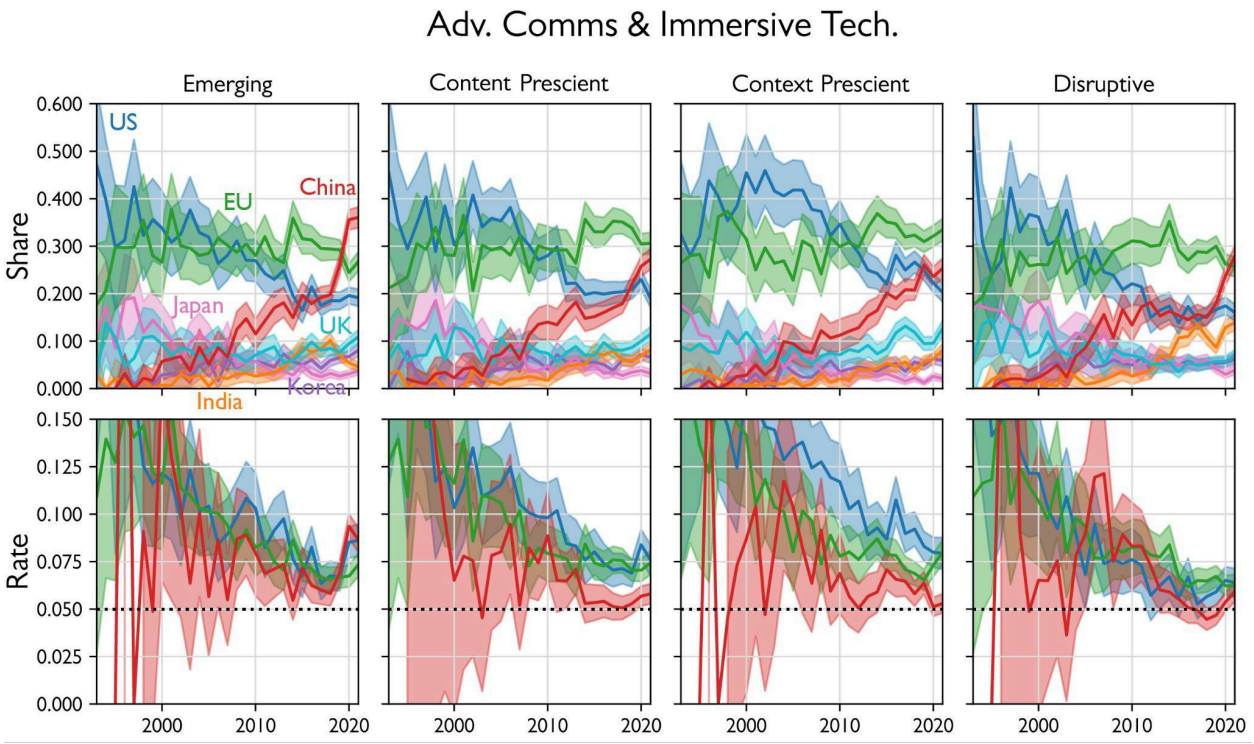
图 S6. 定义创新性工作时采用的百分位数阈值对结果的影响。不同国家的论文在四项创新指标中的占比，分别采用三个阈值：前 1%（第99百分位，颜色最深）、前 5%（第 95 百分位数，颜色中等）和前 10%（第90至第95百分位，颜色最浅）。新兴领域和内容预见性的结果在不同阈值下均保持稳定。对于颠覆性论文，美国和欧盟的趋势稳健，而中国的占比在较高阈值下较低，表明中国论文更多集中在第90至第95百分位范围内。情境预见性呈现出相反的敏感性：中国论文集中在第90至第95百分位范围内，而美国和欧盟的论文则明显更多地出现在第99百分位。

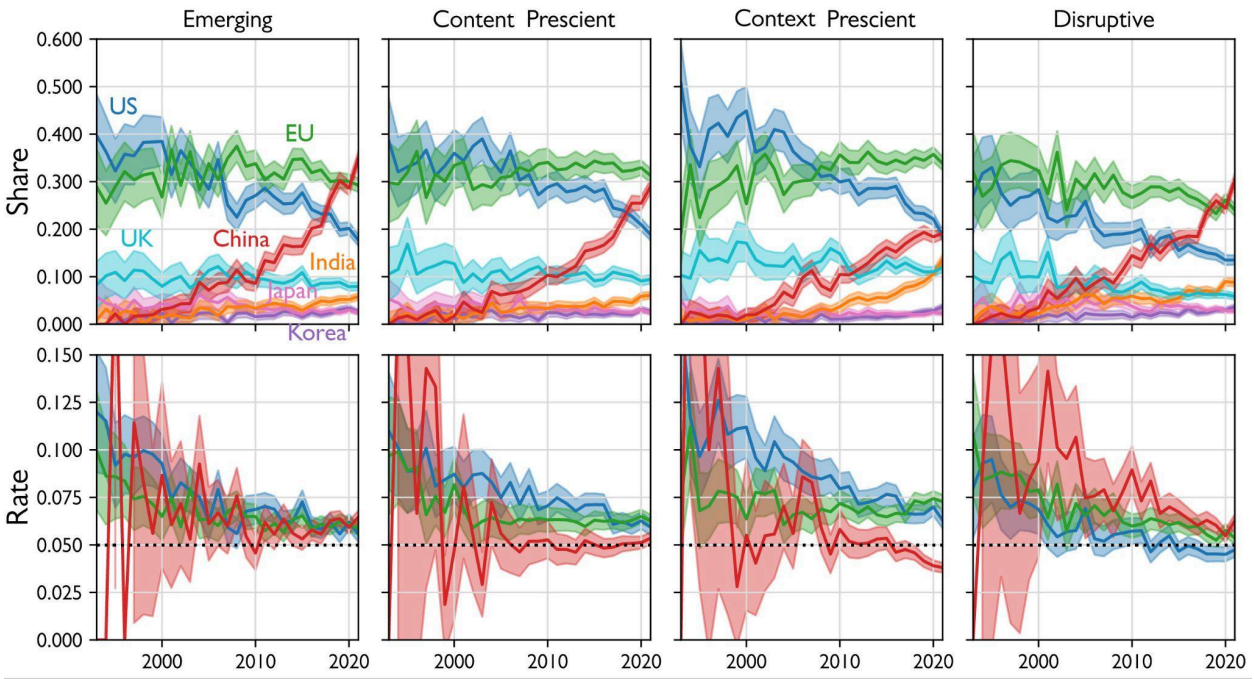


图S7. 结果对作者归属方法的稳健性。针对四项创新指标（新兴领域、颠覆性论文、情境预见性、内容预见性），采用不同策略将论文归属到国家：仅计入通讯作者、仅计入第一作者、仅计入末位作者、计入任一作者（所有相关国家均获全额计入），以及一致归属（仅纳入所有作者均来自同一国家的论文）。中国上升份额的定性趋势

新兴领域以及西方主导地位的下降在所有归因方法中均保持一致，表明这些结果并非国际合作模式或署名惯例的产物。


Biomed. Tech., Genomics, Synth. Bio.
Emerging
Content Prescient
Context Prescient
Disruptive
Share
Rate
US
China
EU
Japan
UK
Korea
India
0.600
0.500
0.400
0.300
0.200
0.100
0.000
0.150
0.125
0.100
0.075
0.050
0.025
0.000
2000
2010
2020
AI, ML, Autonomy
Emerging
Content Prescient
Context Prescient
Disruptive
Share
Rate
US
China
EU
UK
Japan
India
Korea
0.600
0.500
0.400
0.300
0.200
0.100
0.000
0.150
0.125
0.100
0.075
0.050
0.025
0.000
2000
2010
2020


Adv. Comms & Immersive Tech.
Emerging
Content Prescient
Context Prescient
Disruptive
Share
Rate
US
EU
China
Japan
UK
India
Korea
0.600
0.500
0.400
0.300
0.200
0.100
0.000
0.150
0.125
0.100
0.075
0.050
0.025
0.000
2000
2010
2020


Disaster Prevention & Mitigation
Emerging
Content Prescient
Context Prescient
Disruptive
Share
Rate
US
EU
UK
China
India
Japan
Korea
0.600
0.500
0.400
0.300
0.200
0.100
0.000
0.150
0.125
0.100
0.075
0.050
0.025
0.000
2000
2010
2020

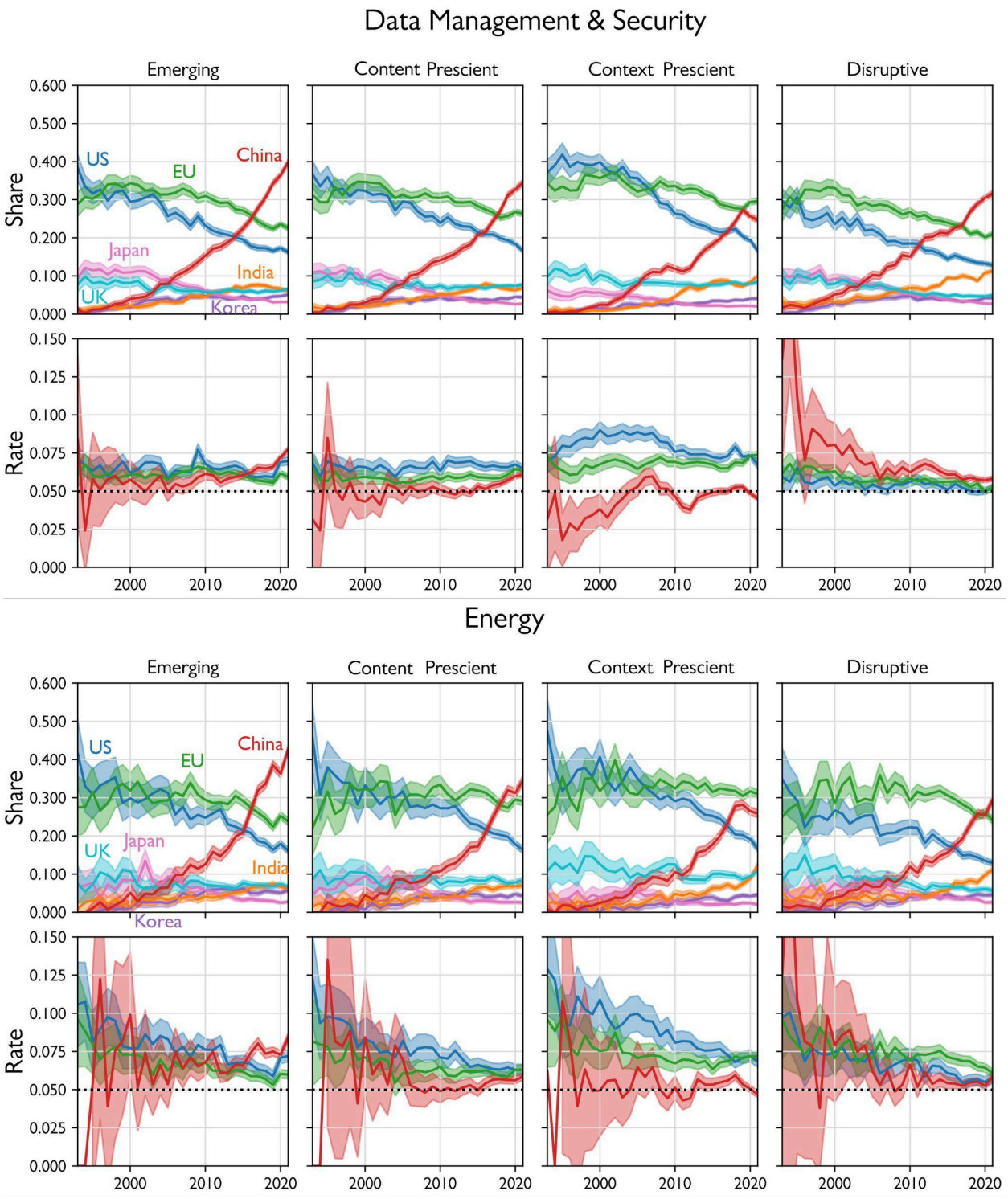


图 S8. 美国国家科学基金会优先研究领域的创新模式。每个面板分别显示与美国国家科学基金会和美国国会认定具有战略重要性的关键技术领域相关的科学领域中，创新论文的占比（上方图表）以及相对于出版物数量的创新率（下方图表）。展示的领域包括：生物医学技术、基因组学和合成生物学；人工智能，

机器学习与自主系统；先进通信与沉浸式技术；灾害预防与减轻；数据管理与安全；以及能源。阴影区域表示95%置信区间。比率图中以虚线水平线标示5%的基准率。结果显示出领域特异性差异：中国在大多数领域的新兴领域中处于领先地位，而美国在前瞻性指标方面仍保持优势，尤其是在人工智能/机器学习和生物医学领域。较小学科领域的置信区间更宽，反映出论文数量较少带来的更大不确定性。

参考文献

1. D. O’Neil，英语作为国际出版的通用语。《世界英语》 **37**，146–165（2018）。2. M. S. Di Bitetti，J. A. Ferreras，以英语发表，否则遭淘汰：在科学出版物中使用英语以外的语言对引用率的影响。《安比奥》 **46**，121–127（2017）。3. B. Perozzi，R. Al-Rfou，S. Skiena，“ DeepWalk：社会表征的在线学习”，载于第*20*届*ACM SIGKDD*国际知识发现与数据挖掘会议论文集（美国计算机协会，纽约，美国纽约州，2014年）*KDD ’14*，第701–710页。4. J. Sourati，J. Evans，利用人类感知人工智能加速科学，*arXiv [*计算机科学-人工智能*]* （2023年）。 http://arxiv.org/abs/2306.01495。