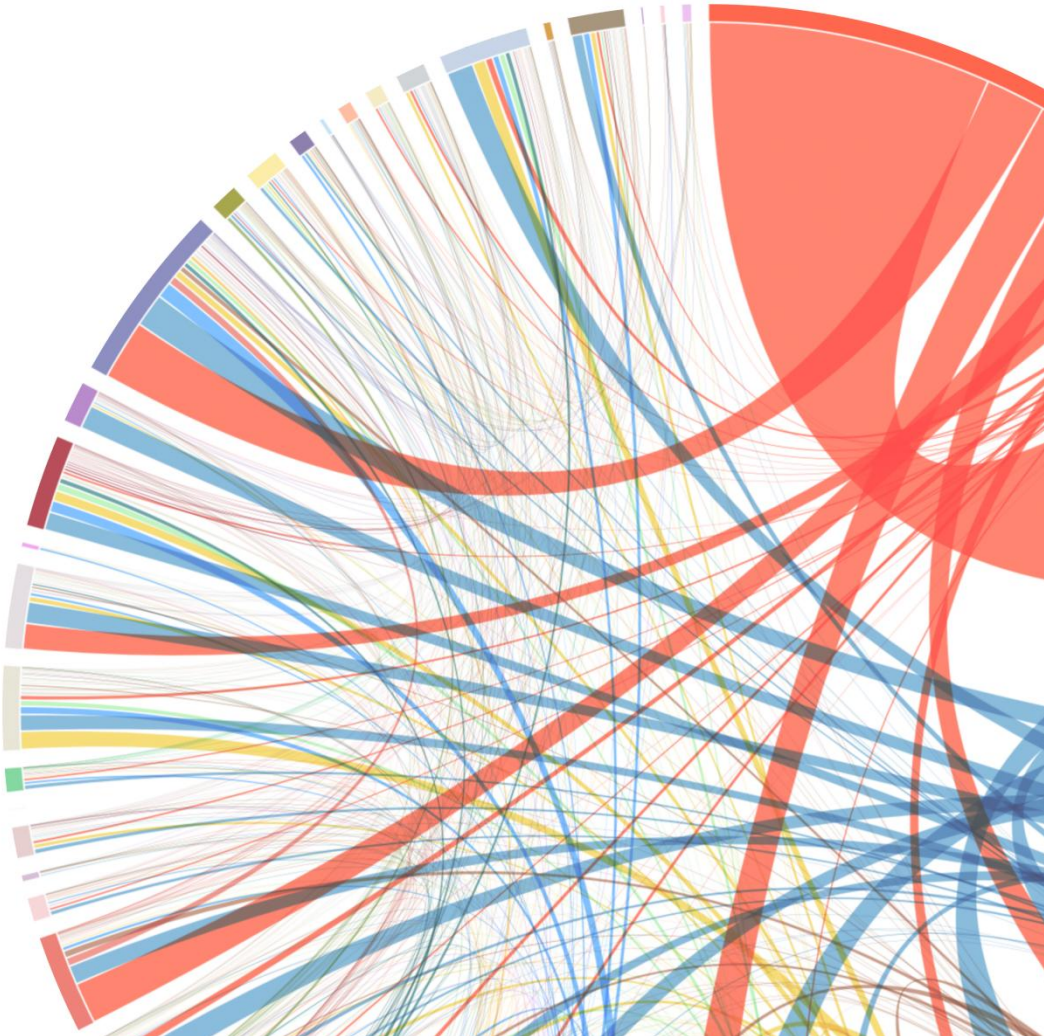
全球
人工智能治理
评估指数

AGILE 指数 2025

2025 年 7 月

- 远期人工智能研究中心 (CLAI)
- 北京前瞻人工智能安全与治理研究院 (Beijing-AISI)
- 人工智能安全与超级对齐北京市重点实验室
- 中国科学院自动化研究所人工智能伦理与治理研究中心



参考引用：

曾毅，鲁恩萌，郭晓阳，皇甫存青，谢佳玮，陈煜，王正奇，梁栋旗，曹功策，王金，阮子喆，关心，Ammar Younas. 全球人工智能治理评估指数（AGILE 指数）2025 [R/OL]. 北京：远期人工智能研究中心，北京前瞻人工智能安全与治理研究院，人工智能安全与超级对齐北京市重点实验室，中国科学院自动化研究所人工智能伦理与治理研究中心，2025. https://agile-index.ai/.

在线平台：

https://agile-index.ai/ （提供支撑数据和持续更新）

发布机构：

远期人工智能研究中心 (CLAI)

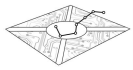
https://long-term-ai.center/

北京前瞻人工智能安全与治理研究院 (Beijing-AISI)

https://beijing-aisi.ac.cn/

人工智能安全与超级对齐北京市重点实验室

https://beijing.safe-ai-and-superalignment.cn/



中国科学院自动化研究所人工智能伦理与治理研究中心

https://ai-ethics-and-governance.institute/

资助信息：

本研究受到新一代人工智能国家科技重大专项 (编号：2022ZD0116202)和北京市科学技术委员会资助。

联系信息：

如需了解更多信息，或有任何意见和反馈，欢迎通过以下方式联系我们：contact@long-term-ai.cn.

报告版本：

本报告正式版本（v1.1.1）发布于 2025 年 7 月 23 日，其预发布版本（v1.0.0-pre）已于 2025 年 7 月 8 日在线公开。

研究团队

首席科学家

曾 毅

中国科学院自动化研究所研究员、人工智能伦理与治理中心主任；人工智能安全与超级对齐北京市重点实验室主任；北京前瞻人工智能安全与治理研究院创始院长；远期人工智能研究中心创始主任。担任国家新一代人工智能治理专委会委员；国家科技伦理委员会人工智能伦理分委员会委员；联合国人工智能高层顾问机构专家；联合国教科文组织人工智能伦理特设专家组专家；世界卫生组织健康领域人工智能伦理与治理专家组专家；世界互联网大会人工智能专委会主任委员。长期从事类脑人工智能，人工智能伦理、安全与治理，人工智能与全球可持续发展研究。

研究人员

鲁恩萌	北京前瞻人工智能安全与治理研究院
郭晓阳	人工智能安全与超级对齐北京市重点实验室
皇甫存青	中国科学院自动化研究所
谢佳玮	人工智能安全与超级对齐北京市重点实验室
陈 煜	人工智能安全与超级对齐北京市重点实验室
王正奇	北京前瞻人工智能安全与治理研究院
梁栋旗	中国科学院自动化研究所
曹功策	北京前瞻人工智能安全与治理研究院
王 金	北京前瞻人工智能安全与治理研究院
阮子喆	中国科学院自动化研究所
关 心	远期人工智能研究中心
Ammar Younas	远期人工智能研究中心

AGILE 指数 2025 更新概览

指标体系与国家覆盖范围的双重扩展：

与 2024 版 AGILE 指数相比，AGILE 指数 2025 在国家覆盖范围上从 14 个国家扩展到 40 个国家，指标数量也从 39 项增加到 43 项。范围的扩展为横向比较和趋势分析提供了更坚实的基础。

数据更新与数据来源扩充：

对 2024 版 AGILE 指数涉及的所有数据进行了彻底审查和更新，其中包含截至 2025 年 3 月可获得的最新信息。此外还整合了 20 多个新的数据来源，以增强评估的覆盖范围和可靠性。

前沿人工智能技术与研究主题分析：

报告新增了生成式人工智能等前沿人工智能技术领域的最新数据，同时文献分析范围扩展到更广泛的人工智能安全与治理主题领域，通过更加细化的领域关键词对研究进展进行了更全面的统计。

年度对比分析：

对于数据定义保持一致的指标，报告增强了逐年比较的分析，以便更好地捕捉趋势并突出随时间的关键变化。

缺失数据插补策略优化：

针对国家覆盖范围的扩大以及对数据质量和指数评分的更严格要求，缺失数据的插补策略也得到了进一步完善。当前插补方法结合了历史表现趋势与指标间的相关性，以确保更高的合理性与一致性。详细内容请参见附录 2 中的“方法论”部分。

执行摘要

人工智能（AI）既是一场前所未有的技术革命，也是全球面临的系统性治理挑战。2024年全球人工智能治理进程显著加速，多边框架持续强化，各国监管举措密集涌现。这一态势进一步凸显出系统性追踪和持续评估全球治理进展的重要性和迫切性。

正是基于此背景，全球人工智能治理评估指数（AI Governance InternationaL Evaluation Index，简称 AGILE 指数）项目于 2023 年启动。首版 AGILE 指数于 2024 年 2 月发布，覆盖了 14 个评估国家，初步建立起了具有可操作性和可比较性的基准评估框架。

在此基础上，全球人工智能治理评估指数 2025（简称 AGILE 指数 2025）在 2024 版基础上进行了系统性优化，进一步平衡了科学严谨性与实践适应性。更新后的指数评估过程在扩大数据多样性的同时，增强了指标的有效性和跨国可比性。AGILE 指数 2025 的评估框架包括 4 大评估方面、17 个维度和 43 项指标，同时紧扣研究进展与政策演化趋势，对 40 个国家进行了评估，评估国家范围涵盖了不同的收入水平、地理区域以及技术发展阶段。在指数评估过程中，团队整合了多源证据，包括政策文件、治理实践、科研产出及风险事件等，构建了统一的比较基准，来更加客观地刻画全球人工智能治理现状，帮助各国识别优势、短板与系统性约束。

通过持续的迭代优化，我们期待 AGILE 指数能够不断推动全球人工智能治理的透明化与可衡量化，为各国绘制数据驱动的人工智能治理能力画像，协助其精准识别自身治理水平与面临的关键挑战，并最终为国家及全球层面人工智能治理体系的建设提供切实可行的洞见。

AGILE 指数 2025 主要发现

总体观察

1. 根据 AGILE 指数 2025 的得分，40 个被评估国家总体可划分为三个梯队。其中，在 AI 发展水平（P1）与 AI 治理工具（P3）两个评估方面上，各国的表现差异相对较大。（第 14 页）
2. 在 2024 年评估的 14 个国家中，相对排名变化呈现出高排名国家动态变化、低排名国家相对稳定的格局，其中值得注意的变化是美国和中国之间的位置互换。（第 16 页）
3. 在所评估的 40 个国家中，AGILE 指数总分与人均 GDP 之间仍然呈现总体正相关关系。相对滞后的国家需要提升 AI 治理准备水平，以增强应对能力。（第 16 页）
4. 各个国家在 AGILE 指数不同评估方面上的表现呈现出四种不同的 AI 发展与治理类型：全面领先型、治理超前型、治理滞后型和基础建设型。（第 17 页）
5. 高收入国家组在 AI 发展水平（P1）和 AI 治理工具（P3）方面明显优于中高和中低收入国家组，而后者的 AI 风险暴露度更低且社会对 AI 的接受度普遍更高。（第 19 页）

AI 发展水平

1. 美国和中国在 AI 发展水平方面各自展现出独特优势，其他国家也通过差异化优势一同丰富了全球 AI 生态。（第 21 页）
2. 自 2010 年以来，生成式人工智能全球专利数量呈现出指数级增长。在此期间，生成式人工智能专利总量增长约 30 倍，生成式人工智能应用专利数量增长约 25 倍。2018 年之后增长速度显著加快。（第 23 页）

AI 治理环境

1. 2024 年记录在案的 AI 风险事件数量较 2023 年大幅上升，增长了约 100%。其中，美国的 AI 风险事件数量增长了 1.8 倍，而其他国家的平均增长速度更快，达到 2.1 倍。（第 24 页）
2. 与 AI 内部安全与外部安全、AI 与人权以及数据治理相关的风险事件数量最多，占所有记录在案的 AI 风险事件数量的一半。（第 25 页）

3. 尽管高收入国家往往在整体治理方面得分更高——这表明它们在 AI 治理方面具有更强的内在准备度和基础，但实际经验仍显示，在不同国家背景下积极推进 AI 治理仍有很大的努力空间。（第 26 页）

AI 治理工具

1. 所有 40 个被评估的国家都已发布了国家层面的 AI 战略，但在战略制定的结构化方法上存在差异。（第 28 页）
2. 自 2024 年以来，AI 立法呈现出明显的加速趋势。一些国家已经制定了国家层面专门针对 AI 的综合性法律法规，还有一些国家则制定了针对 AI 垂直领域的法律法规。（第 29 页）
3. 40 个国家都以多种形式参与了全球 AI 治理的相关机制，其中法国、日本、韩国和新加坡的参与度最高。（第 30 页）

AI 治理成效

1. 公众对 AI 应用与服务的认知度、信任度与积极态度总体上与各国人均 GDP 呈负相关。（第 32 页）
2. 越来越多的女性参与到了 AI 相关的研究中，活跃研究者中男女比例持续下降，到 2025 年已降至 2.2:1。（第 34 页）
3. 以人均 GDP 为代表的经济发展水平在一定程度上与社会弱势群体的数字普惠程度呈正相关，但较高的经济发展水平并不必然带来高水平的数字包容性。（第 35 页）
4. 在有影响力的开放 AI 模型与数据集方面，中国和美国合计占评估所覆盖 40 个国家总量的 70%以上。（第 36 页）
5. 北美和西欧在开源 AI 生态上占据主导，美国以占 40 国总量 30.43%的贡献遥遥领先。与此同时，中国和印度也展现出强劲的参与度。（第 37 页）
6. 对 AI 治理相关议题的研究关注持续上升，相关出版物在全部 AI 出版物中的占比在 2024 年已达到 14.4%。在所评估的 40 个国家中，来自中国与美国的 AI 治理相关研究占比合计达到 54%，贡献超过一半。（第 38 页）
7. 在 AI 治理研究的国际合著中，中国、美国、加拿大、德国、英国和澳大利亚之间的合

著最为普遍，在 40 个评估国家的所有合著出版物中占比约 70%。（第 40 页）

8. 美国和中国在面向可持续发展目标（SDGs）的 AI 研究中处于主导地位，在 40 个国家中共同贡献了约 60%的相关出版物。与此同时，其他国家也在不同可持续发展目标领域做出了重要贡献。（第 41 页）
9. SDG3（良好健康与福祉）、SDG11（可持续城市和社区）以及 SDG9（产业、创新和基础设施）是当前研究关注度最高的 AI 赋能可持续发展目标方向。同时，高收入国家组在社会层面的目标方向上分配了更高的 AI 研究关注度，而中高收入和中低收入国家组在经济、环境相关的目标上有着相对更高的 AI 研究关注度。（第 42 页）

目录

I.
AGILE 指数

1.1. 理论框架

全球人工智能治理评估指数（AGILE 指数）以“治理水平应同发展水平相匹配”为核心设计原则。随着 AI 经历不同的创新阶段，治理机制应实现动态调整，在技术进步与社会治理之间建立协同共生的关系。AGILE 指数的评估框架旨在促进 AI 价值创造加速与新兴风险防控之间的平衡，确保治理成熟度能够与技术复杂性同步提升。

AGILE 指数 2025 采用了多层次的评估体系，包括：

- 4 大评估方面（Pillars）：AI 发展水平、AI 治理环境、AI 治理工具以及 AI 治理成效。
- 17 个评估维度（Dimensions）：每个评估维度聚焦具体的治理重心，将宏观原则与评估目标相衔接。
- 43 项评估指标（Indicators）：提供细化的指标项，将评估目标转化为可量化的证据和可操作的评价标准。

图 1 AGILE 指数 2025：评估方面与评估维度

II.
评估概览

2.1. 评分结果

表 1 AGILE 指数 2025：各国总分、评估方面得分和维度分

国家	总分	排名	D1	D2	D3	P1	D4	D5	P2	D6	D7	D8	D9	D10	D11	D12	P3	D13	D14	D15	D16	D17	P4
中国	70.1	1	75.5	50.6	55.4	60.5	48.5	68.7	58.6	100.0	100.0	100.0	0.0	100.0	83.3	87.5	81.5	100.0	99.5	45.4	91.0	63.8	79.9
美国	69.9	2	82.6	100.0	100.0	94.2	0.0	80.0	40.0	100.0	100.0	100.0	100.0	100.0	50.0	87.5	91.1	48.2	21.7	27.0	100.0	74.7	54.3
德国	68.8	3	52.5	78.8	43.6	58.3	73.0	86.7	79.8	100.0	100.0	100.0	0.0	100.0	100.0	93.5	84.8	27.8	42.9	56.0	77.3	57.3	52.2
韩国	67.8	4	59.8	46.0	43.3	49.7	58.5	86.0	72.3	100.0	100.0	100.0	100.0	100.0	50.0	99.5	92.8	96.4	51.0	50.6	45.8	38.8	56.5
英国	67.6	5	63.8	64.3	74.2	67.4	12.5	86.5	49.5	100.0	100.0	100.0	100.0	0.0	83.3	93.5	96.7	55.4	32.6	60.6	75.0	60.2	56.8
新加坡	65.7	6	49.7	60.6	74.7	61.6	43.5	83.0	63.2	75.0	100.0	100.0	100.0	100.0	100.0	99.5	70.8	71.1	74.0	82.6	63.8	44.7	67.2
法国	65.6	7	38.9	66.2	50.2	51.8	57.5	84.6	71.1	100.0	100.0	100.0	100.0	100.0	100.0	99.5	99.9	16.7	27.7	51.3	67.3	34.8	39.6
加拿大	63.5	8	49.8	55.0	56.7	53.8	37.0	83.4	60.2	75.0	100.0	100.0	100.0	100.0	50.0	93.5	88.4	57.3	27.3	59.5	61.5	51.7	51.5
日本	61.4	9	47.2	72.0	24.0	47.7	72.0	84.9	78.5	100.0	100.0	100.0	100.0	0.0	16.7	99.5	73.7	77.8	31.1	43.8	42.5	33.2	45.7
芬兰	60.3	10	40.7	89.6	65.2	65.2	82.5	87.8	85.2	100.0	100.0	0.0	0.0	0.0	66.7	57.5	46.3	95.7	13.5	57.2	18.3	38.2	44.6
荷兰	58.9	11	36.3	73.1	37.8	49.1	68.0	85.1	76.5	100.0	0.0	0.0	100.0	0.0	100.0	87.5	55.4	58.7	34.5	69.7	68.8	42.5	54.8
阿联酋	56.4	12	25.2	47.6	67.1	46.6	70.5	77.4	73.9	50.0	100.0	100.0	100.0	0.0	33.3	23.5	58.1	41.4	55.8	83.8	24.3	28.5	46.8
瑞典	56.0	13	41.2	63.6	92.4	65.7	76.0	85.9	81.0	75.0	100.0	0.0	0.0	0.0	66.7	57.5	42.7	33.9	24.5	55.1	26.5	32.0	34.4
澳大利亚	55.0	14	48.1	65.7	20.5	44.8	37.5	85.1	61.3	75.0	100.0	100.0	0.0	100.0	50.0	87.5	73.2	39.7	22.9	29.9	47.3	63.5	40.6
丹麦	54.8	15	37.4	68.9	53.1	53.1	72.5	88.8	80.6	100.0	0.0	0.0	0.0	0.0	66.7	57.5	32.0	87.2	53.5	74.8	24.0	28.2	53.5
瑞士	54.8	16	55.3	98.3	35.4	63.0	58.5	84.7	71.6	75.0	0.0	100.0	0.0	0.0	0.0	79.5	36.4	46.0	35.8	75.4	52.8	31.5	48.3
挪威	54.7	17	34.4	55.7	45.0	45.0	79.0	86.4	82.7	100.0	0.0	0.0	0.0	0.0	100.0	57.5	36.8	83.8	10.3	88.3	48.3	41.0	54.3
沙特阿拉伯	54.5	18	18.4	35.8	27.1	27.1	94.0	74.3	84.1	75.0	100.0	100.0	100.0	100.0	33.3	79.5	84.0	11.0	22.9	22.3	11.3	47.2	22.9
意大利	53.8	19	27.1	79.0	24.7	43.6	66.0	79.2	72.6	100.0	100.0	0.0	0.0	0.0	100.0	87.5	55.4	26.9	57.2	46.8	33.3	54.7	43.7
西班牙	51.7	20	30.6	58.6	20.7	36.6	66.0	83.7	74.9	100.0	100.0	0.0	0.0	0.0	66.7	79.5	49.5	37.6	44.6	60.4	38.8	47.3	45.7
马来西亚	51.2	21	35.3	57.2	46.3	46.3	66.0	72.2	69.1	100.0	100.0	100.0	0.0	0.0	0.0	50.0	50.0	17.4	72.0	66.8	9.5	32.2	39.6
印度尼西亚	51.0	22	8.6	39.3	24.0	24.0	89.5	67.5	78.5	100.0	100.0	100.0	0.0	0.0	0.0	29.5	47.1	53.2	96.1	68.8	17.8	36.8	54.5
土耳其	46.7	23	6.9	20.8	13.8	13.8	94.0	70.9	82.5	100.0	100.0	100.0	0.0	0.0	16.7	79.5	56.6	52.0	67.6	16.5	7.5	25.3	33.8
爱尔兰	46.6	24	34.8	57.4	46.1	46.1	47.0	83.1	65.1	100.0	0.0	0.0	0.0	0.0	66.7	73.5	34.3	49.4	32.5	67.5	16.5	39.3	41.0
泰国	45.8	25	25.6	19.0	22.3	22.3	80.5	72.6	76.6	100.0	0.0	100.0	0.0	0.0	16.7	7.5	32.0	31.5	96.6	81.8	12.5	40.0	52.5
葡萄牙	45.4	26	29.9	27.0	28.4	28.4	84.5	79.6	82.1	100.0	0.0	0.0	0.0	0.0	66.7	57.5	32.0	70.1	21.8	49.0	12.5	41.8	39.0
以色列	45.0	27	42.5	48.7	69.5	53.5	28.5	75.6	52.0	100.0	0.0	100.0	0.0	0.0	33.3	57.5	41.5	47.1	33.0	31.0	27.3	26.5	33.0
比利时	45.0	28	23.8	28.6	26.2	26.2	61.0	80.2	70.6	100.0	100.0	0.0	0.0	0.0	66.7	57.5	46.3	53.1	21.9	54.5	36.0	18.8	36.9
智利	44.9	29	14.3	52.3	33.3	33.3	93.5	76.4	84.9	100.0	0.0	0.0	0.0	0.0	50.0	29.5	25.6	25.7	56.5	61.3	11.8	24.5	35.9
俄罗斯	44.1	30	18.2	32.0	25.1	25.1	41.0	65.0	53.0	100.0	100.0	100.0	0.0	100.0	66.7	50.0	73.8	47.5	24.6	2.4	31.3	17.3	24.6
墨西哥	43.8	31	12.2	10.2	11.2	11.2	89.5	66.1	77.8	75.0	100.0	100.0	0.0	0.0	16.7	29.5	45.9	72.1	84.3	13.8	10.8	20.0	40.2
波兰	43.1	32	28.2	42.4	35.3	35.3	73.0	78.1	75.6	75.0	0.0	0.0	0.0	0.0	66.7	57.5	28.5	60.0	37.3	24.6	25.8	18.2	33.2
新西兰	41.8	33	29.2	69.2	49.2	49.2	35.0	86.7	60.8	75.0	0.0	100.0	0.0	0.0	0.0	73.5	35.5	23.9	37.1	5.7	16.8	26.0	21.9
匈牙利	41.3	34	29.2	15.8	22.5	22.5	88.0	75.0	81.5	100.0	0.0	0.0	0.0	0.0	66.7	57.5	32.0	35.1	53.8	36.3	11.0	9.8	29.2
印度	41.0	35	27.9	31.3	40.8	33.3	23.5	65.0	44.2	50.0	0.0	0.0	100.0	0.0	0.0	79.5	32.8	45.9	68.0	40.8	63.5	49.7	53.6
巴西	40.8	36	16.9	30.9	16.2	21.3	83.5	72.5	78.0	100.0	0.0	0.0	0.0	0.0	50.0	79.5	32.8	33.1	52.7	6.1	31.0	31.7	30.9
秘鲁	40.8	37	22.4	13.8	18.1	18.1	94.0	68.2	81.1	75.0	0.0	0.0	0.0	0.0	16.7	7.5	14.2	53.3	88.0	76.9	3.5	26.7	49.7
阿根廷	37.2	38	7.3	11.1	9.2	9.2	84.5	67.9	76.2	100.0	0.0	100.0	0.0	0.0	0.0	7.5	29.6	7.1	62.4	82.8	10.0	6.7	33.8
哥伦比亚	37.1	39	13.0	17.8	15.4	15.4	90.5	68.7	79.6	100.0	0.0	0.0	0.0	0.0	50.0	7.5	22.5	43.4	69.0	0.0	20.3	21.8	30.9
南非	33.9	40	5.4	28.9	17.1	17.1	58.0	63.6	60.8	75.0	0.0	0.0	0.0	0.0	0.0	0.0	10.7	40.6	75.4	73.3	5.0	39.8	46.8

2.2. 总体观察

主要发现 1：

根据 AGILE 指数 2025 的得分，40 个被评估国家总体可划分为三个梯队。其中，在 AI 发展水平（P1）与 AI 治理工具（P3）两个评估方面上，各国的表现差异相对较大。

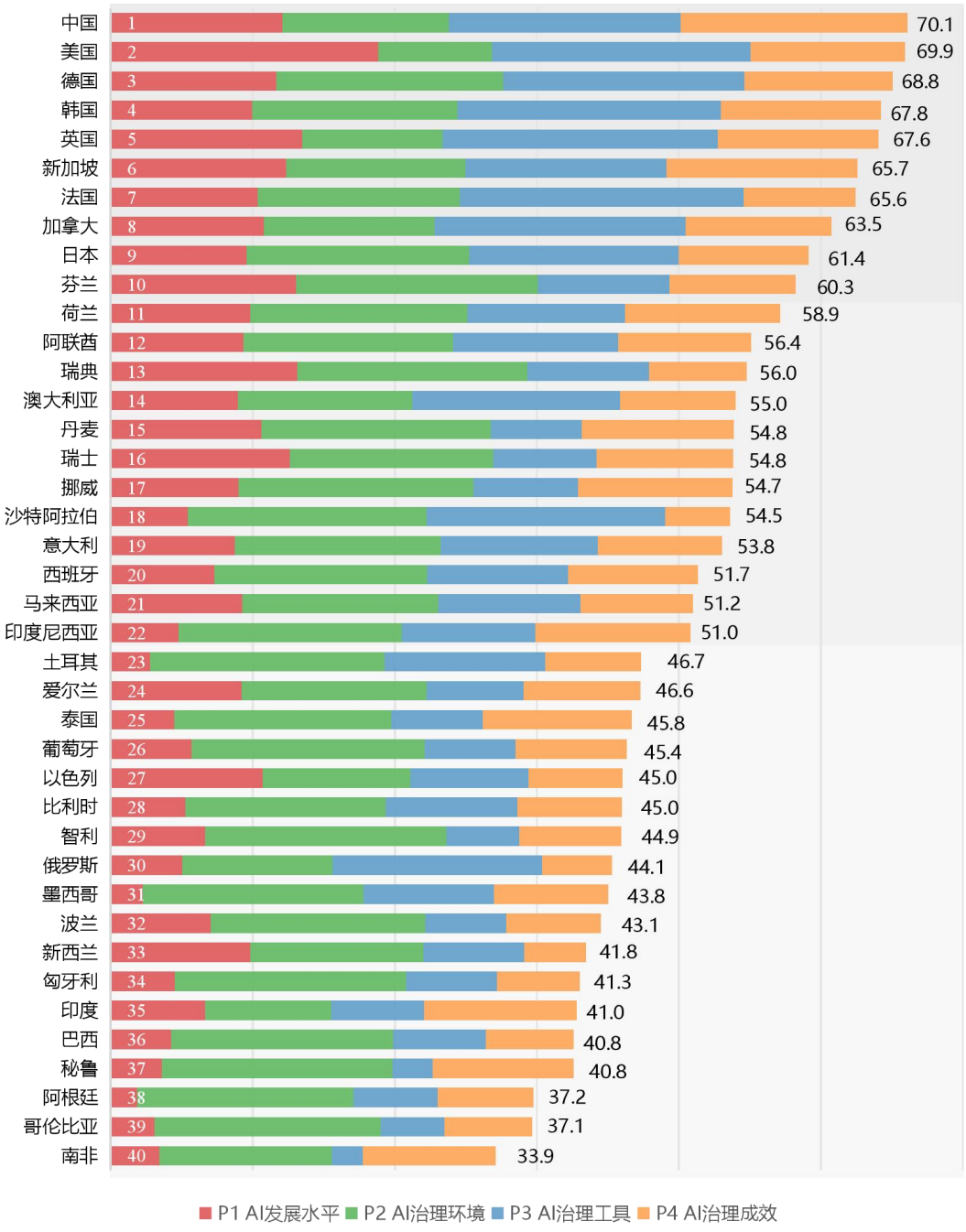


图 2 AGILE 指数 2025：各国总分与排名

根据 AGILE 指数 2025 的得分分布，40 个被评估国家可以大致分为三个梯队。第一梯队国家的得分高于 60，包括中国、美国和德国等；第二梯队国家的得分在 50 到 60 之间，例如荷兰、阿拉伯联合酋长国和瑞典等；第三梯队国家的得分则是低于 50。

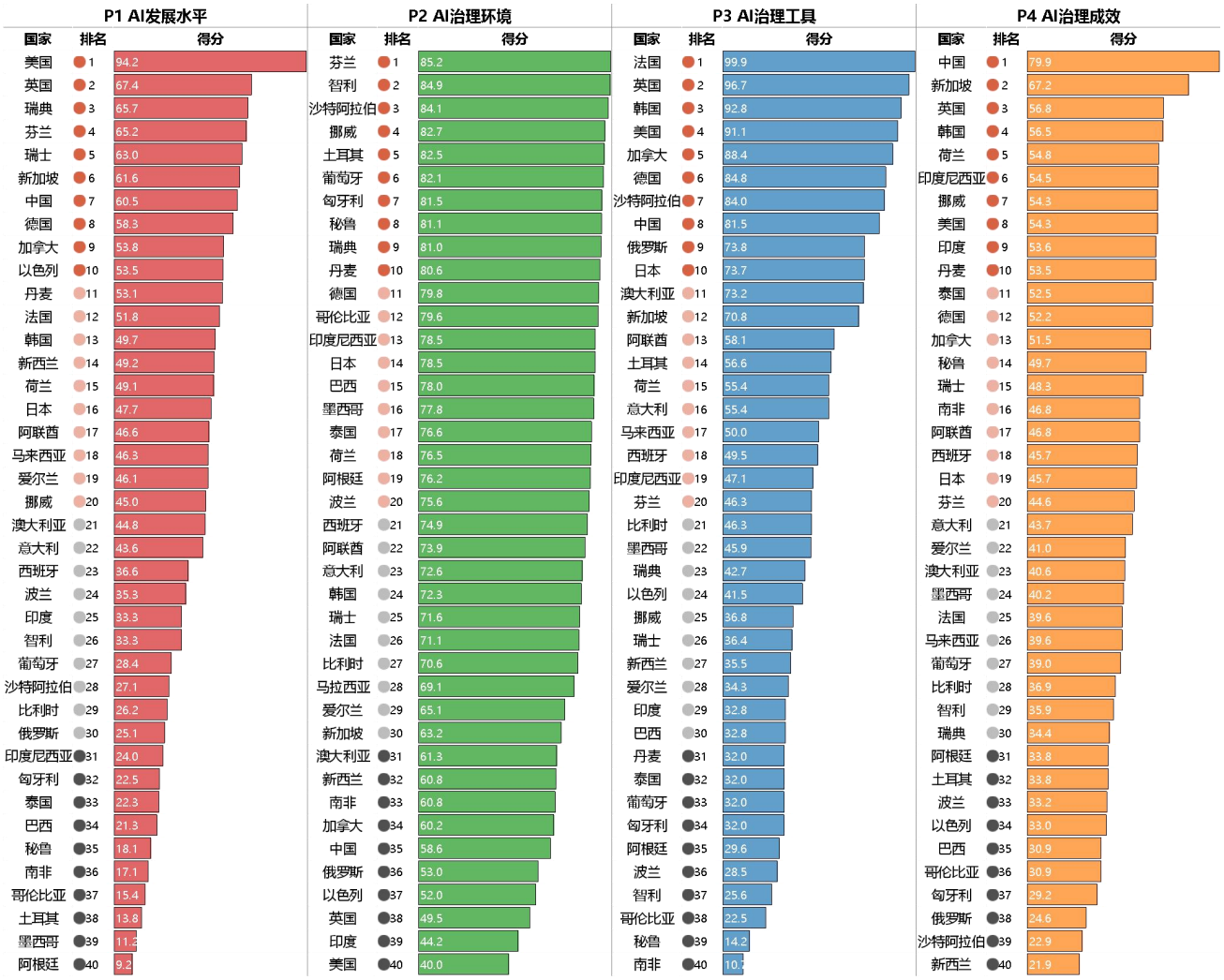


图 3 AGILE 指数 2025：各国在不同评估方面上的得分与排名

在四个评估方面中，P1（AI 发展水平）和 P3（AI 治理工具）的得分分布明显比 P2（AI 治理环境）和 P4（AI 治理成效）更为分散，显示出更显著的分层差异。领先国家与落后国家之间的得分差距在 P1 和 P3 中尤为显著（P1 最高分：94.2，最低分：9.2；P3 最高分：99.9，最低分：10.7）。相比之下，P2 和 P4 的得分分布更为集中，不同梯队之间的差异比 P1 和 P3 中观察到的要小。

主要发现 2：

在 2024 年评估的 14 个国家中，相对排名变化呈现出高排名国家动态变化、低排名国家相对稳定的格局，其中值得注意的变化是美国和中国之间的位置互换。

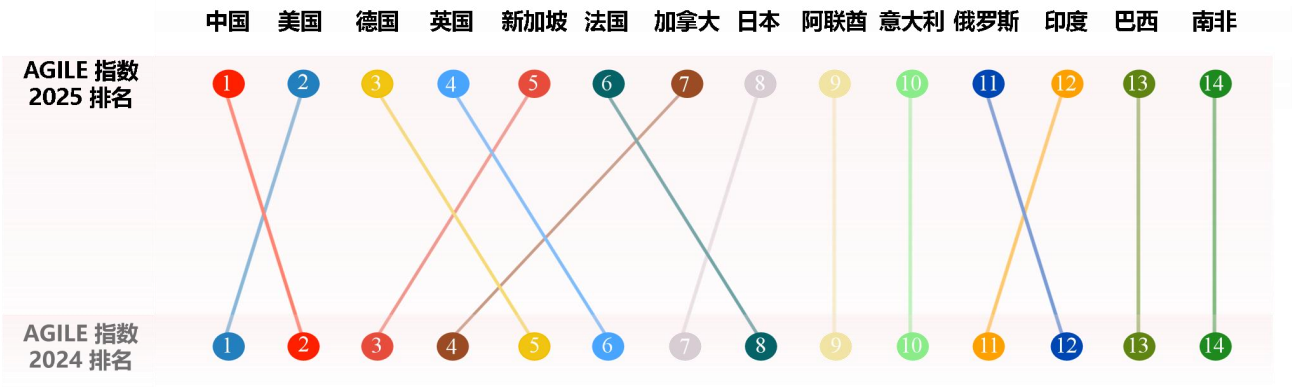


图 4 AGILE 指数 2024 评估的 14 个国家在此次评估中的相对排名变化

2024 年所评估的 14 个国家在 AGILE 指数 2025 中的相对排名发生了变化。排名靠前的国家出现了明显的位次变动，而排名靠后的国家则相对稳定。美国的相对排名下滑至第二位，主要是由于其在 AI 立法方面趋于宽松的政策趋势所产生的影响。随着美国第 14179 号行政命令——《消除美国在人工智能领域领导地位的障碍》的出台，规定废除第 14110 号行政命令——《安全、可靠和可信地开发和使用人工智能》。然而，新的行政命令并不具备等同于第 14110 号行政命令的综合性，这导致其在相关评估指标上的得分受到影响。与此同时，中国凭借更为连贯且稳定的 AI 治理政策上升至第一位。一些国家，如新加坡和加拿大，其相对排名出现了一些波动，而巴西和南非等国家的排名则相对稳定。总体来看，排名前列的国家波动较大，而排名后段的国家则变化不大。

主要发现 3：

在所评估的 40 个国家中，AGILE 指数总分与人均 GDP 之间仍然呈现总体正相关关系。相对滞后的国家需要提升 AI 治理准备水平，以增强应对能力。

本年度的 AGILE 指数继续显示出人均 GDP 与总指数得分之间的正相关关系。总体而言，人均 GDP 较高的国家往往在 AGILE 指数上得分更高，表明经济发展与 AI 治理的加强可能存在关联。然而，随着评估范围的扩大，数据也呈现出一些重要的例外情况：部分人均 GDP 较

低的国家在治理方面表现出相对较高的得分，而一些经济较为发达的国家却得分不高。这说明，尽管发展是治理的必要基础，但它并不是决定 AI 治理水平的唯一因素。

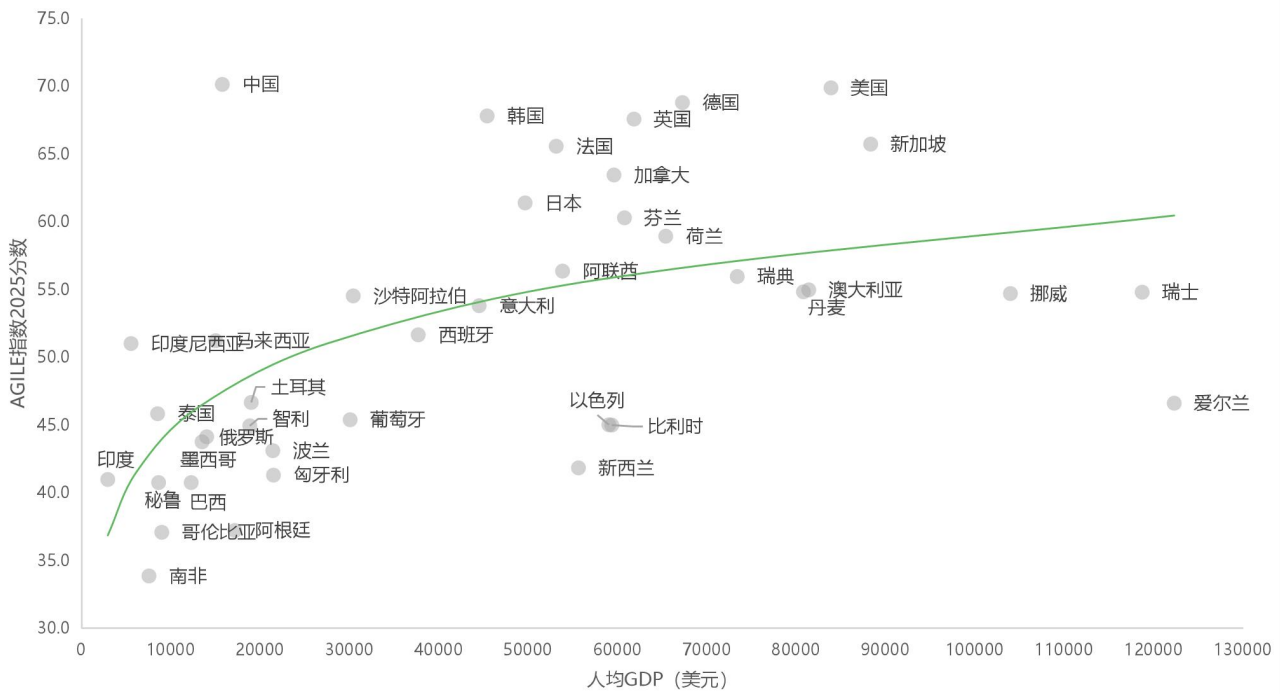


图 5 40 国 AGILE 指数 2025 得分与其人均 GDP 水平总体呈正相关关系

主要发现 4：

各个国家在 AGILE 指数不同评估方面上的表现呈现出四种不同的 AI 发展与治理类型：全面领先型、治理超前型、治理滞后型和基础建设型。

对 40 个国家在 AGILE 指数 2025 四个评估方面得分的进一步分析，可以看出形成了四种不同的发展与治理类型：

- **全面领先型：**包括美国、新加坡、中国和英国在内的国家在四个评估方面上均表现出色、得分均衡，展现出较为全面的发展与治理能力。
- **治理超前型：**与第一类国家相比，法国、韩国和加拿大等国家在治理环境和治理工具方面保持较高水平，但在 AI 研发和治理成效方面相对滞后。
- **治理滞后型：**爱尔兰、以色列和新西兰等国家表现出 AI“发展快、治理慢”的不匹配状态。这些国家在 AI 研究和应用方面取得了快速进展，但在治理工具开发和政策框架建设等领域的表现相对滞后。

- **基础建设型：**以印度和南非等国家为代表，这类国家在四个评估方面上的得分普遍均较低，尤其在 P1（AI 发展水平）和 P3（AI 治理工具）方面存在明显短板，反映出其在 AI 发展水平和治理工具方面都亟需加强基础建设。

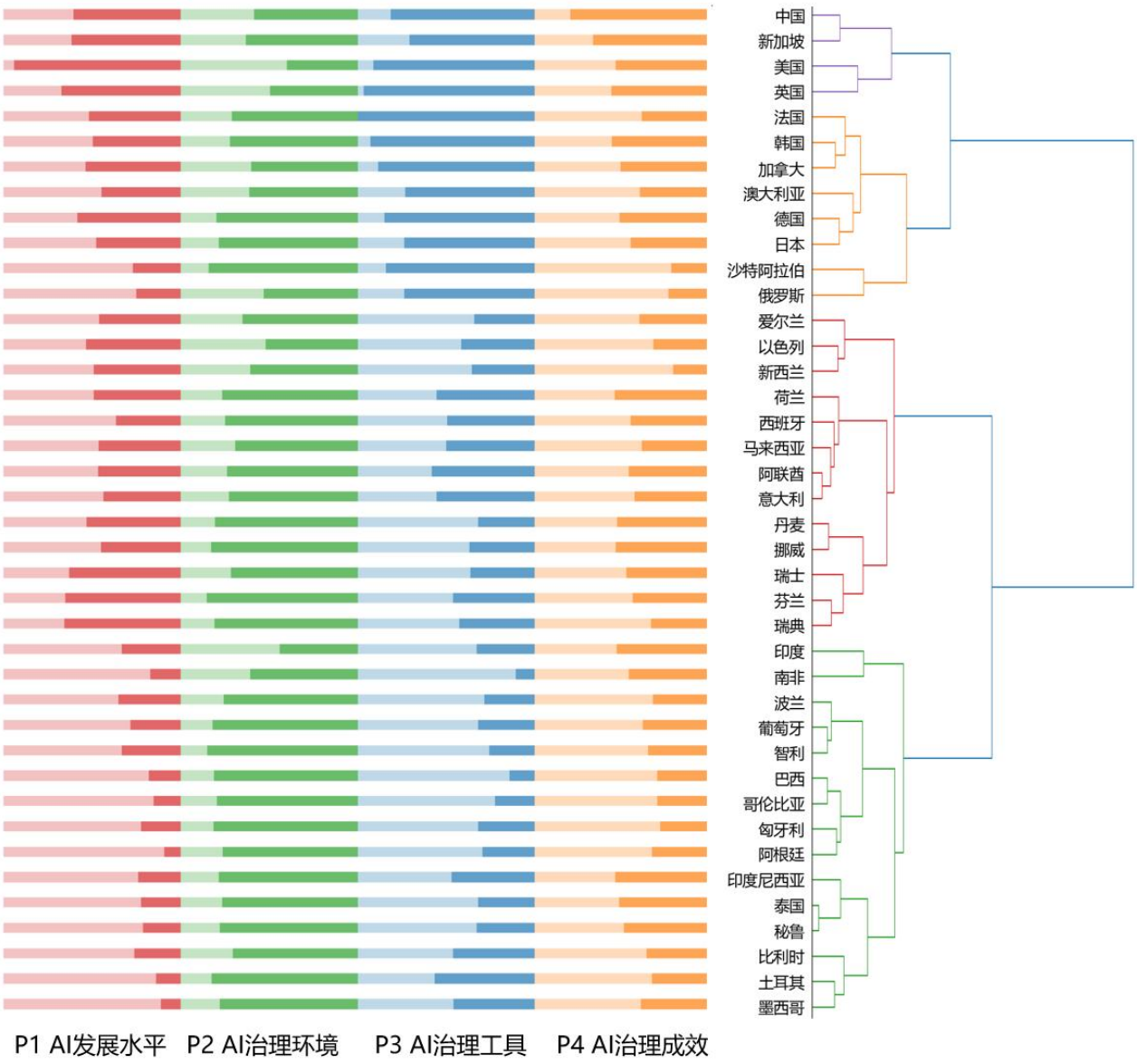


图 6 40 国 AGILE 指数 2025 各评估方面得分的分布情况及国家类型

主要发现 5：

高收入国家组在 AI 发展水平（P1）和 AI 治理工具（P3）方面明显优于中高和中低收入国家组，而后者的 AI 风险暴露度更低且社会对 AI 的接受度普遍更高。

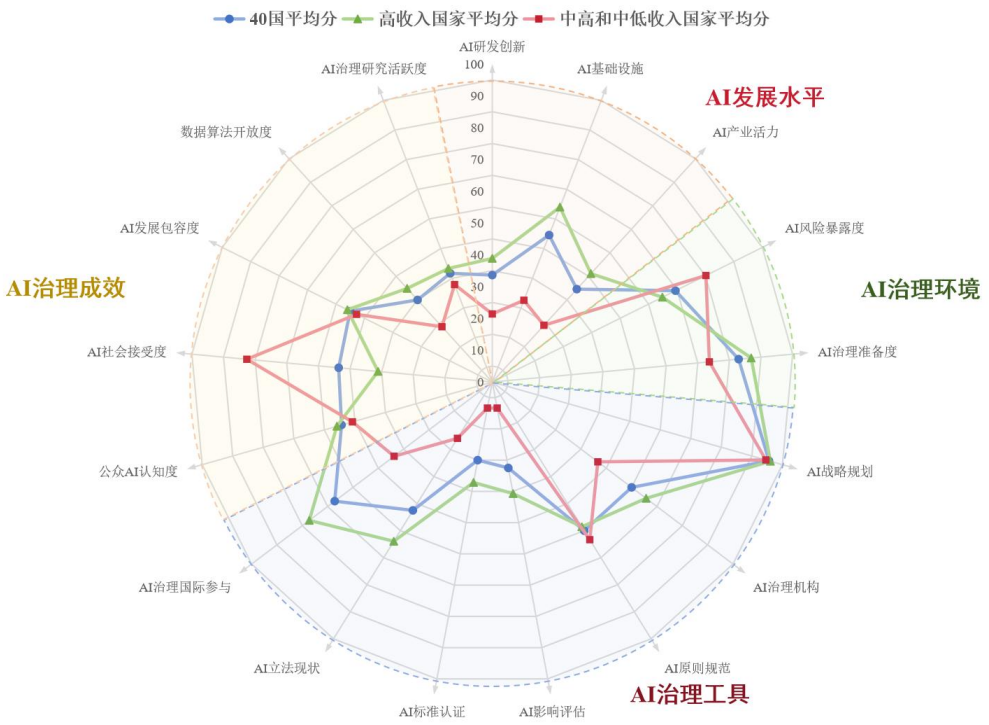


图 7 40 国、高收入国家组、中高和中低收入国家组在各评估维度上的平均分

如果将 40 个国家按照人均收入水平划分为高收入国家、中高和中低收入国家两组[1]，则高收入国家组在四个评估方面——AI 发展水平（P1）、AI 治理环境（P2）、AI 治理工具（P3）以及 AI 治理成效（P4）——的平均得分分别为 48、71、58 和 43；中高和中低收入国家组在上述四个方面上的平均得分分别为 24、72、38 和 46。

相比之下，高收入国家组在 P1 和 P3 上的表现显著优于中高和中低收入国家，这反映了其在技术进步和治理基础设施方面的优势。然而，在 AI 治理环境（P2）和 AI 治理成效（P4）上，中高和中低收入国家组略占优势，这主要得益于其较低的人工智能风险暴露程度以及更高的社会接受度，这些因素有助于提升其治理环境与治理成效的相对表现。

[1] 注：本报告中高收入国家、中高收入国家和中低收入国家的划分依据参照世界银行的收入分类标准。
https://blogs.worldbank.org/zh/opendata/world-bank-country-classifications-by-income-level-for-2024-2025

III.
分析与发现

3.1. 评估方面 1：AI 发展水平

评估方面 1 概览：

AGILE 指数从三个维度评估 AI 发展水平：AI 研发创新、AI 基础设施和 AI 产业活力。

表 2 40 国 AI 发展水平总体情况

42万+	20万+	1.6万+	375	1.15万+	8千+
发表过AI相关出版物的 **研究者**	新发表的AI相关 **出版物**	新授权的AI **专利**	已开发的大规模AI **系统**	全球500强榜单各国上榜 **超级计算机** 的 RMAX 总体性能 [EFlop/s]	**数据中心**
2024 年4 月至2025 年3 月期间			*截至2025 年3 月*		

在所评估的 40 个国家中，过去一年内共有超过 42 万名研究人员发表了超过 20 万篇与 AI 相关的出版物，并且有超过 1.6 万项专利获得授权。截至 2025 年 3 月，这些国家已经开发了 375 个大型 AI 系统，拥有累计超过 11,500 EFlop/s 的超级计算机算力，并拥有 8,000 余个托管数据中心来支持 AI 相关的研发活动。

主要发现 1. 1：

美国和中国在 AI 发展水平方面各自展现出独特优势，其他国家也通过差异化优势一同丰富了全球 AI 生态。

美国和中国在 AI 发展方面各有优势。两国在过去一年中的 AI 相关出版物数量、AI 领域活跃研究人员数量以及开发的大型 AI 系统数量中占比约 60%。与此同时，中国在 AI 专利授权数量方面保持领先地位，而美国则在超级计算机和数据中心基础设施数量、私人投资金额以及新获融资的 AI 企业数量方面占据主导地位。

此外，英国、德国、日本和加拿大通过各自的差异化优势在 AI 生态中占据重要位置：德国在 AI 研发创新（占 AI 相关出版物的 5.7%，活跃 AI 研究人员的 5.58%）和国家超级计算机总浮点运算能力（占 Top500 超级计算机性能总和的 3.43%）方面表现出色；英国在 AI 研发创

新和 AI 基础设施（占大型 AI 系统的 5.33%，数据中心的 4.69%）以及 AI 产业活力（占 AI 私人投资的 3.16%，新获融资的 AI 企业的 6.72%）方面表现出色；日本在国家超级计算机总浮点运算能力方面（占 Top500 超级计算机性能总和的 8.08%）展现出优势。加拿大在 AI 发展水平方面表现均衡（占出版物的 2.97%，研究人员的 2.14%，数据中心的 3.15%，私人投资的 2.02%，以及新获融资的 AI 企业的 3.65%）。与此同时，印度在金砖国家中表现突出：在 8 项 AI 发展水平评估指标中，有 5 项指标在金砖国家中位列第二，在这些领域有着显著的影响力。

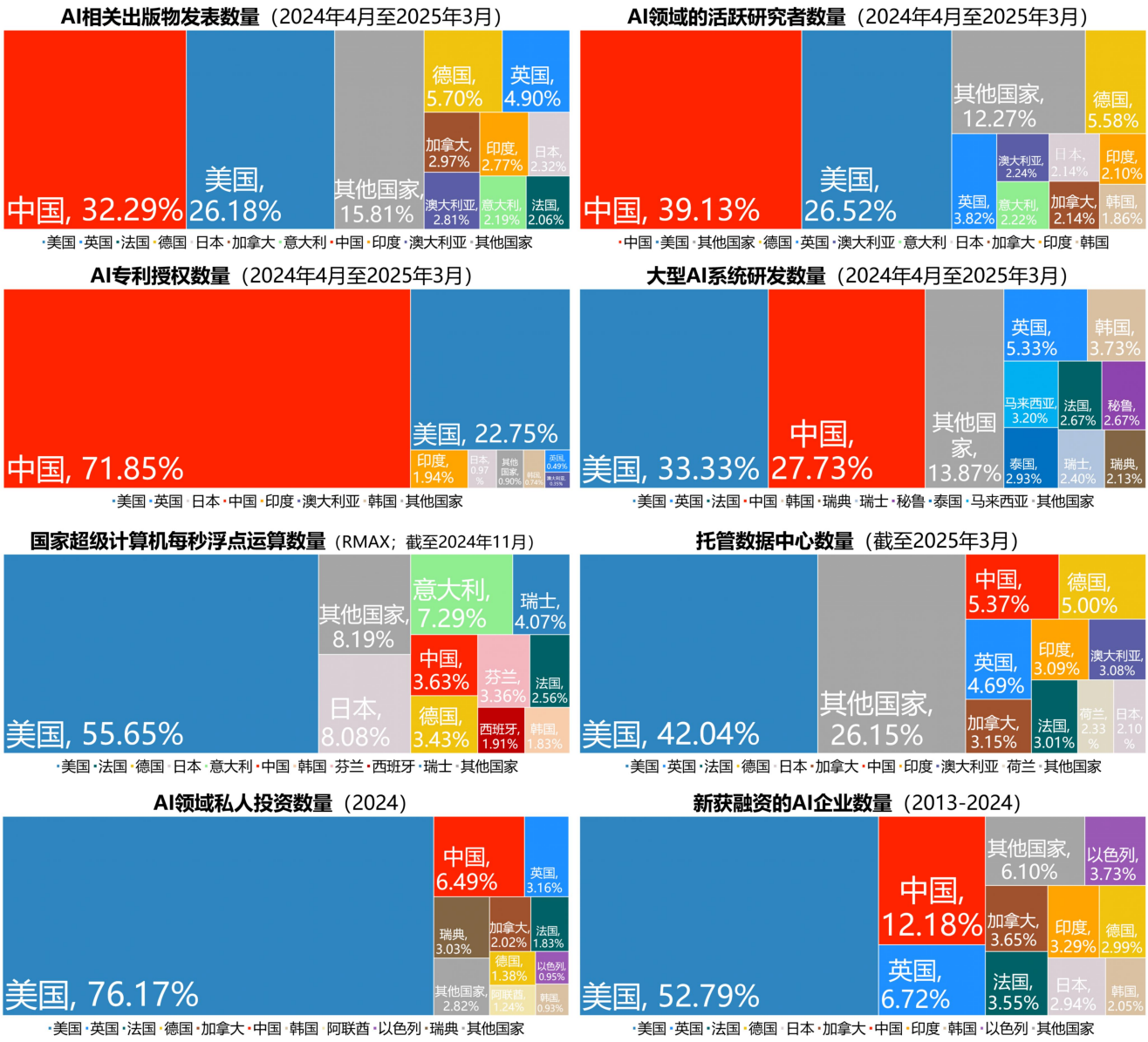


图 8 40 国在 AI 发展水平关键指标中的占比

数据来源：DBLP, Our World in Data, Data Center Map, The TOP500 List of Supercomputer, ICT Development Index 2024, Artificial Intelligence Index Report 2025

主要发现 1.2：

自 2010 年以来，生成式人工智能全球专利数量呈现出指数级增长。在此期间，生成式人工智能专利总量增长约 30 倍，生成式人工智能应用专利数量增长约 25 倍。2018 年之后增长速度显著加快。

从 2010 年到 2023 年，生成式人工智能（GenAI）专利总量和生成式人工智能应用专利数量都呈现出快速增长的趋势。生成式人工智能专利总量从 2010 年的 1,169 项增加到 2023 年的 36,389 项，增长了约 31.13 倍。生成式人工智能应用专利数量从 2010 年的 633 项增加到 2023 年的 16,046 项，增长了约 25.35 倍。2018 年之后，这两者的增长速度显著加快，呈现指数级增长态势，表明生成式人工智能领域的研发和专利相关活动进入了快速扩张阶段，创新活动不断加强。

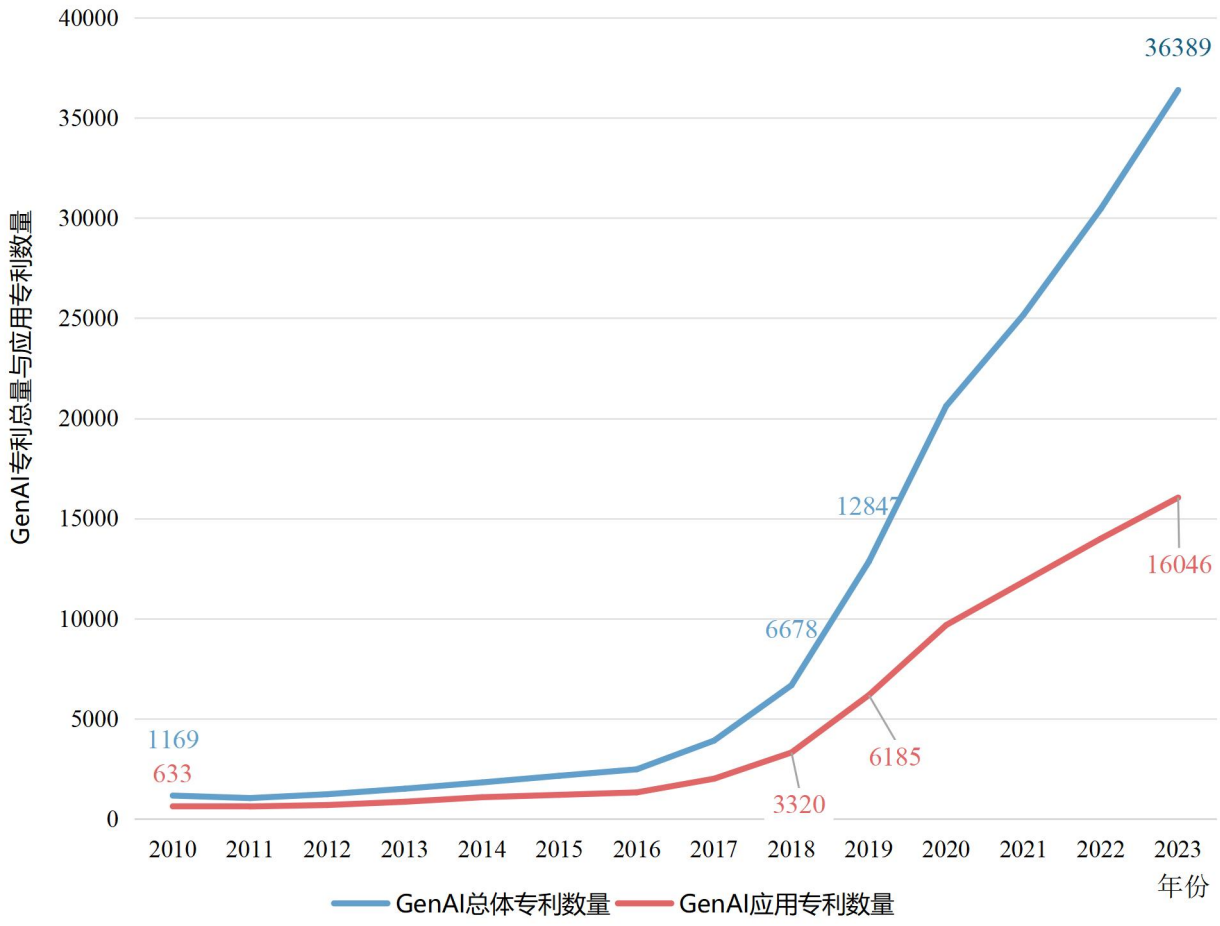


图 9 40 个国家的生成式人工智能专利总量与应用专利数量的年度发展趋势

数据来源：WIPO Patent Landscape Report on Generative AI 2024（数据集）

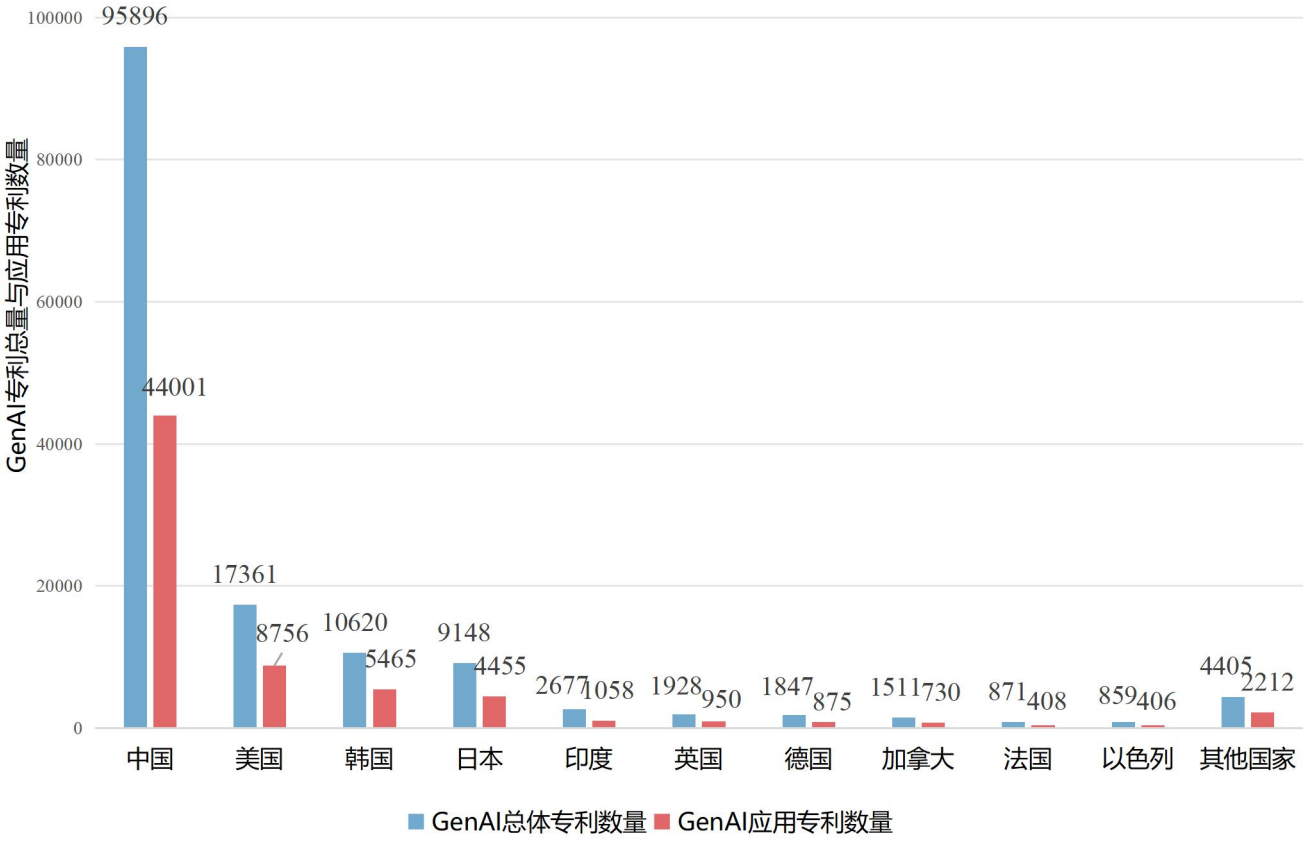


图 10 1997 年至 2023 年各国授权的生成式人工智能专利数量

数据来源：WIPO Patent Landscape Report on Generative AI 2024（数据集）

全球生成式人工智能授权专利高度集中，中国占生成式人工智能总体专利数量的 65%，超过了美国、韩国和日本的总和。在与应用相关的专利数量方面，中国的应用专利数量的比例达到了 63%，显示出在总量和实际应用方面的主导优势。其他国家存在显著差距，其中美国在生成式人工智能总体专利数量和应用专利数量方面排名第二，约为 12%，其他国家则低于 8%。

3.2. 评估方面 2：AI 治理环境

主要发现 2.1：

2024 年全球记录在案的 AI 风险事件数量较 2023 年翻了一倍。其中，美国的 AI 风险事件数量增长了 80%，而其他国家的增长速度更快，达到了 110%。

根据经合组织 AI 事件监测站（OECD AI Incidents Monitor, OECD AIM）的数据，2024 年

记录在案的 AI 风险事件数量相较于 2023 年保持了快速增长的态势。2024 年记录的 AI 风险事件数量增加了约 4000 起，与 2023 年相比增长率达到 104%，表明迫切需要一个有效的 AI 治理体系来跟进 AI 技术的快速发展。

尽管美国的 AI 风险事件数量也在迅速增加，但其增长率（2023 年至 2024 年增长 80%）相较于整体增长率（2023 年至 2024 年增长 104%）略低。2023 年，美国占全球 AI 风险事件总数的 31%，但这个比例在 2024 年下降到了 27%——其他国家所占的份额正在上升。

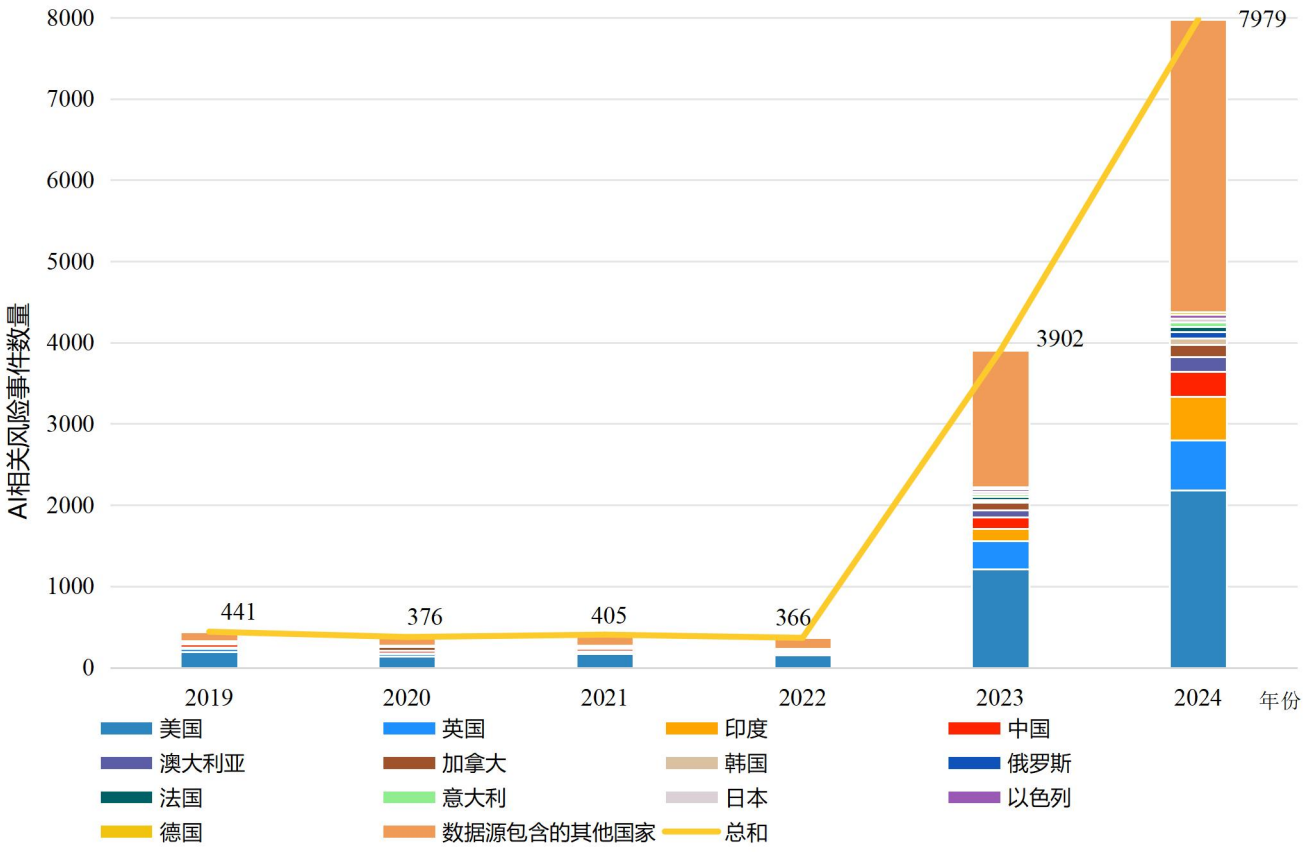


图 11 2017 年至 2024 年记录在案的 AI 风险事件数量

数据来源：OECD AIM

主要发现 2.2：

与 AI 内部安全与外部安全、AI 与人权以及数据治理相关的风险事件数量最多，占所有记录在案的 AI 风险事件数量的一半。

根据 OECD AIM 截至 2024 年 10 月记录的 AI 风险事件案例，最为频繁的风险涉及稳健性

与数字安全、尊重人权以及隐私与数据治理，表明技术与运行安全问题仍然是当前 AI 发展中最突出的关切。然而，与人类福祉、民主以及人类自主性相关的事件相对较少——这些领域通常与长期和系统性影响有关。这凸显了双重需求：一方面，应对短期安全风险仍然至关重要；另一方面，同样重要的是要预见并缓解那些不太明显但可能对社会价值产生深远影响的风险。因此，平衡的治理方法必须同时整合短期风险管控和长期伦理预见，为 AI 发展奠定可持续的基础。

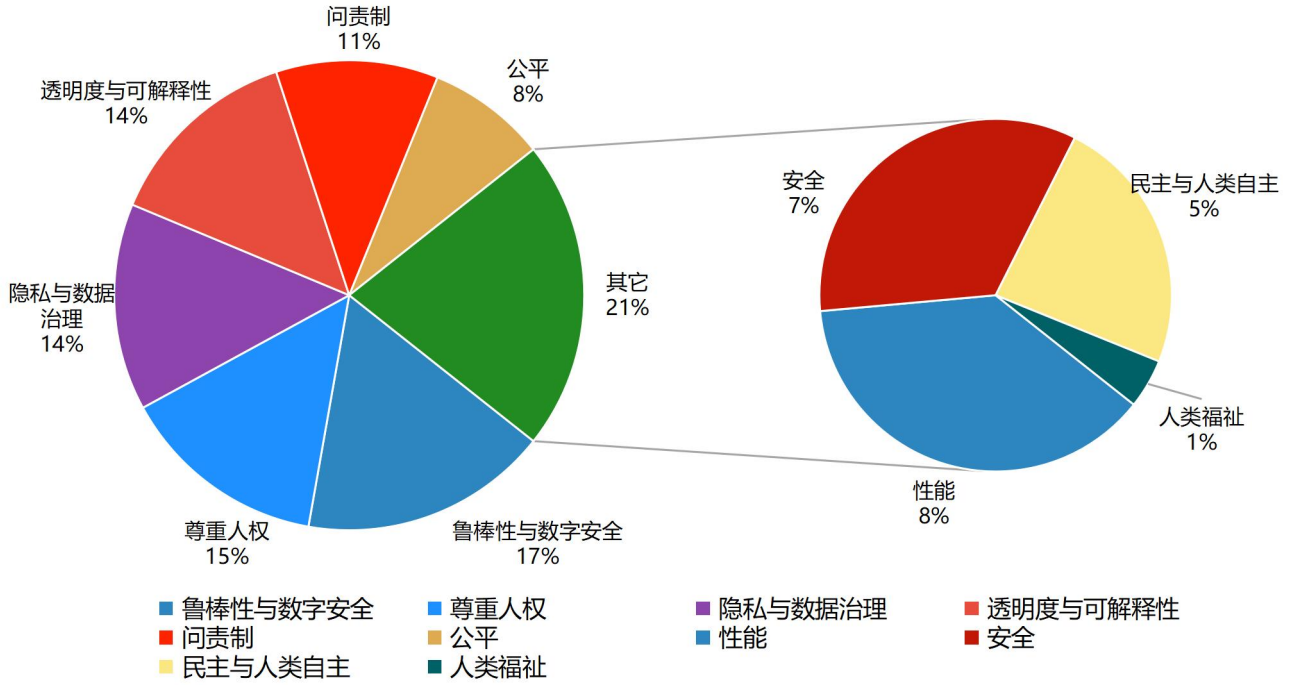


图 12 按主题分类的 AI 风险事件分布

数据来源: OECD AIM

主要发现 2.3：

尽管高收入国家往往在整体治理方面得分更高——这表明它们在 AI 治理方面具有更强的内在准备度和基础，但实际经验仍显示，在不同国家背景下积极推进 AI 治理仍有很大的努力空间。

在 AI 治理准备度的三维评估体系（包括对国家治理整体水平评估表现、国家数字治理整体水平评估表现以及国家可持续发展整体水平评估表现的综合评估）中，高收入国家的表现大多优于中高和中低收入国家组的表现。例如，丹麦和芬兰等高收入国家凭借其成熟的治理体系、先进的数字基础设施以及在可持续发展方面取得的显著进展，在评估体系中占据优势，成为治理准备度的“第一梯队”。然而，应该看到高收入国家在治理准备度上的优势是长期发展的积

累结果，并不构成 AI 治理的“门槛”——技术迭代的平等性和治理需求的多样性使得每个国家都有机会根据自身需求来推进 AI 治理。例如，中国尽管仅作为一个中高收入国家，在整体治理准备度方面相对高收入国家组处于较低的水平，但在 AI 治理工具和 AI 治理成效评估方面上仍取得了较高的表现。

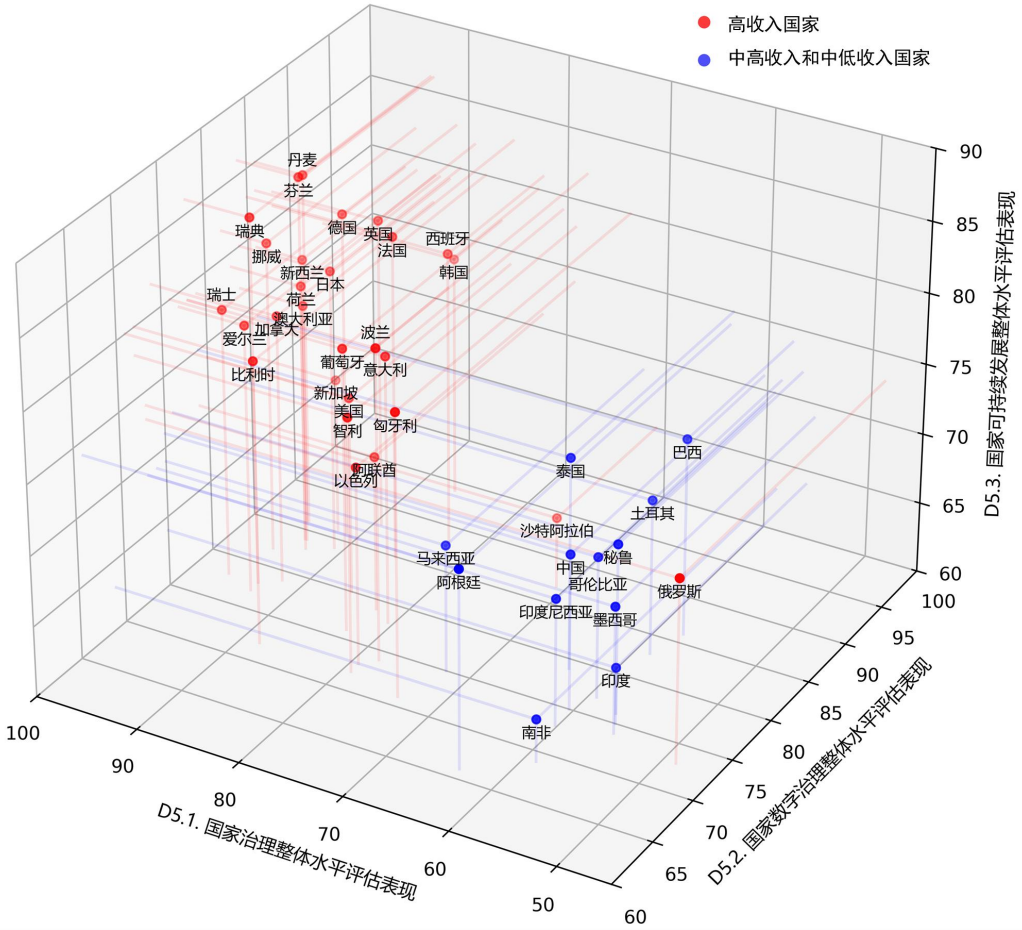


图 13 40 个国家整体治理准备度评分的构成（分数来自相关指数）

数据来源: Worldwide Governance Indicators, Human Development Index, Global Cybersecurity Index, Effective Governance of Data, GovTech Maturity Index, E-Government Development Index, E-Participation Index, 2025 Sustainable Development Index

3.3. 评估方面 3：AI 治理工具

评估方面 3 概览：

AI 治理依赖于多样化的工具与路径，它们各自承担不同但互补的职能。实现这些治理工

具的有效协同与部署，是构建和健全 AI 监管体系与推动可持续 AI 创新的关键基础。战略文件与伦理原则为治理提供高层指引，技术评估与标准则构成科学监管的依据；而法律法规及制度框架将治理目标具体化，界定可执行的规则与职责。在此基础上，全球合作机制促进多方对话与包容，汇聚多元视角，推动 AI 负责任的发展来造福全人类。

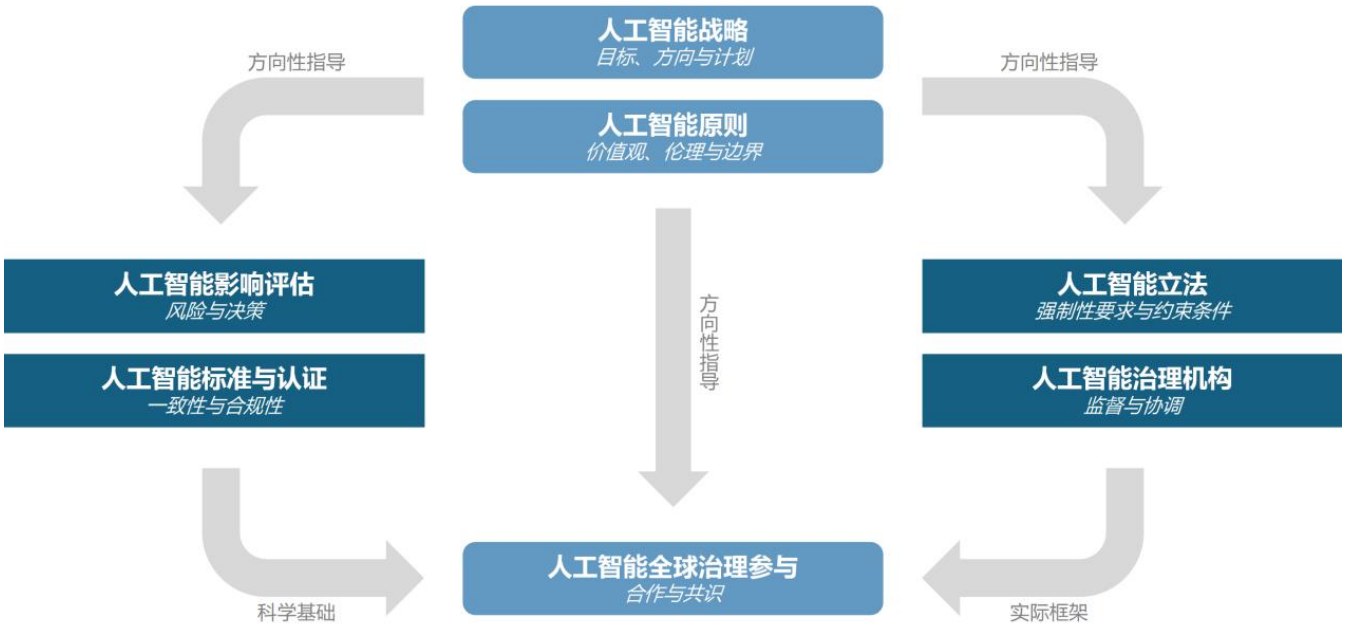


图 14 AI 治理工具：整体视角

主要发现 3.1：

所有 40 个被评估的国家都已发布了国家层面的 AI 战略，但在战略制定的结构化方法上存在差异。

在 AI 战略这一维度上，40 个被评估的国家整体表现良好。所有被评估的国家目前均已制定国家层面的 AI 战略，其中包括在上一轮评估中尚未发布战略的南非。从战略时间框架来看，瑞士每年更新其数字战略，以指导 AI 发展；法国、意大利、马来西亚和土耳其等国制定了为期 2 至 5 年的短期战略。相对而言，印度尼西亚的 AI 战略覆盖长达 25 年的时间跨度。从战略结构上看，阿根廷、爱尔兰和美国等国采用了模块化结构，围绕不同发展方面分别制定目标；中国、意大利和秘鲁等国则采用垂直型结构，系统呈现当前现状、战略目标与实施路径；法国、印度和南非等国家则更倾向于采用全面性和混合型战略布局。

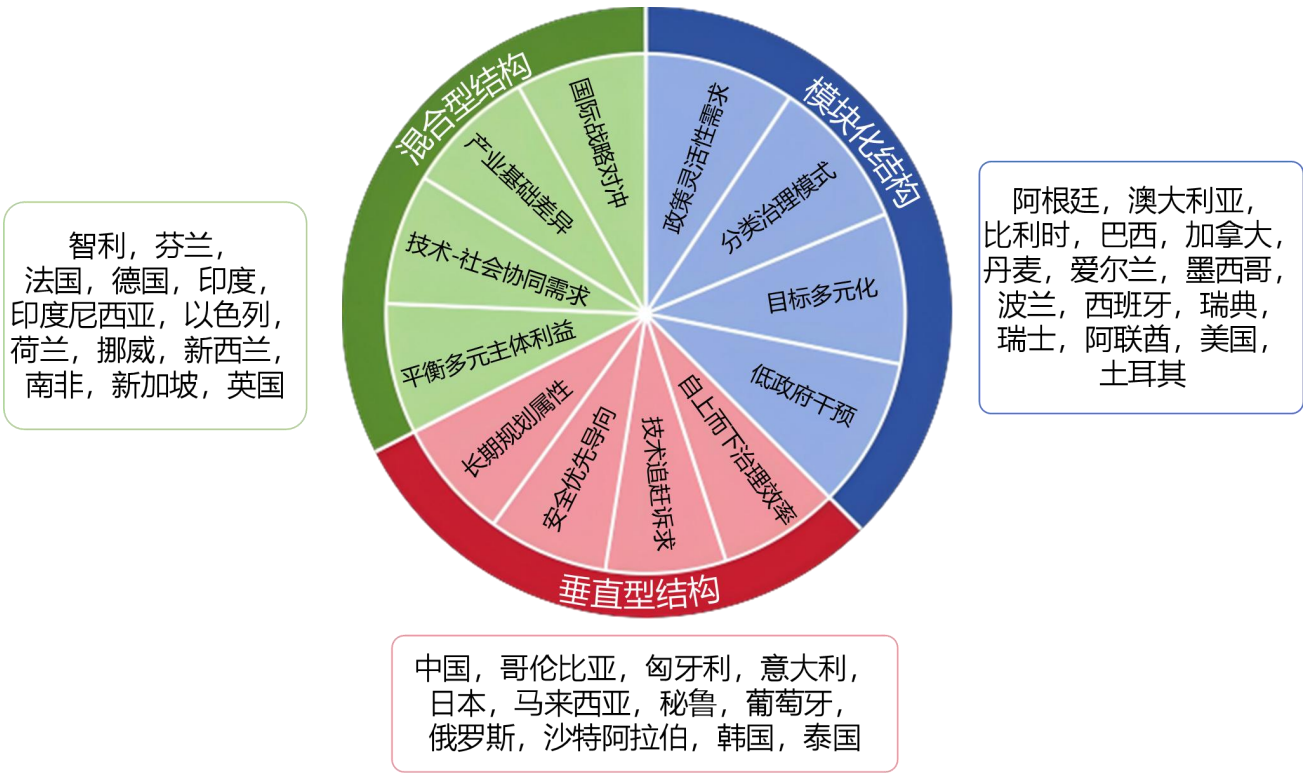


图 15 40 个国家 AI 战略的不同结构类型

主要发现 3.2：

自 2024 年以来，AI 立法呈现出明显的加速趋势。一些国家已经制定了国家层面专门针对 AI 的综合性法律法规，还有一些国家则制定了针对 AI 垂直领域的法律法规。

在被评估的 40 个国家中，AI 立法呈现出综合性法律与垂直领域法律并存的局面。其中，有 26 个国家已经颁布或正在制定国家层面专门针对 AI 的综合性法律法规。在这些国家中，欧盟的 14 个国家以及挪威都在执行欧盟《人工智能法》（*Artificial Intelligence Act*），这是世界上第一部综合性的 AI 法律，于 2024 年 8 月 1 日生效。韩国于 2024 年 12 月 26 日通过了《人工智能基本法》（*The Basic Act on the Development of AI and the Establishment of Trust*），计划于 2026 年 1 月开始实施。此外，包括中国、法国、日本、秘鲁和土耳其在内的 10 多个国家目前正在起草各自专门针对 AI 的国家层面的法律法规。

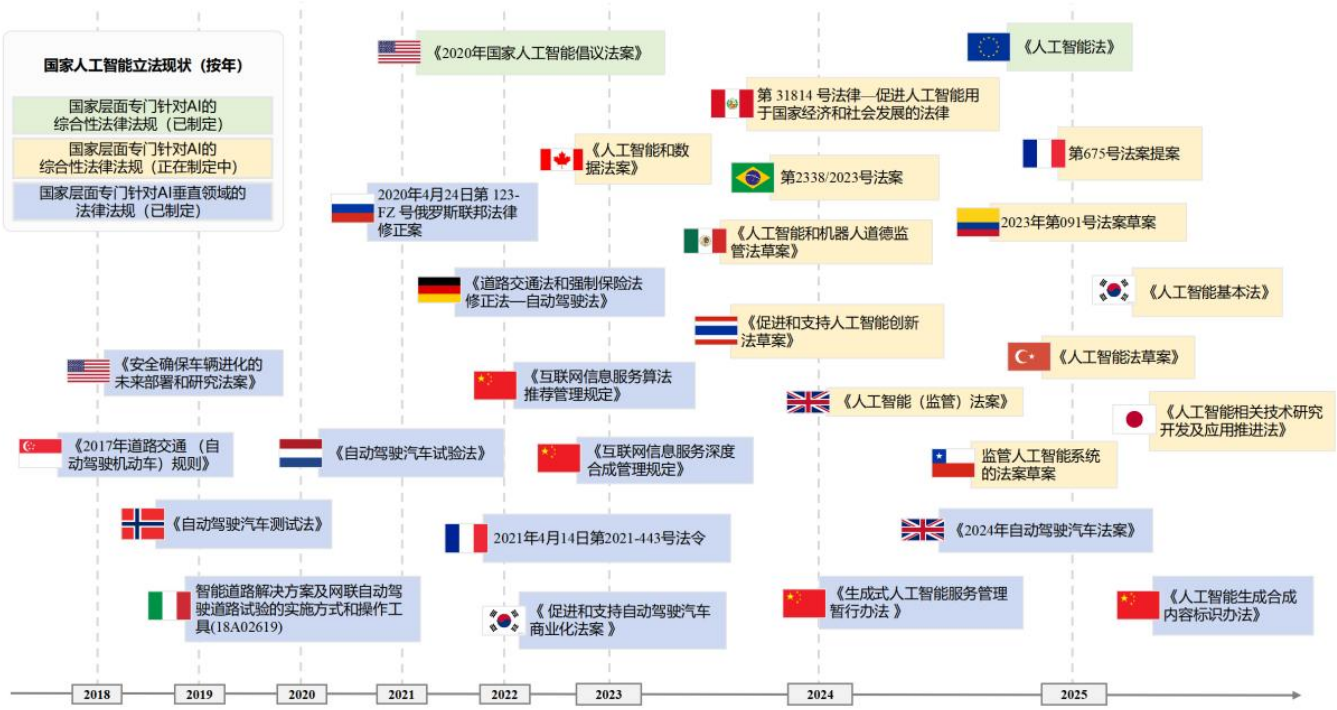


图 16 各国 AI 立法状况（按年份）

在推进综合性立法的同时，各国还通过补充条款或修正案积极将 AI 治理融入现有法律框架。目前，在 40 个国家中，有 24 个国家已经制定了与 AI 直接相关的数据或信息保护法；另有 11 个国家颁布了聚焦垂直领域的国家级 AI 法律，主要集中在两个领域：自动驾驶和生成式 AI。

从立法趋势来看，美国、意大利和挪威在 2017 - 2018 年就在自动驾驶领域启动了相关立法。自 2023 年以来，中国开始在生成式 AI 领域推进垂直立法，这一领域也正逐渐成为治理新焦点。这反映出 AI 立法作为治理工具正朝着专业化和细分化发展的趋势。

主要发现 3. 3：

40 个国家都以多种形式参与了全球 AI 治理的相关机制，其中法国、日本、韩国和新加坡的参与度最高。

所有 40 个被评估的国家都以不同程度参与了各种全球 AI 治理机制，突显了国际合作在 AI 治理中的重要性。其中，法国、日本、韩国和新加坡最为活跃，参与了此次 AGILE 指数评估调研的所有 AI 国际治理机制/倡议。在 40 个国家中，除美国和以色列外，其他国家都签署了联合国教科文组织的《人工智能伦理问题建议书》（*Recommendation on the Ethics of Artificial*

Intelligence）。在联合国未来峰会通过的《全球数字契约》（*Global Digital Compact*）是迄今为止得到最广泛支持的 AI 治理文件，在此次评估国家中仅有俄罗斯出于主权和安全因素未参与签署。在联合国机制之外，关键的全球治理框架包括在英国举行的 AI 安全峰会、在韩国举行的 AI 首尔峰会以及在法国举行的 AI 行动峰会、由荷兰和韩国牵头的军事领域负责任使用 AI 峰会以及由美国主导的国际 AI 安全研究所网络。全球 AI 治理格局呈现出多边合作并行和多元参与的鲜明特点。尽管距离实现充分包容的参与模式与协同应对机制仍有差距，但诸如 AI 安全、能力建设等核心议题已获得广泛响应。

表 3 各国在部分 AI 国际治理机制/倡议中的参与情况

	人工智能国际治理机制/倡议							
国家	联合国教科文组织：《人工智能伦理问题建议书》(2021)	人工智能安全峰会：《布莱切利宣言》(2023)	2023军事领域负责任使用人工智能峰会：《行动呼吁》(2023)	人工智能首尔峰会：《首尔部长声明》(2024)	联合国：《全球数字契约》(2024)	2024军事领域负责任使用人工智能峰会：《行动蓝图》(2024)	人工智能安全研究所国际网络：初始成员(2024)	人工智能行动峰会：《关于发展包容、可持续的人工智能造福人类与地球的声明》(2025)
法国	√	√	√	√*	√	√	√	√
日本	√	√	√	√*	√	√	√	√
新加坡	√	√	√	√*	√	√	√	√
韩国	√	√	√	√*	√	√	√	√
加拿大	√	√	√	√*	√	×	√	√
德国	√	√	√	√*	√	√	×	√
荷兰	√	√	√	√	√	√	×	√
英国	√	√	√	√*	√	√	√	×
澳大利亚	√	√	×	√*	√	×	√	√
意大利	√	√	√	√*	√	×	×	√
美国	×	√	√	√*	√	√	√	×
智利	√	√	×	√	√	×	×	√
中国	√	√	√	×	√	×	×	√
印度	√	√	×	√	√	×	×	√
印度尼西亚	√	√	×	√	√	×	×	√
西班牙	√	√	×	√	√	×	×	√
瑞士	√	√	×	√	√	×	×	√
巴西	√	√	×	×	√	×	×	√
爱尔兰	√	√	×	×	√	×	×	√
墨西哥	√	×	×	√	√	×	×	√
新西兰	√	×	×	√	√	×	×	√
沙特阿拉伯	√	√	×	√	√	×	×	×
土耳其	√	√	×	√	√	×	×	×
阿联酋	√	√	×	√	√	×	×	×
比利时	√	×	×	×	√	×	×	√
哥伦比亚	√	×	×	×	√	×	×	√
丹麦	√	×	×	×	√	×	×	√
芬兰	√	×	×	×	√	×	×	√
匈牙利	√	×	×	×	√	×	×	√
以色列	×	√	×	√	√	×	×	×
挪威	√	×	×	×	√	×	×	√
秘鲁	√	×	×	×	√	×	×	√
波兰	√	×	×	×	√	×	×	√
葡萄牙	√	×	×	×	√	×	×	√
南非	√	×	×	×	√	×	×	√
瑞典	√	×	×	×	√	×	×	√
泰国	√	×	×	×	√	×	×	√
阿根廷	√	×	×	×	√	×	×	×
马来西亚	√	×	×	×	√	×	×	×
俄罗斯	√	×	×	×	×	×	×	×

注：标记*的国家同时参与签署了人工智能首尔峰会《首尔宣言》（2024）。

3.4. 评估方面 4：AI 治理成效

主要发现 4.1：

公众对 AI 应用与服务的认知度、信任度与积极态度总体上与各国人均 GDP 呈负相关。

在人均 GDP 与公众对产品和服务中 AI 应用的认知程度之间，呈现出一定的负相关趋势：整体上收入水平越高的国家，公众对 AI 技术的特性及其社会影响的认知往往较低。部分发展中国家，如中国、土耳其、印度尼西亚、墨西哥和马来西亚，在公众认知方面表现出高于同类国家的水平。在高收入国家组中，亚洲国家如韩国和新加坡的认知水平明显高于发达国家的整体趋势。

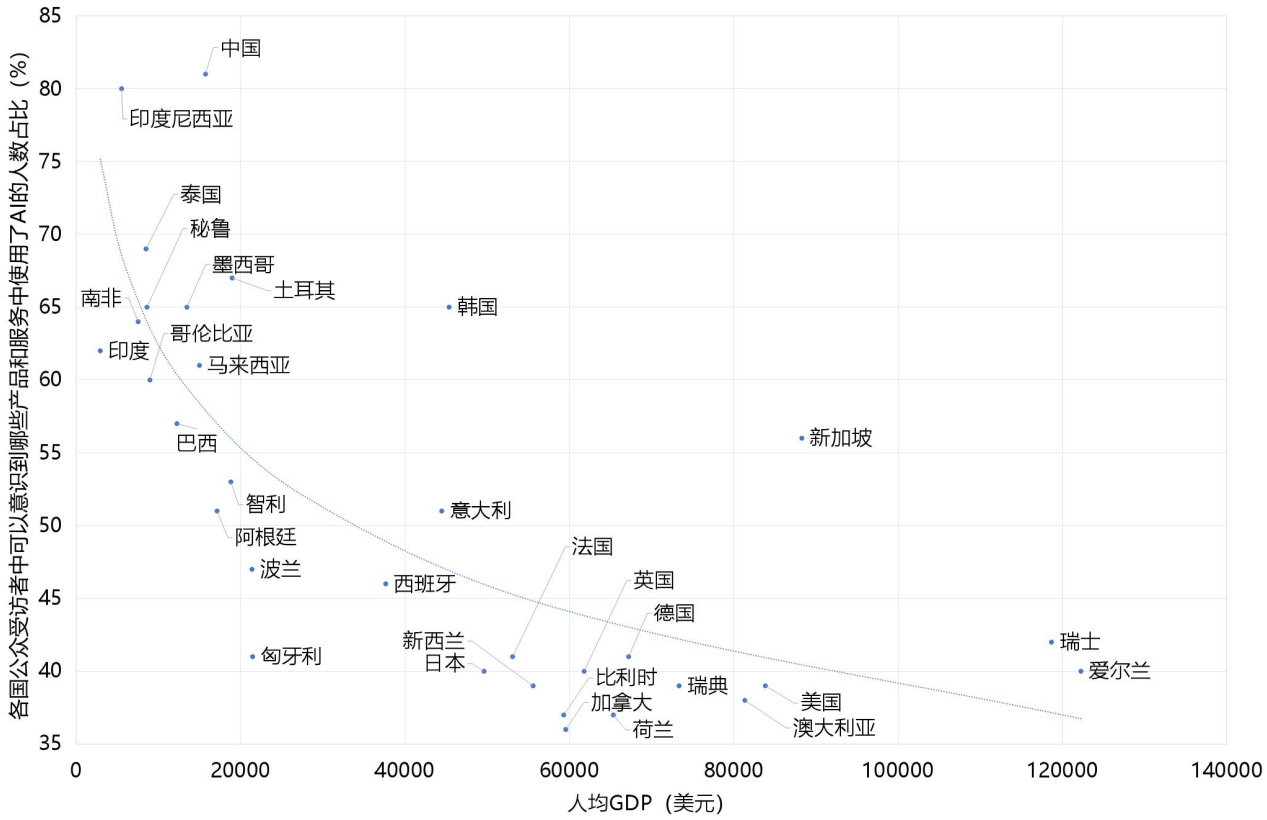


图 17 人均 GDP 与公众对产品和服务中 AI 应用的认知水平

数据来源: THE IPSOS AI MONITOR 2024

在 AI 应用的公众信任度以及对 AI 发展的积极态度方面，中国、印度尼西亚、泰国和墨西哥等新兴经济体表现较为突出。相比之下，瑞典、美国、法国、比利时和加拿大等国家的公众在上述方面则持更为审慎的态度。

在对 AI 使用态度方面，各国公众普遍认可 AI 在推动创新和提升效率方面的潜力：尤其是在个人娱乐选择、完成任务时长、个人健康情况、个人工作情况等领域。然而，公众对于 AI 涉及的伦理与现实风险仍持谨慎态度，特别是在公众对 AI 的信任程度、就业市场以及网络虚假信息数量等方面。

表 4 公众对 AI 应用的态度和预期

数据来源：THE IPSOSAI MONITOR 2024

(注：对 32 个国家 75 岁以下成年受访者进行在线调查，表中数据为选择各选项的受访者占比)

评估维度	D14.3. 公众对AI应用的信任程度		D14.1. 公众对AI发展的积极态度		D14.2. 公众对AI使用的积极或中性预期							
国家/问/答	各国受访公众中同意以下观点者的占比（%）					在未来3–5年，受访公众对以下领域人工智能应用持积极或中立预期者的占比（%）。						
	我相信使用AI的公司会保护我的个人数据。	我相信AI不会歧视或偏袒任何群体。	使用AI产品和服务让我感到兴奋。	使用AI产品和服务的利大于弊。	未来3-5年，使用AI产品和服务将深刻改变我的日常生活。	个人娱乐选择	完成任务时长	个人健康情况	个人工作情况	经济情况	网络虚假信息数量	就业市场情况
中国	66	76	80	83	86	92	92	89	89	87	70	63
印度尼西亚	66	68	76	80	80	89	93	90	87	84	73	69
泰国	68	73	76	77	78	90	89	88	85	82	73	72
墨西哥	64	76	70	70	76	87	90	84	83	81	69	68
秘鲁	62	74	67	70	76	88	90	86	81	78	73	69
哥伦比亚	50	66	63	68	77	88	88	80	76	72	65	62
南非	62	72	61	62	76	80	86	79	76	69	63	58
马来西亚	52	62	68	63	71	82	82	83	79	72	61	68
新加坡	56	60	64	66	79	85	85	78	70	77	54	59
土耳其	46	61	70	69	76	78	81	73	77	64	58	63
韩国	37	51	73	66	79	83	87	80	74	70	61	49
阿根廷	46	60	52	57	67	85	84	79	78	70	66	65
印度	60	63	63	62	65	68	70	69	66	68	64	64
智利	43	59	46	60	69	82	86	79	72	58	65	58
巴西	45	56	56	56	62	78	81	76	76	67	57	58
意大利	58	61	49	53	60	76	78	76	74	62	51	53
匈牙利	64	64	47	51	64	71	75	75	67	64	41	45
西班牙	48	52	45	50	59	79	81	76	73	61	51	49
德国	43	48	46	47	59	77	79	73	73	62	46	53
荷兰	44	38	40	36	63	76	81	81	81	65	44	56
瑞士	43	43	42	42	55	78	78	78	80	64	46	52
新西兰	39	43	43	48	64	76	80	77	78	61	37	42
波兰	45	53	44	44	56	74	76	71	70	67	43	42
日本	27	38	47	48	63	75	77	76	70	63	47	49
爱尔兰	42	42	40	45	59	76	74	72	76	63	44	47
英国	41	44	38	46	58	73	79	75	74	59	41	45
澳大利亚	32	42	39	44	61	74	68	73	71	55	39	42
瑞典	35	33	36	43	52	69	77	72	79	59	34	45
美国	33	41	34	39	58	67	75	73	73	51	39	44
法国	35	41	39	41	57	69	75	67	68	51	38	45
比利时	40	37	33	38	61	69	74	70	73	53	39	39
加拿大	28	41	37	40	61	73	75	68	73	50	37	40

主要发现 4.2：

越来越多的女性参与到了 AI 相关的研究中，活跃研究者中男女比例持续下降，到 2025 年已降至 2.2:1。

根据对 DBLP 数据的统计分析总体趋势显示，AI 相关领域活跃研究者中的女性占比近年来在稳步增长，相关出版物中作者的男女比例已由 2017 年的 4.0:1 降至 2025 年的 2.2:1，反映出女性在 AI 科研领域的参与度持续提升。

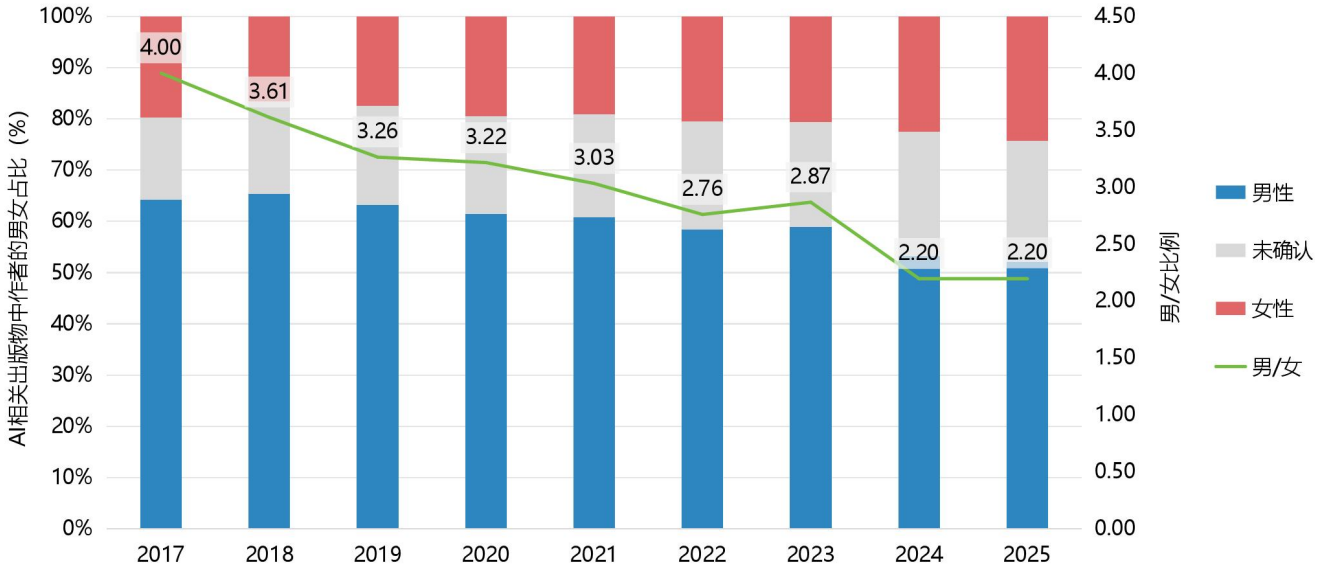


图 18 AI 相关出版物作者的男女比例 2017-2025

数据来源：基于对 DBLP 文献数据库的统计分析（注：性别信息根据各国人名习惯，通过 DBLP 中出版物作者的姓名推断得出。结果可能受到中性命名或跨文化差异所带来的影响）

在本次评估的 40 个国家中，性别比例最为平衡的是泰国，其次为中国、印度尼西亚和新加坡；北美和西欧国家整体上并无明显优势。部分发达国家，如德国、日本等，仍存在 AI 研究人员中性别比例严重失衡的情况。

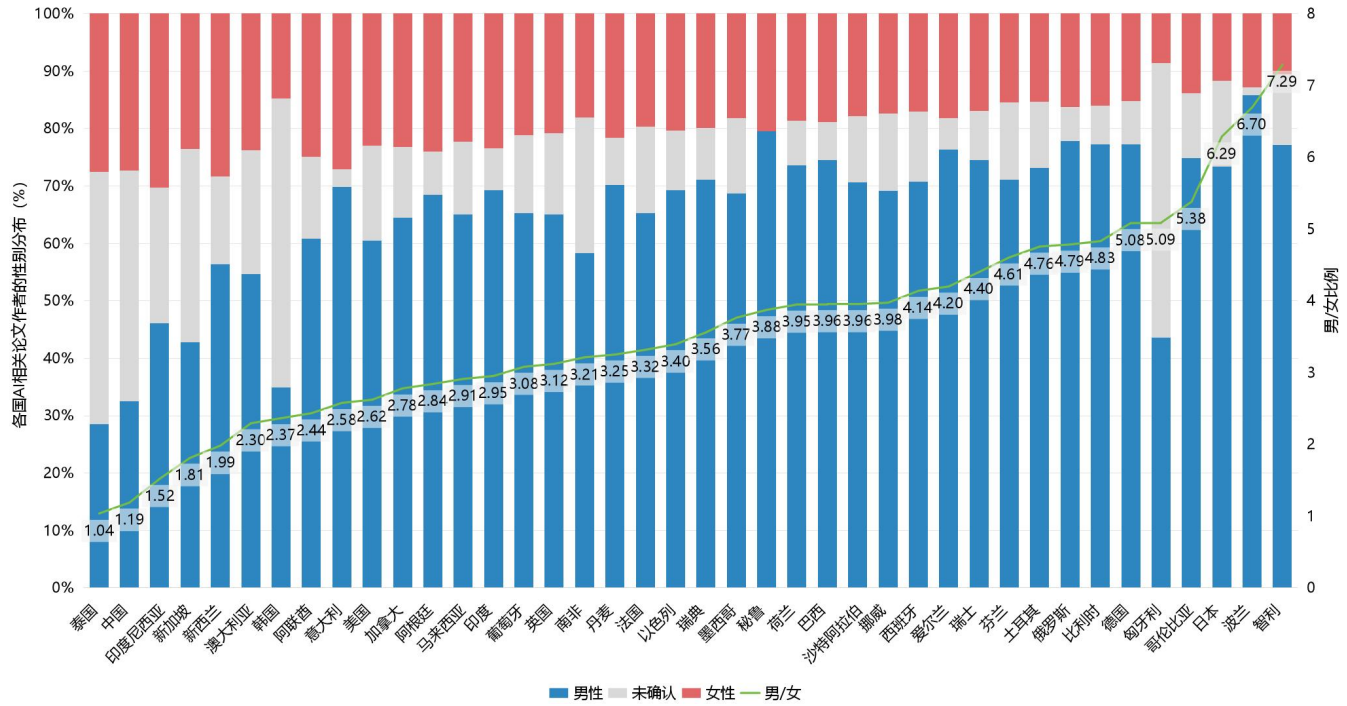


图 19 40 国 AI 相关出版物作者的性别占比情况

数据来源：基于对 DBLP 文献数据库的统计分析（2024 年 4 月至 2025 年 3 月期间）

主要发现 4.3：

以人均 GDP 为代表的经济发展水平在一定程度上与社会弱势群体的数字普惠程度呈正相关，但较高的经济发展水平并不必然带来高水平的数字包容性。

在不同人均 GDP 水平的国家中，弱势群体（老年人和低收入人群）的互联网接入与使用情况呈现出一定趋势。总体来看，人均 GDP 水平更高的国家，这两类人群的互联网使用比例相对更高，体现了经济发展水平与弱势群体数字普惠间的正向关联。然而同样可以看到，美国、澳大利亚和以色列等部分发达国家在相关表现上与其经济表现所带来的预期差距明显。因此，经济实力和技术优势并不必然惠及弱势群体，实现真正包容的技术发展仍需长期的社会政策与治理努力支撑。

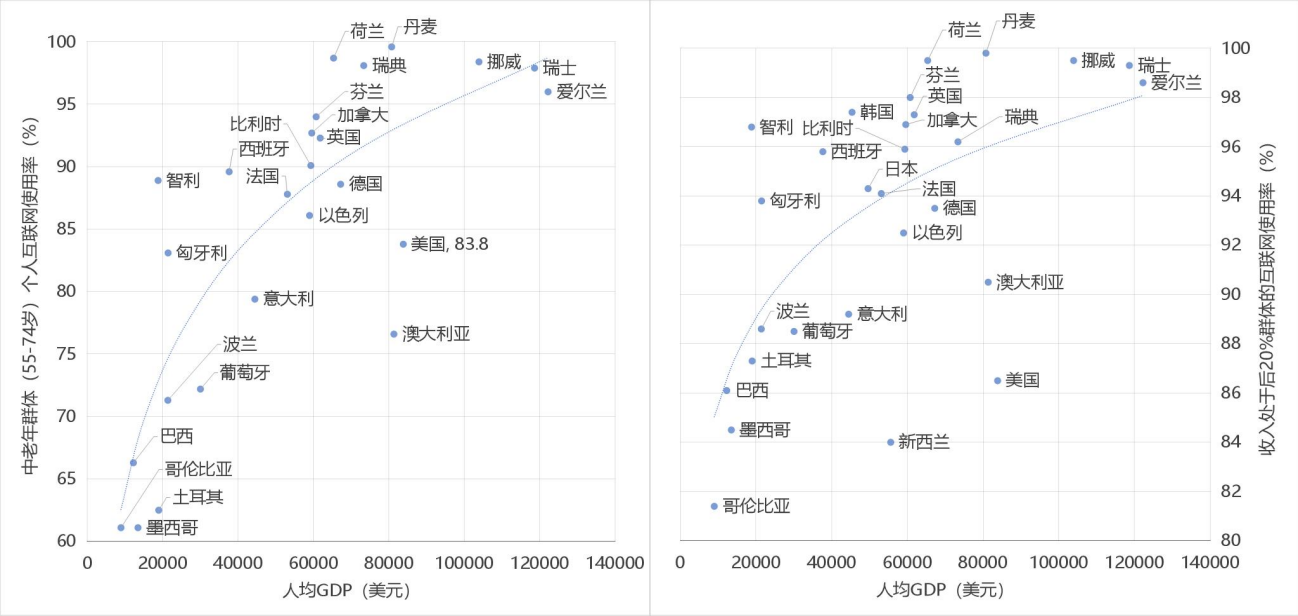


图 20 人均 GDP 与弱势群体的互联网使用与接入情况

数据来源: OECD Going Digital Toolkit

主要发现 4.4：

在有影响力的开放 AI 模型与数据集方面，中国和美国合计占 40 个国家总量的 70%以上。

根据对 Hugging Face 开源社区的统计分析显示，中国共发布 2,014 个具有影响力的 AI 模型，占比达 50.3%，位居全球第一；美国以 995 个模型（24.8%）排名第二。在 AI 数据集方面，美国以 922 个具有影响力的数据集位居首位，中国紧随其后，数量为 912 个。中国与美国在开源 AI 生态体系中不仅具备数量优势，也在全球 AI 科研与应用发展中发挥着重要引领作用。

北美和欧洲整体上依然保持优势地位，加拿大、英国、法国、德国和瑞士均进入模型与数据集数量的前十名，但发展中经济体的表现也日益突出。除中国外，印度在 AI 数据开放方面表现较好，与中国一同成为进入前十的发展中国家。巴西、智利和印度尼西亚等国也通过其在 AI 科研与应用方面的探索，为全球南方在该领域赢得更多关注。

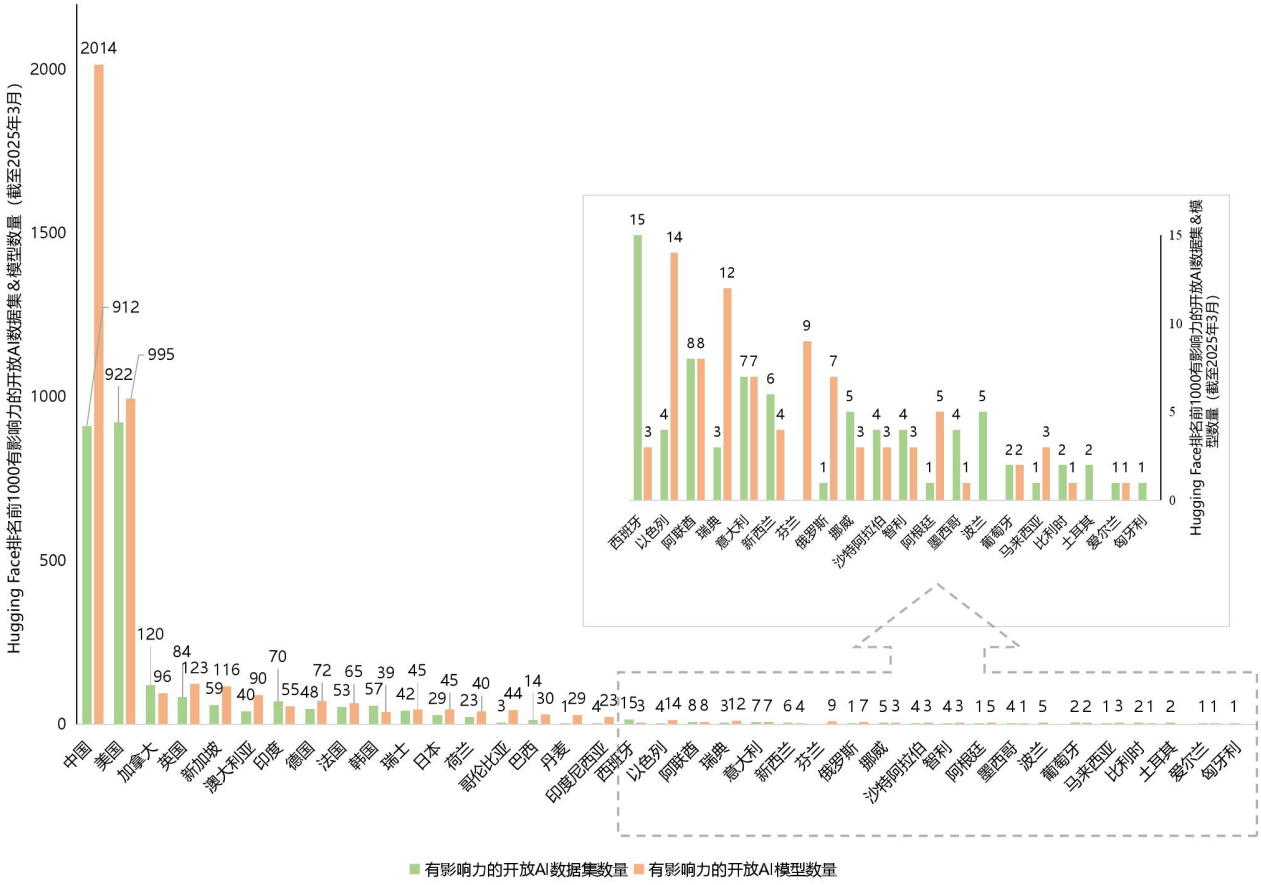


图 21 具有影响力的开源 AI 数据集与模型数量

数据来源：基于对 Hugging Face 数据分析（截至 2025 年 3 月）

主要发现 4.5：

北美和西欧在开源 AI 生态上占据主导，美国以占 40 国 30.43%的开源贡献量遥遥领先。与此同时，中国和印度也展现出强劲的参与度。

从 GitHub 平台上热门 AI 开源项目的贡献看，美国以 202,118 名贡献者、占 40 个国家开源贡献总量的 30.43%而处于首位，显示出其在技术投入、开发生态与教育体系等方面的综合优势。德国（82,912 人，12.48%）、荷兰（73,558 人，11.07%）和挪威（58,525 人，8.81%）等西欧发达国家亦排名靠前。与此同时，中国（62,789 人，9.45%）与印度（14,815 人，2.23%）分别位列第四与第八，反映出两国在编程人才规模、教育普及程度与技术社区活跃度方面也奠定了坚实的基础。

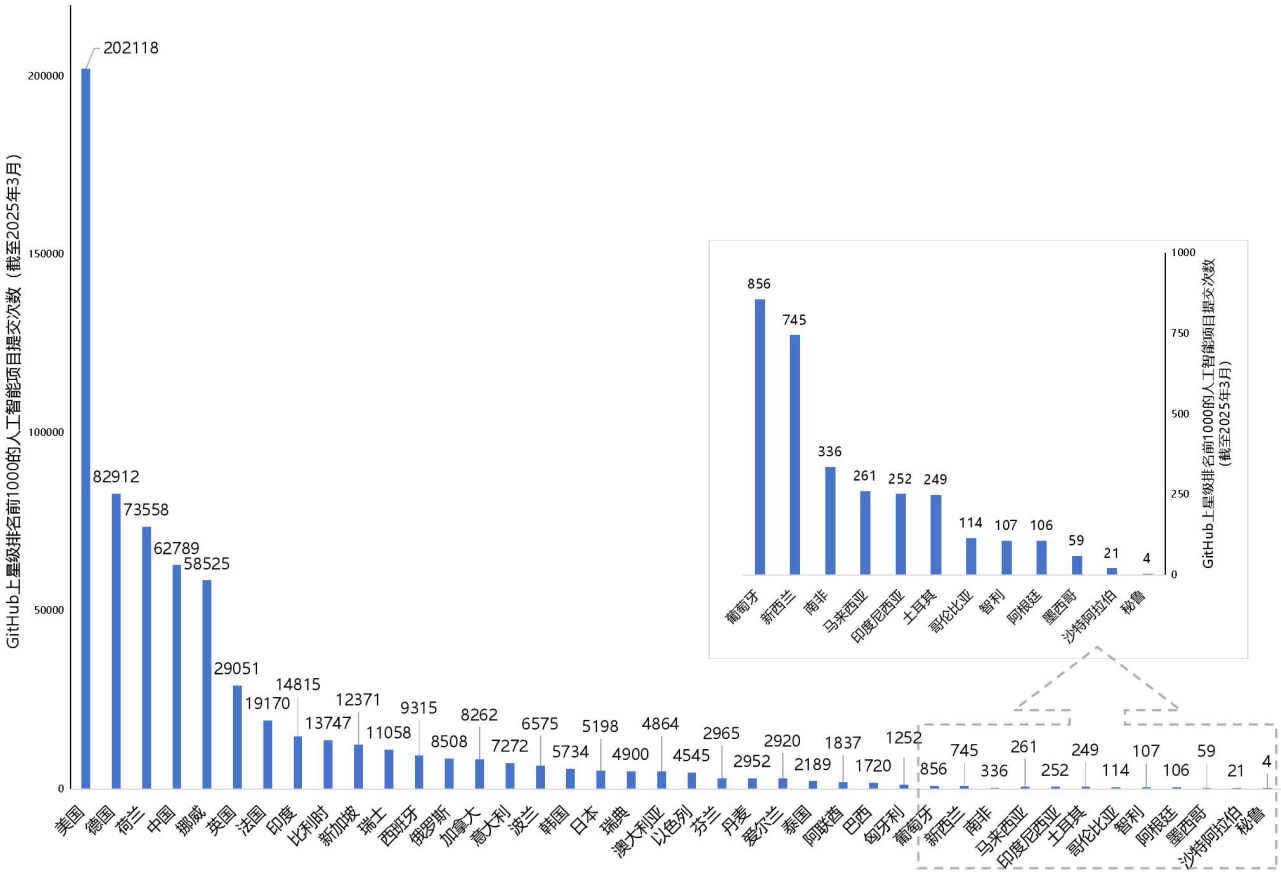


图 22 各国在 GitHub 热门 AI 开源项目上的贡献总量（按国家统计提交次数）

数据来源：基于对 GitHub 数据分析（截至 2025 年 3 月）

主要发现 4.6：

对 AI 治理相关议题的研究关注持续上升，相关出版物在全部 AI 出版物中的占比在 2024 年已达到 14.4%。在所评估的 40 个国家中，来自中国与美国的 AI 治理相关研究占比合计达到 54%，贡献超过一半。

尽管近年来人工智能领域论文总量有所波动，但聚焦治理方向的研究占比持续攀升——从 2020 年的 10.4%增至 2025 年的 14.4%。治理正逐渐成为人工智能研究版图中不可或缺的组成部分。在这 40 个国家中，人工智能治理论文数量主要由美国和中国贡献，占比分别为 28%和 26%左右。德国、英国和澳大利亚也作出了重要贡献。与去年相比，德国在占比排名上从第四位升至第三位，超越了英国。

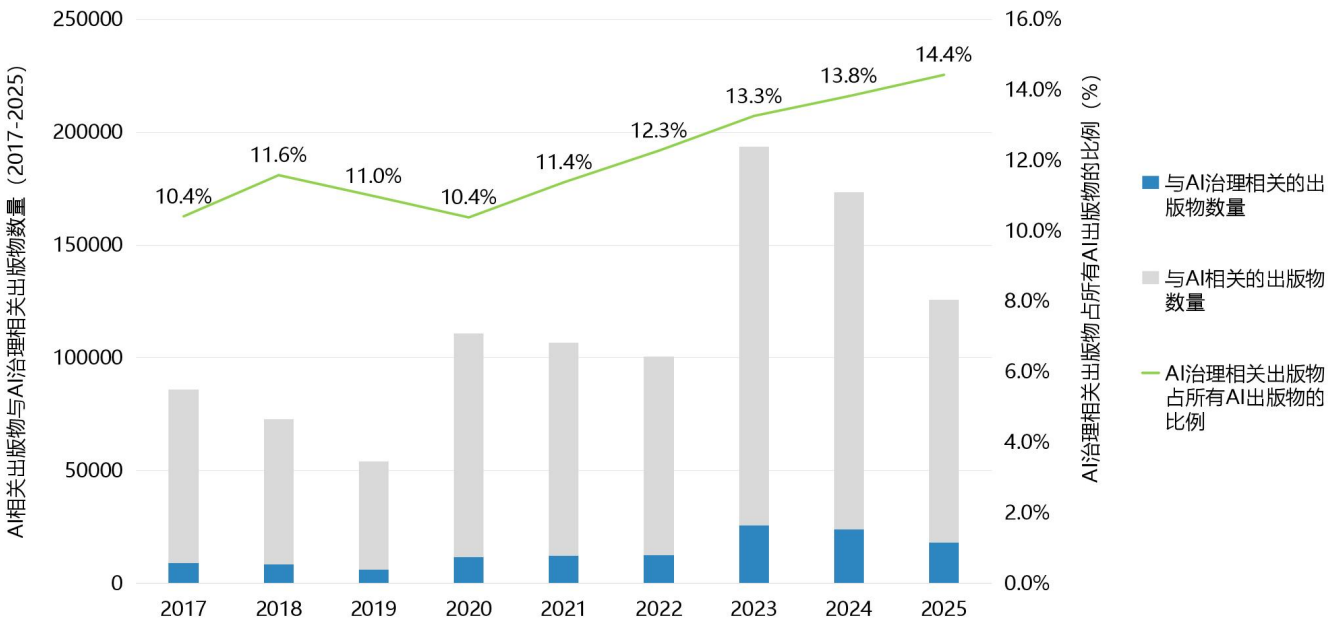


图 23 AI 治理相关出版物数量及其在所有 AI 相关出版物总量中的占比（2017－2025）

数据来源：基于对 DBLP 数据库的统计分析（截至 2025 年 3 月）

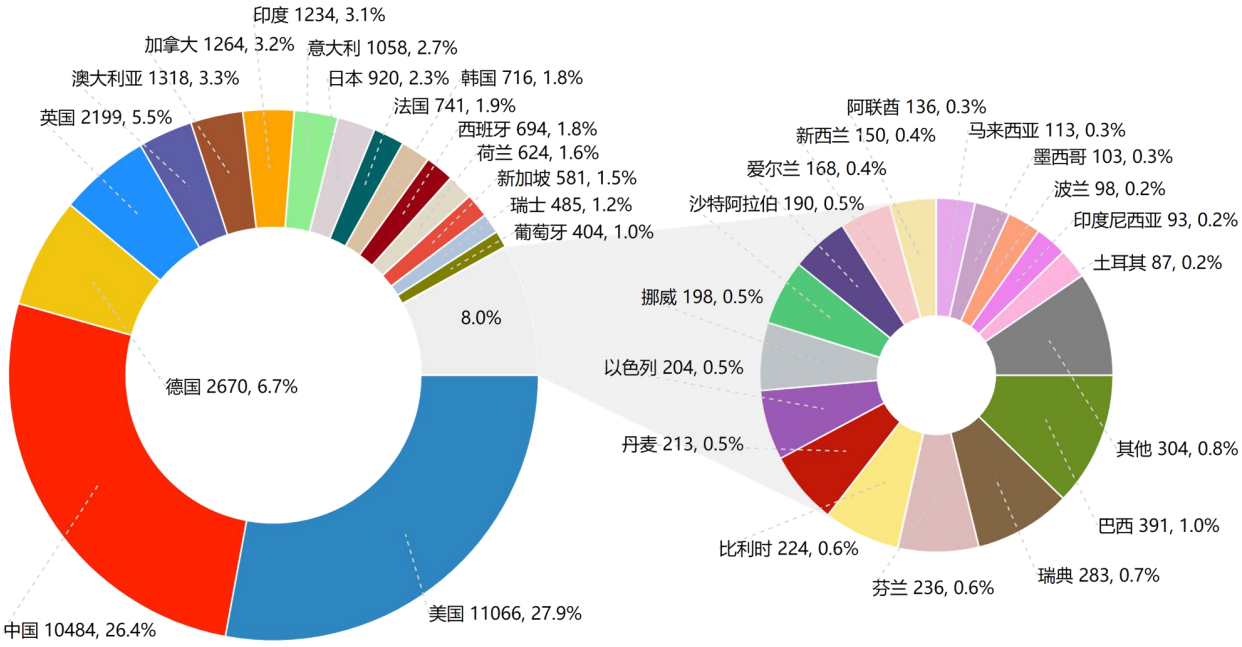


图 24 各国在 AI 治理相关出版物中的占比

数据来源：基于对 DBLP 文献数据库的统计分析（2024 年 4 月至 2025 年 3 月期间）

主要发现 4.7：

在 AI 治理研究的国际合著中，中国、美国、加拿大、德国、英国和澳大利亚之间的合著最为普遍，在 40 个评估国家的所有合著出版物中占比约 70%。

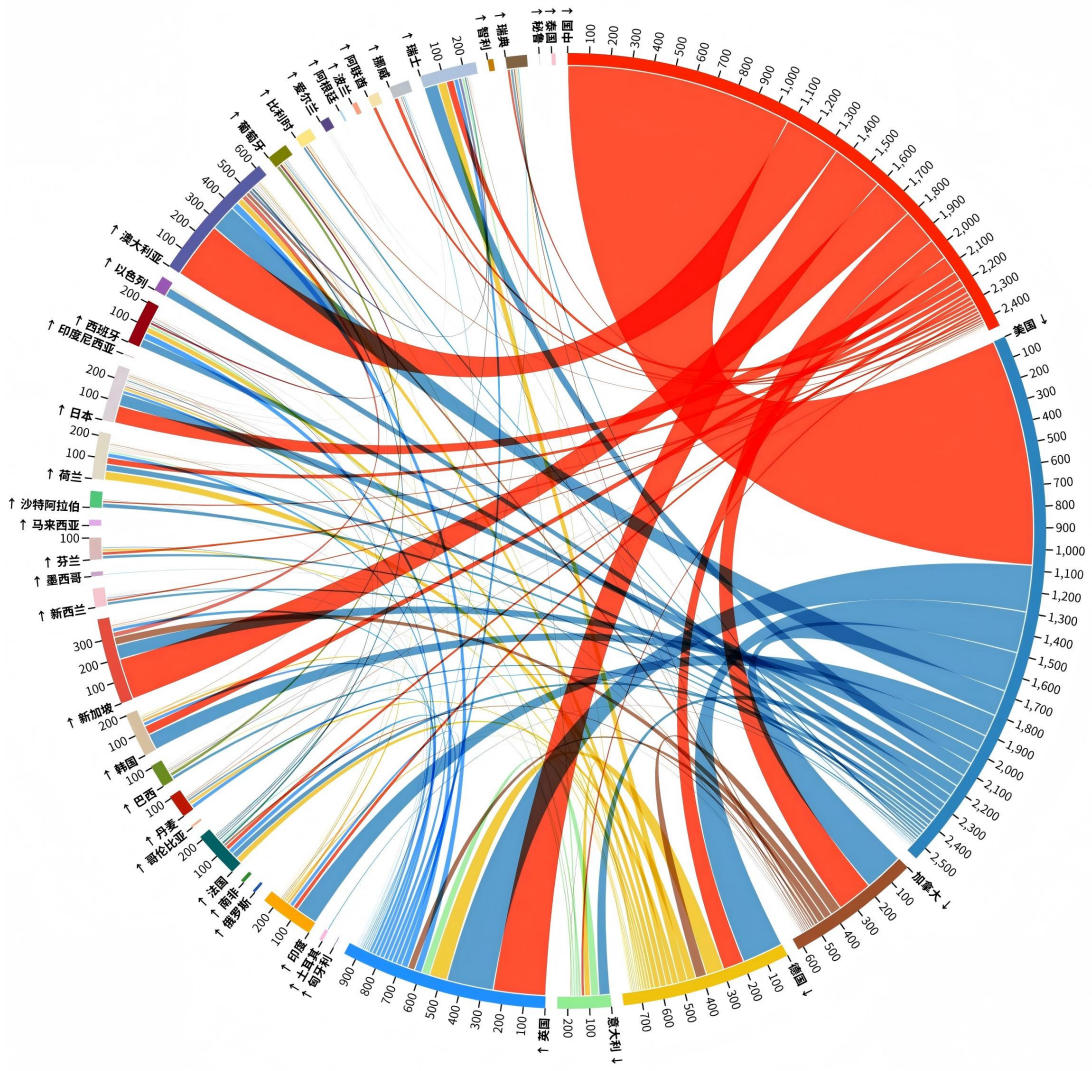


图 25 40 国 AI 治理相关出版物数量与作者国际合作情况

数据来源：基于对 DBLP 文献数据库的统计分析（2024 年 4 月至 2025 年 3 月期间）

对 AI 治理相关出版物合作情况的分析揭示了一个显著的国际合作模式。中国、美国、加拿大、德国、英国和澳大利亚在该领域的合作研究中处于领先地位，在所调研的 40 个国家中，这些国家之间的合著出版物约占 40 个评估国家全部合著出版物的 70%。与此同时，在这 40 个国家的 AI 治理相关出版物总量中，仅有 14.7%涉及这些国家相互之间的国际合著，这凸显了进一步深化 AI 治理研究协作的必要性。

主要发现 4.8：

美国和中国在面向可持续发展目标（SDGs）的 AI 研究中处于主导地位，在 40 个国家中共同贡献了约 60%的相关出版物。与此同时，其他国家也在不同可持续发展目标领域做出了重要贡献。

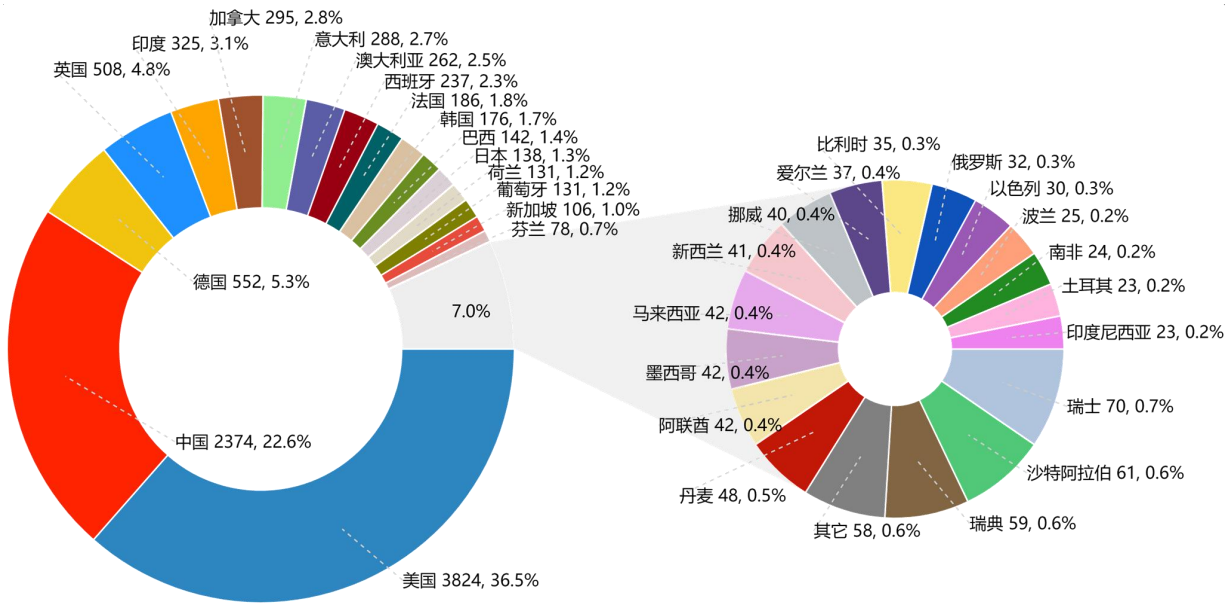


图 26 各国在可持续发展目标相关的 AI 出版物上的占比

数据来源：基于对 DBLP 文献数据库的统计分析（2024 年 4 月至 2025 年 3 月期间）

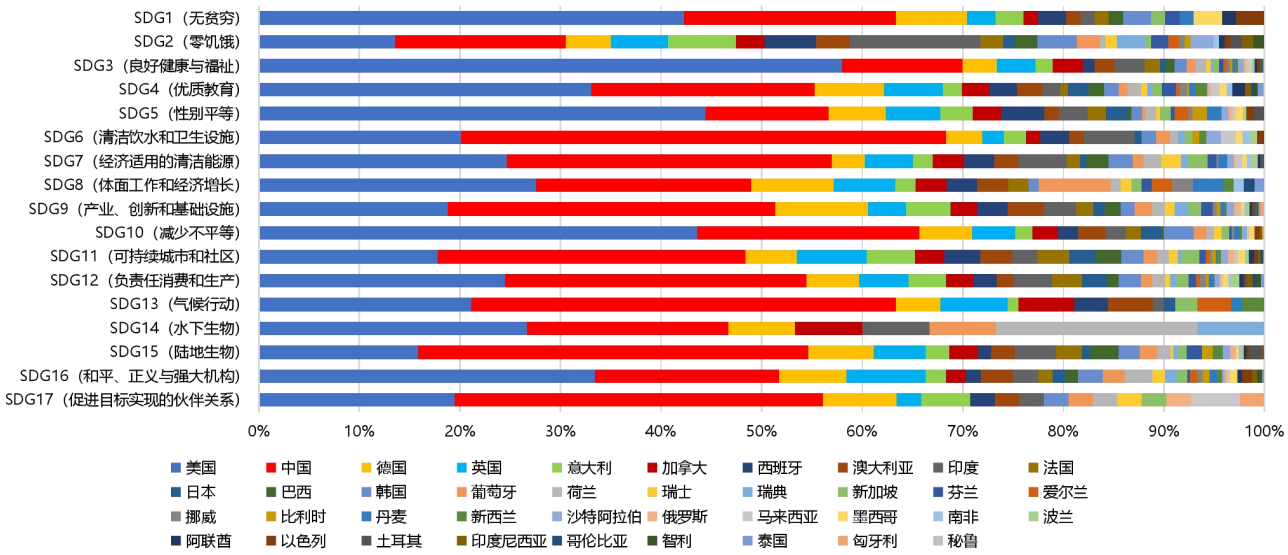


图 27 各国 AI 出版物在不同可持续发展目标上的分布情况

数据来源：基于对 DBLP 文献数据库的统计分析（2024 年 4 月至 2025 年 3 月期间）

40 个国家在推动 AI 促进实现可持续发展目标方面贡献各异，展现了它们在“智能向善”方面的不同努力和侧重。总体来看，美国和中国合计贡献了占 40 国总量 59.1%的相关出版物，德国和英国也在几乎所有 AI 与可持续发展目标领域发挥了重要作用。与此同时，其他国家也在不同目标领域做出了重要贡献。例如，印度在 AI 与 SDG2（零饥饿）、SDG6（清洁饮水和卫生设施）、SDG7（经济适用的清洁能源）、SDG14（水下生物）等领域的研究中贡献可观；荷兰在 AI 与 SDG14（水下生物）领域亦有显著表现。

主要发现 4.9：

SDG3（良好健康与福祉）、SDG11（可持续城市和社区）以及 SDG9（产业、创新和基础设施）是当前研究关注度最高的 AI 赋能可持续发展目标方向。同时，高收入国家组在社会层面的目标方向上分配了更高的 AI 研究关注度，而中高收入和中低收入国家组在经济、环境相关的目标上有着相对更高的 AI 研究关注度。

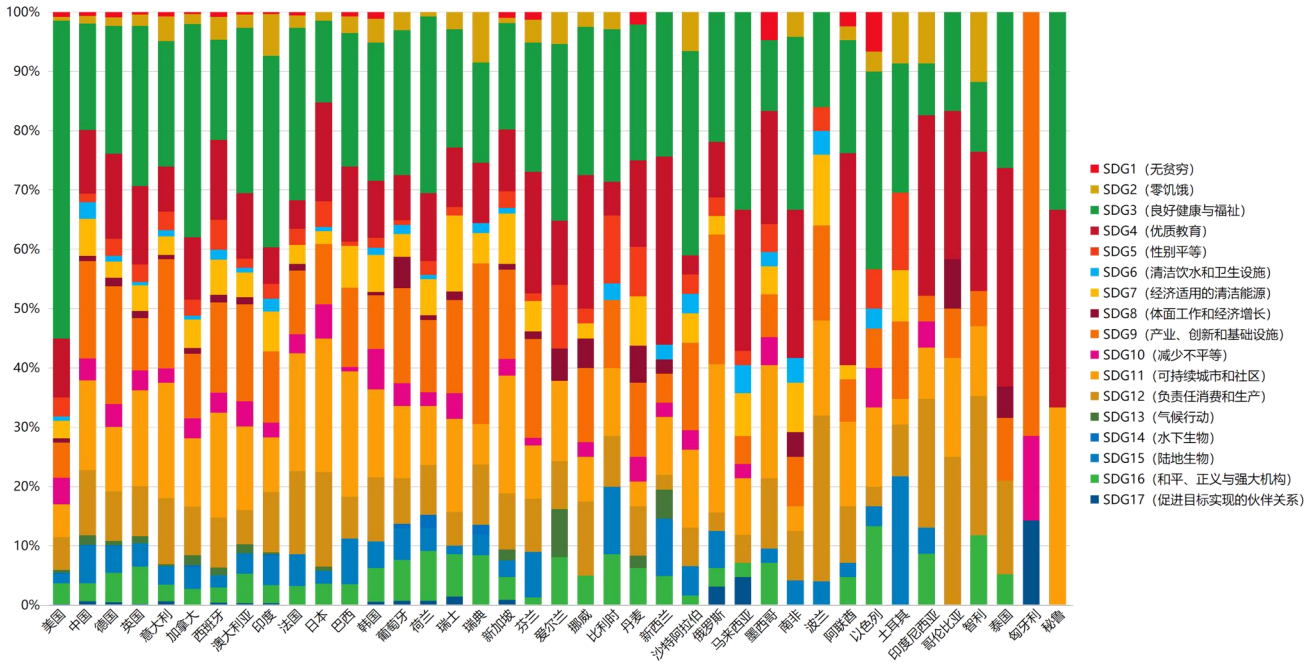


图 28 各国 AI 促进可持续发展目标相关研究中不同目标的分布情况

数据来源：基于对 DBLP 文献数据库的统计分析（2024 年 4 月至 2025 年 3 月期间）

40 个国家在 AI 促进不同可持续发展目标的研究关注度上存在一定共性。其中，SDG3（良好健康与福祉）是最受关注的议题，其次依次为 SDG11（可持续城市和社区）、SDG9（产业、创新和基础设施）、SDG4（优质教育）及 SDG12（负责任消费和生产）。

同时，对比两个国家组在 17 个可持续发展目标的研究关注度分配可见，高收入国家组在 SDG3（良好健康与福祉）、SDG5（性别平等）、SDG16（和平、正义与强大机构）等目标方向上的关注度相对高于中高收入和中低收入国家组；而后者则在 SDG6（清洁饮水和卫生设施）、SDG7（经济适用的清洁能源）、SDG9（产业、创新和基础设施）、SDG11（可持续城市和社区）等目标方向上有着相对更高的研究关注度。

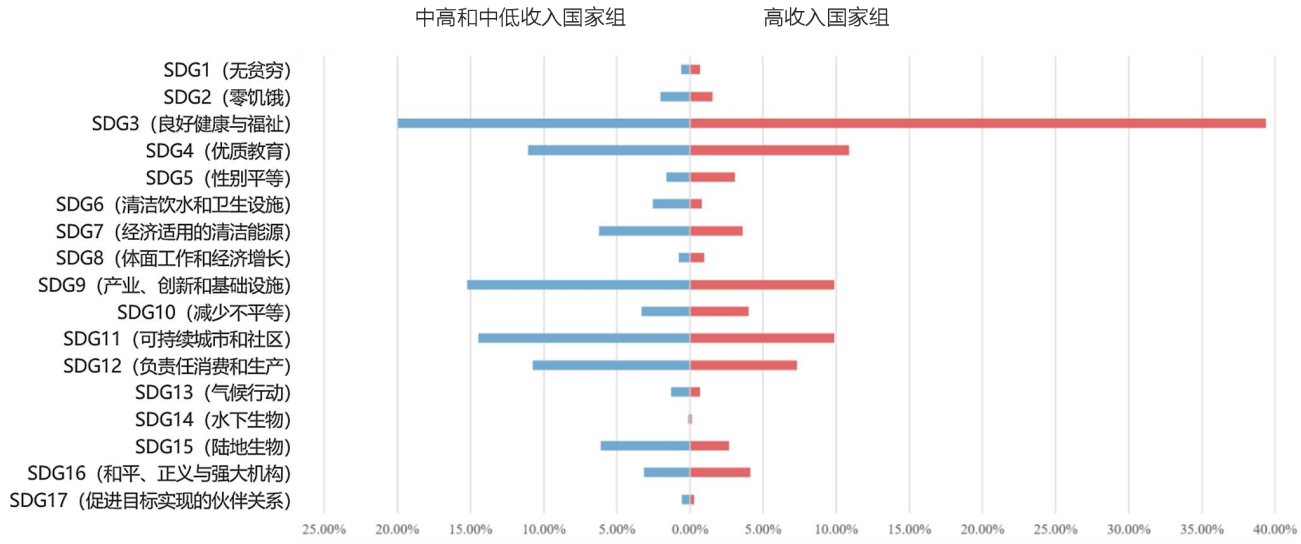


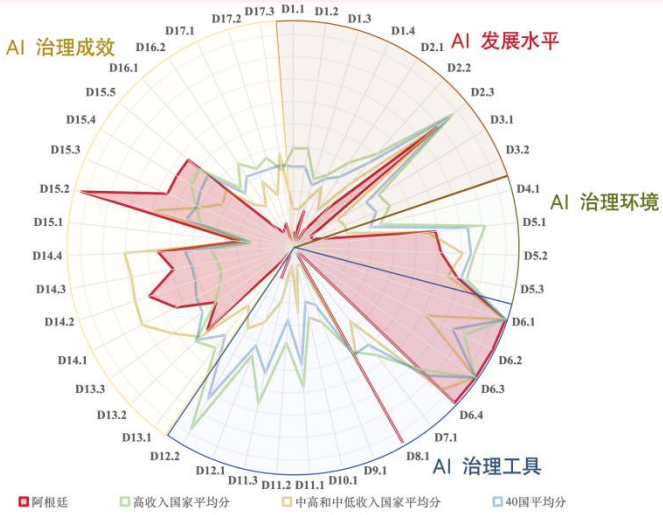
图 29 高收入国家组与中高和中低收入国家组在不同可持续发展目标上的 AI 研究关注度分配对比

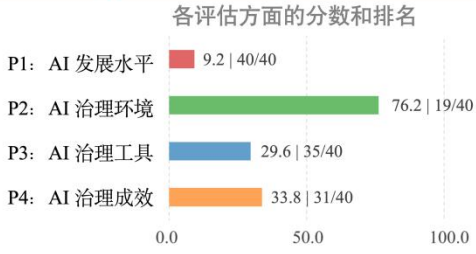
数据来源：基于对 DBLP 文献数据库的统计分析（2024 年 4 月至 2025 年 3 月期间；横轴代表两个国家组在各个可持续发展目标上的 AI 出版物在所有与可持续发展目标相关的 AI 出版物中的占比）

IV.
国家概况

阿根廷

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
38/40	4600万	17,144 $	中高收入



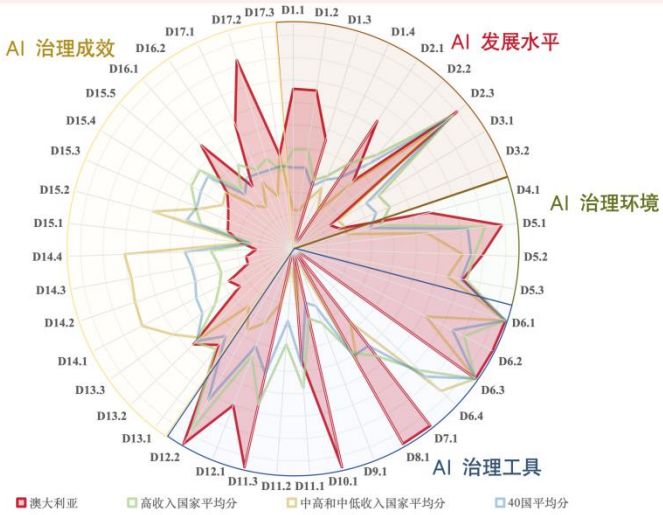


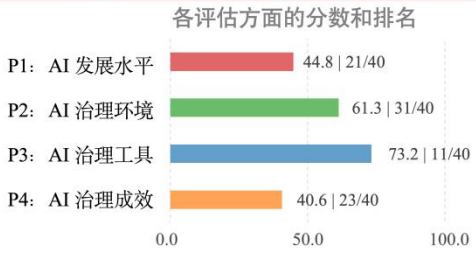
表现较好的部分：
D15 AI 发展包容度（排名：3/40）
D4 AI 风险暴露度（排名：9/40）

表现欠佳的部分：
D13 公众AI认知度（排名：40/40）
D17 AI 治理研究活跃度（排名：40/40）

澳大利亚

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
14/40	2700万	81,355 $	高收入



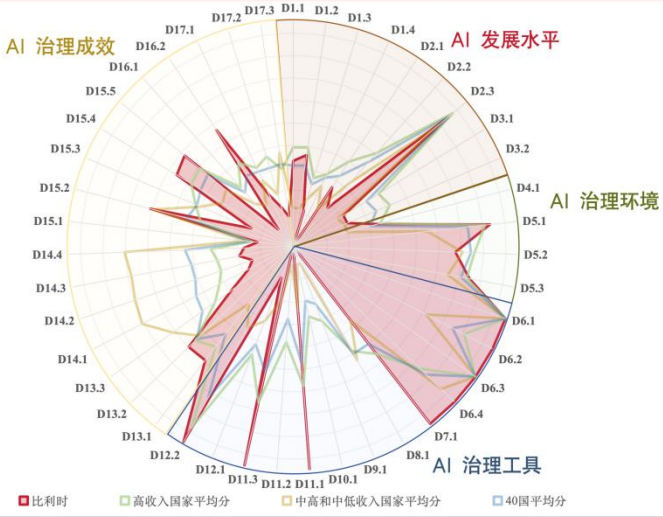


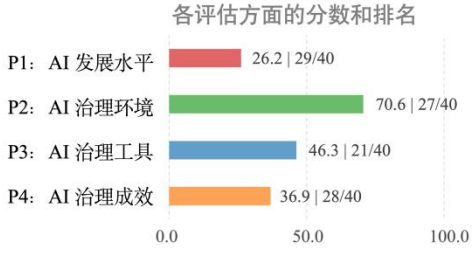
表现较好的部分：
D17 AI 治理研究活跃度（排名：3/40）
D1 AI 研发创新（排名：9/40）

表现欠佳的部分：
D14 AI 社会接受度（排名：35/40）
D4 AI 风险暴露度（排名：34/40）

比利时

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
28/40	1200万	59,317 $	高收入



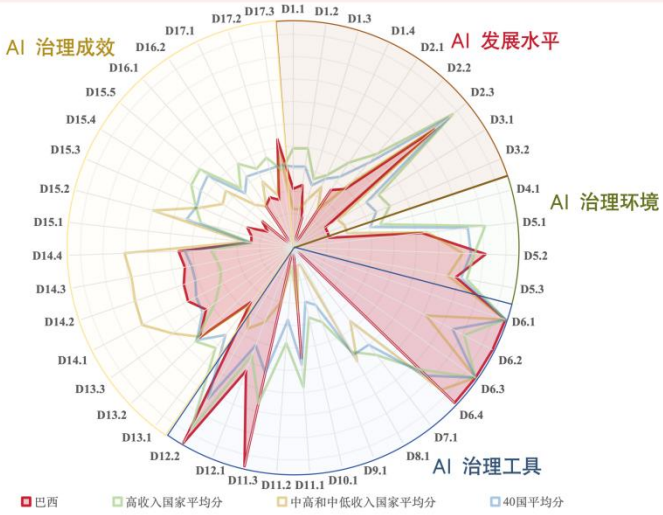


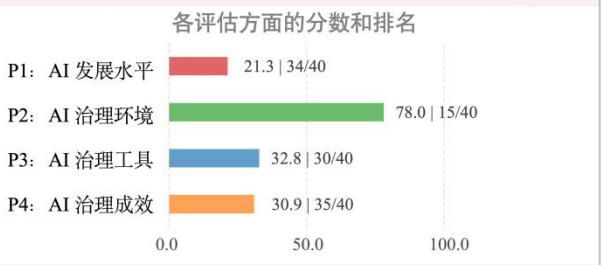
表现较好的部分：
D13 公众AI认知度（排名：16/40）
D16 数据算法开放度（排名：16/40）

表现欠佳的部分：
D14 AI 社会接受度（排名：36/40）
D17 AI 治理研究活跃度（排名：36/40）

巴西

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
36/40	2.12亿	12,263 $	中高收入



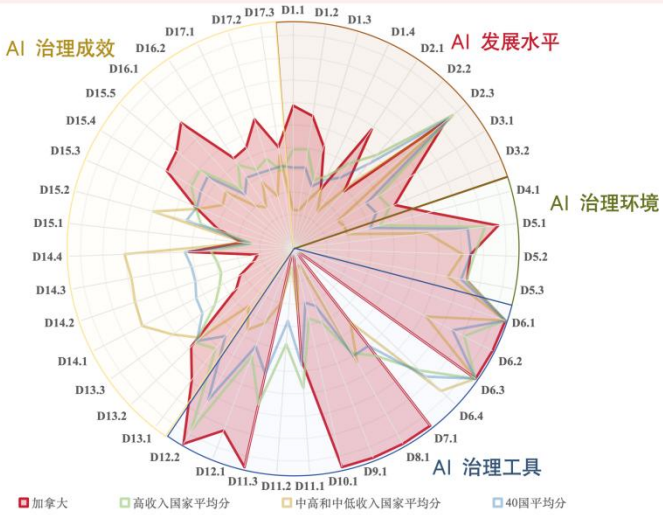


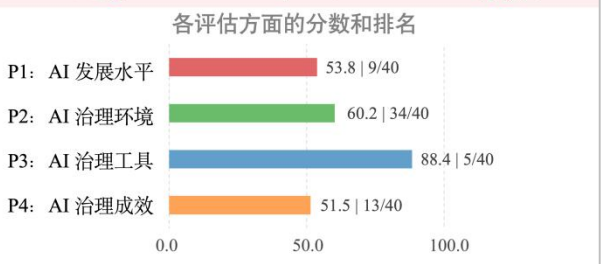
表现较好的部分：
D4 AI 风险暴露度 （排名：11/40）
D14 AI 社会接受度（排名：18/40）

表现欠佳的部分：
D15 AI 发展包容度（排名：37/40）
D3 AI 产业活力（排名：36/40）

加拿大

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
8/40	4000万	59,564 $	高收入



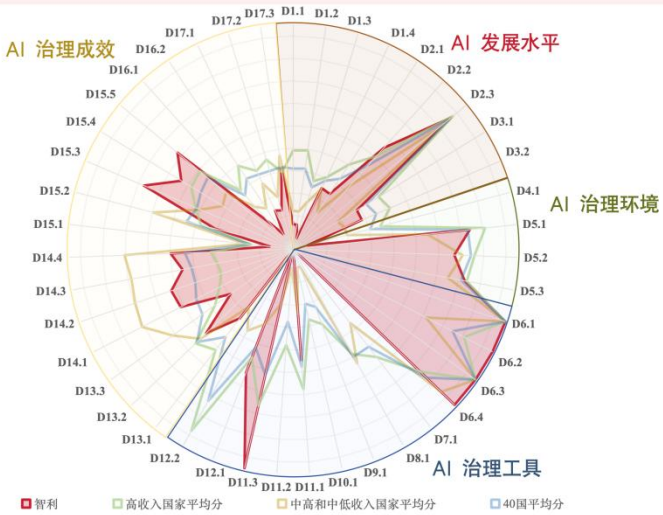


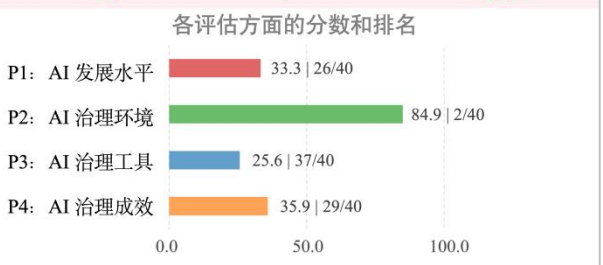
表现较好的部分：
P3 AI 治理工具（排名：5/40）
D1 AI 研发创新（排名：7/40）

表现欠佳的部分：
D4 AI 风险暴露度 （排名：35/40）
D14 AI 社会接受度（排名：31/40）

智利

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
29/40	2000万	18,820 $	高收入



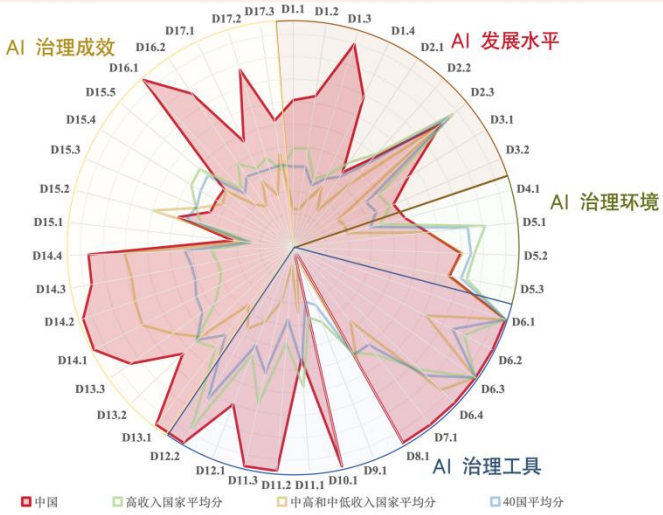


表现较好的部分：
D4 AI 风险暴露度 （排名：4/40）
D14 AI 社会接受度（排名：14/40）

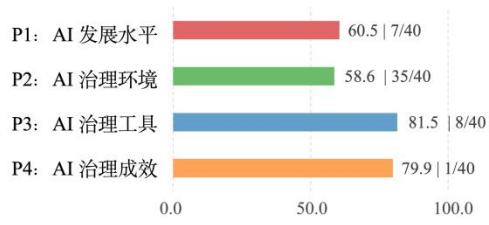
表现欠佳的部分：
P3 AI 治理工具（排名：37/40）
D13 公众AI认知度（排名：35/40）

中国

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
1/40	14.51亿	15,745 $	中高收入



各评估方面的分数和排名

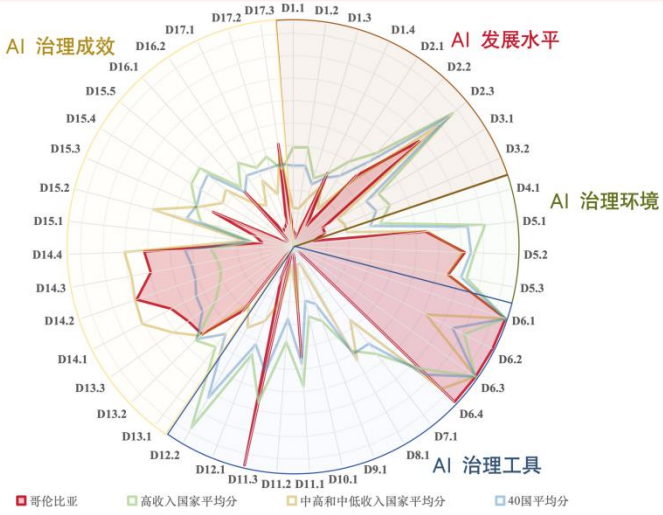


表现较好的部分：
D13 公众AI认知度（排名：1/40）
D14 AI 社会接受度（排名：1/40）

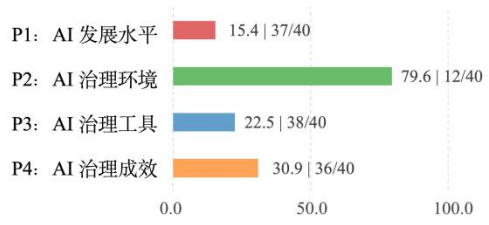
表现欠佳的部分：
D5 AI 治理准备度（排名：32/40）
D4 AI 风险暴露度 （排名：30/40）

哥伦比亚

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
39/40	5300万	8,973 $	中高收入



各评估方面的分数和排名

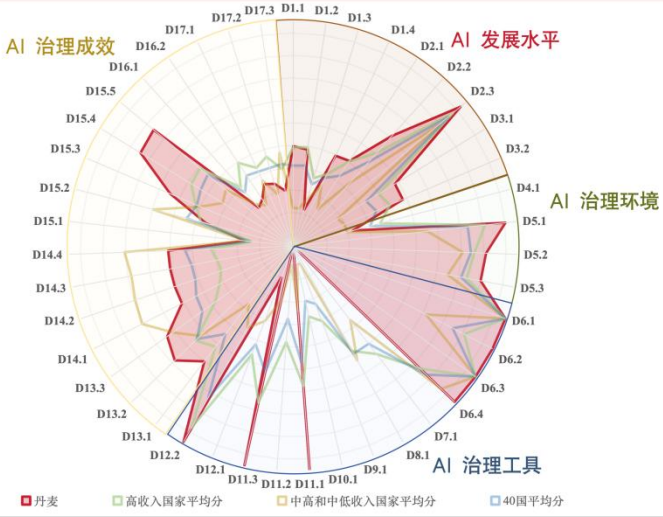


表现较好的部分：
D4 AI 风险暴露度 （排名：5/40）
D14 AI 社会接受度（排名：9/40）

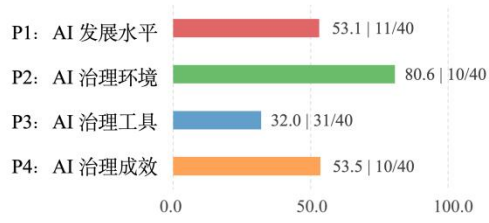
表现欠佳的部分：
D15 AI 发展包容度（排名：40/40）
P3 AI 治理工具（排名：38/40）

丹麦

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
15/40	597.7万	80,759 $	高收入



各评估方面的分数和排名

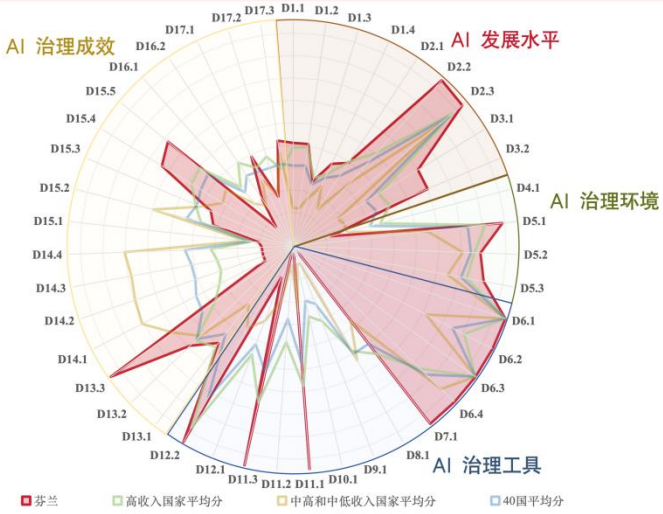


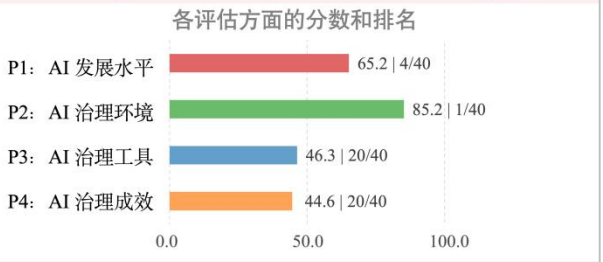
表现较好的部分：
D5 AI 治理准备度（排名：1/40）
D13 公众AI认知度（排名：4/40）

表现欠佳的部分：
P3 AI 治理工具（排名：31/40）
D17 AI 治理研究活跃度 （排名：28/40）

芬兰

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
10/40	561.7万	60,739 $	高收入



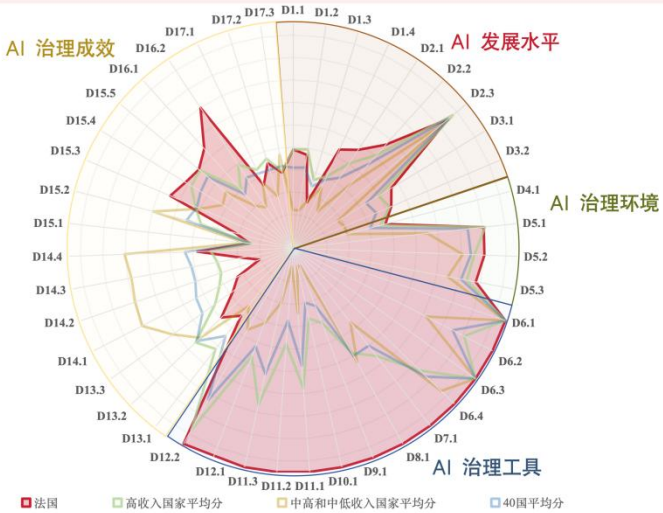


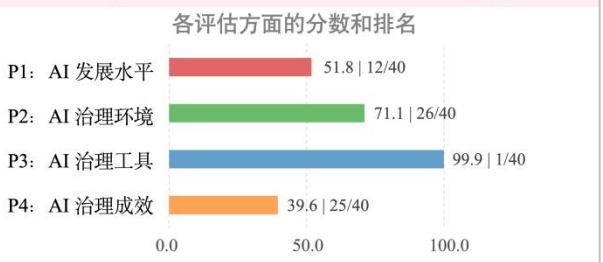
表现较好的部分：
D5 AI 治理准备度（排名：2/40）
D2 AI 基础设施（排名：3/40）

表现欠佳的部分：
D14 AI 社会接受度（排名：39/40）
D16 数据算法开放度（排名：24/40）

法国

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
7/40	6700万	53,101 $	高收入



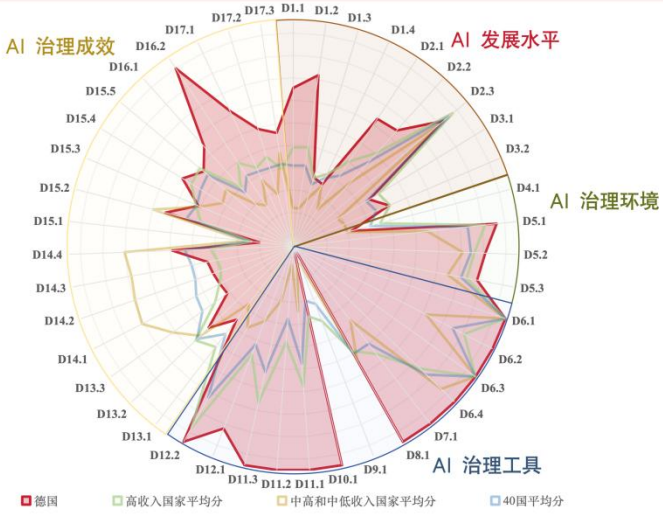


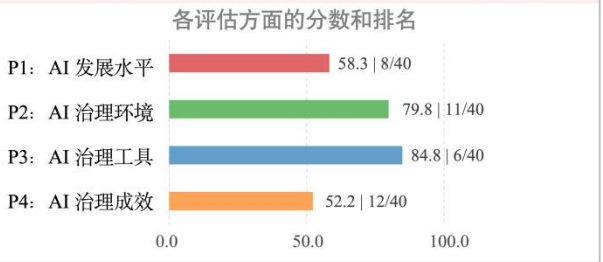
表现较好的部分：
P3 AI 治理工具（排名：1/40）
D16 数据算法开放度（排名：6/40）

表现欠佳的部分：
D13 公众AI认知度（排名：38/40）
D14 AI 社会接受度（排名：30/40）

德国

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
3/40	7200万	67,226 $	高收入



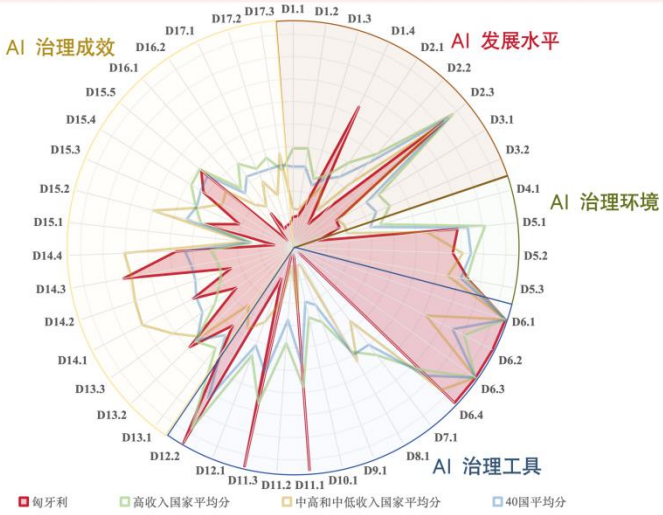


表现较好的部分：
D5 AI 治理准备度（排名：3/40）
D16 数据算法开放度（排名：3/40）

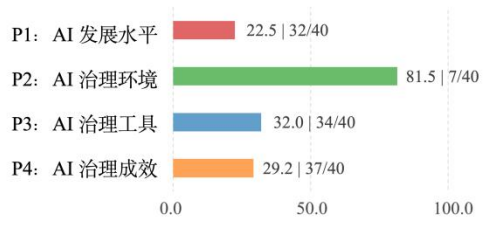
表现欠佳的部分：
D13 公众AI认知度（排名：33/40）
D14 AI 社会接受度（排名：21/40）

匈牙利

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
34/40	1000万	21,468 $	高收入



各评估方面的分数和排名

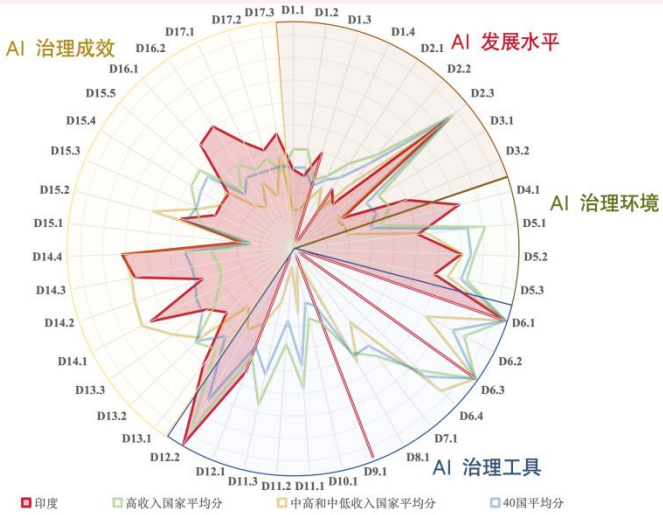


表现较好的部分：
D4 AI 风险暴露度 （排名：8/40）
D14 AI 社会接受度（排名：17/40）

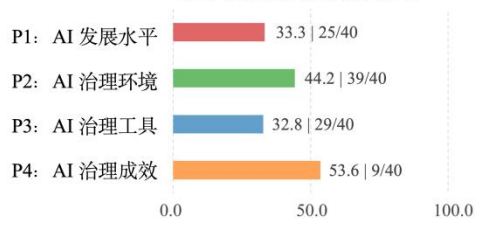
表现欠佳的部分：
D17 AI 治理研究活跃度 （排名：39/40）
D2 AI 基础设施（排名：37/40）

印度

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
35/40	14.51亿	2,932 $	中低收入



各评估方面的分数和排名

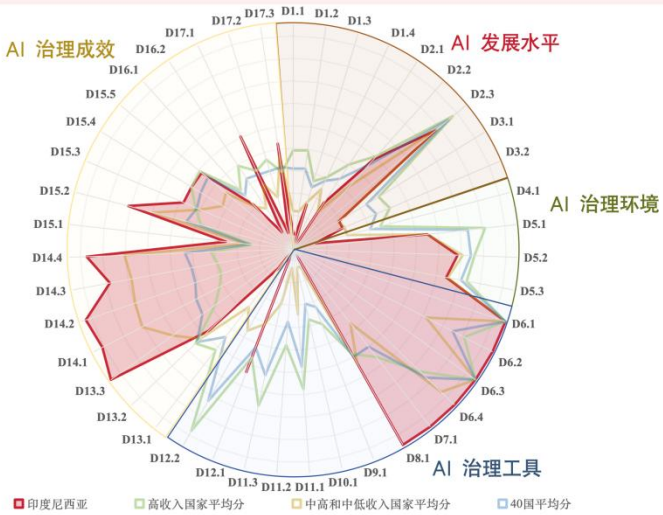


表现较好的部分：
D16 数据算法开放度（排名：7/40）
D17 AI 治理研究活跃度 （排名：8/40）

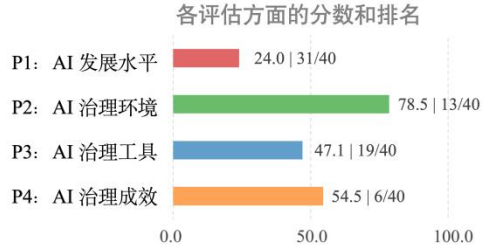
表现欠佳的部分：
D5 AI 治理准备度（排名：39/40）
D4 AI 风险暴露度 （排名：38/40）

印度尼西亚

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
22/40	2.83亿	5,531 $	中高收入



各评估方面的分数和排名

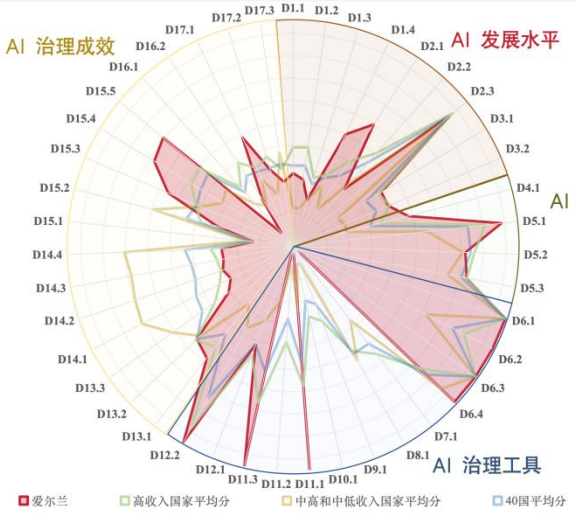


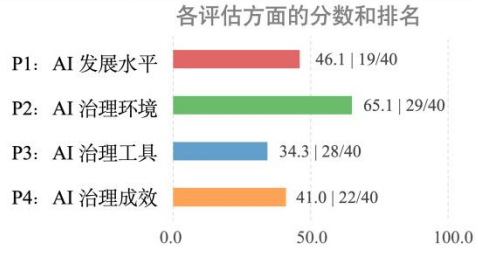
表现较好的部分：
D14 AI 社会接受度（排名：3/40）
D4 AI 风险暴露度 （排名：6/40）

表现欠佳的部分：
D1 AI 研发创新（排名：37/40）
D5 AI 治理准备度（排名：36/40）

爱尔兰

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
24/40	525.5万	122,243 $	高收入



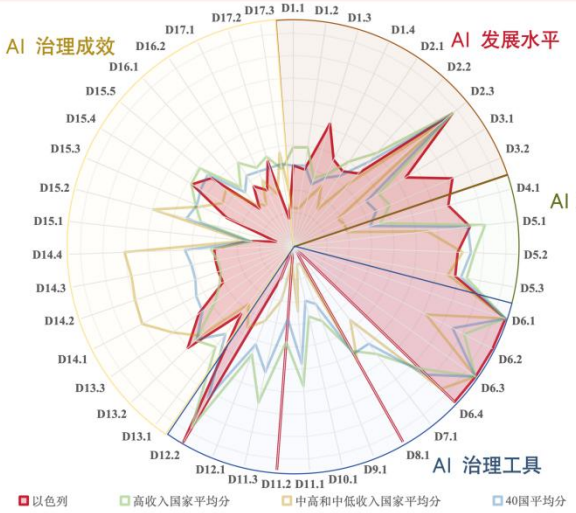


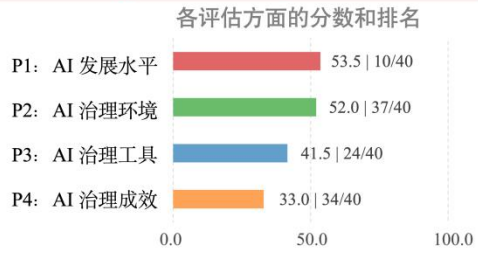
表现较好的部分：
D15 AI 发展包容度（排名：12/40）
D3 AI 产业活力（排名：14/40）

表现欠佳的部分：
D4 AI 风险暴露度 （排名：31/40）
D16 数据算法开放度（排名：28/40）

以色列

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
27/40	525.5万	68,992 $	高收入



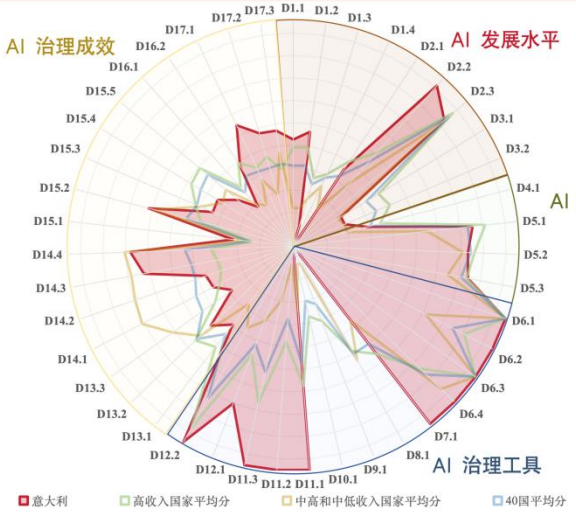


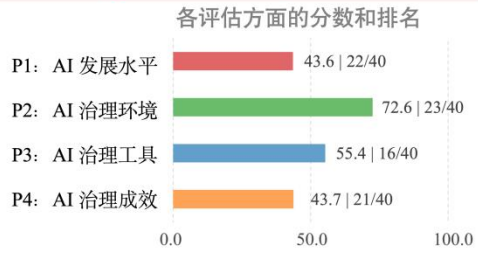
表现较好的部分：
D3 AI 产业活力（排名：5/40）
D1 AI 研发创新（排名：11/40）

表现欠佳的部分：
D4 AI 风险暴露度 （排名：37/40）
D17 AI 治理研究活跃度 （排名：30/40）

意大利

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
19/40	5900万	44,479 $	高收入



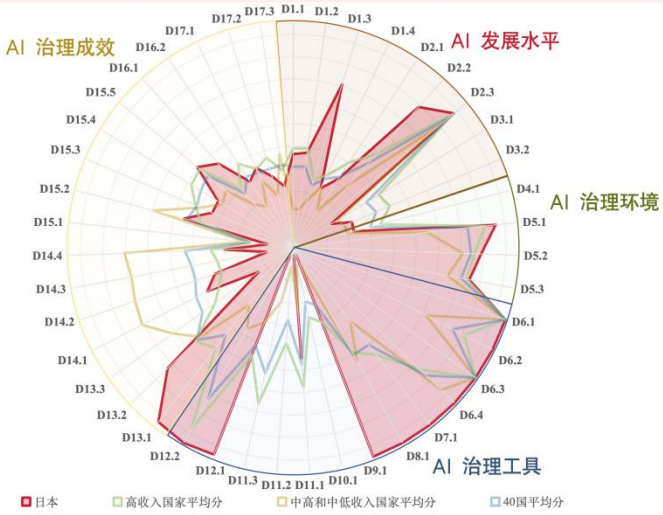


表现较好的部分：
D2 AI 基础设施（排名：4/40）
D17 AI 治理研究活跃度 （排名：6/40）

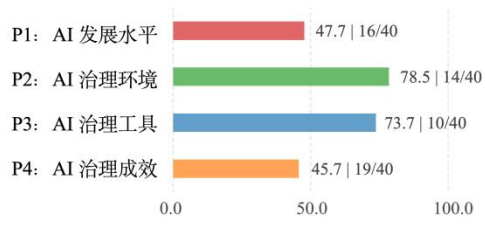
表现欠佳的部分：
D13 公众AI认知度（排名：34/40）
D3 AI 产业活力（排名：27/40）

日本

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
9/40	12400万	49,620 $	高收入



各评估方面的分数和排名

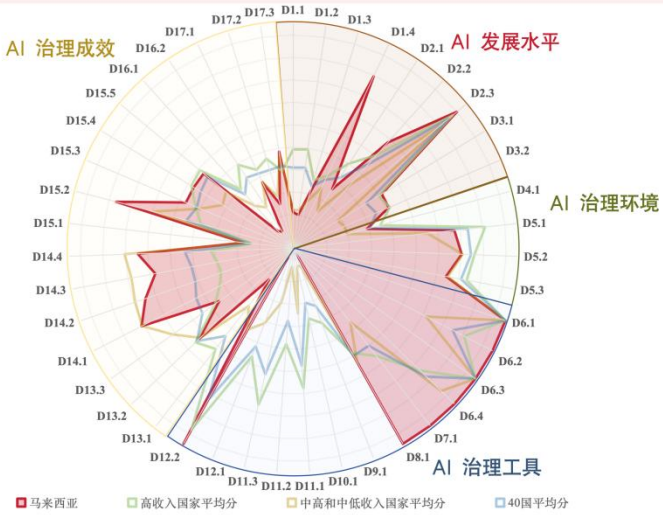


表现较好的部分：
D13 公众AI认知度（排名：6/40）
D2 AI 基础设施（排名：7/40）

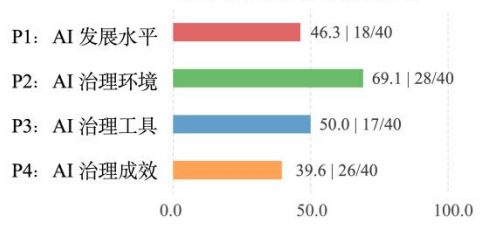
表现欠佳的部分：
D14 AI 社会接受度（排名：29/40）
D3 AI 产业活力（排名：28/40）

马来西亚

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
21/40	3600万	15,003 $	中高收入



各评估方面的分数和排名

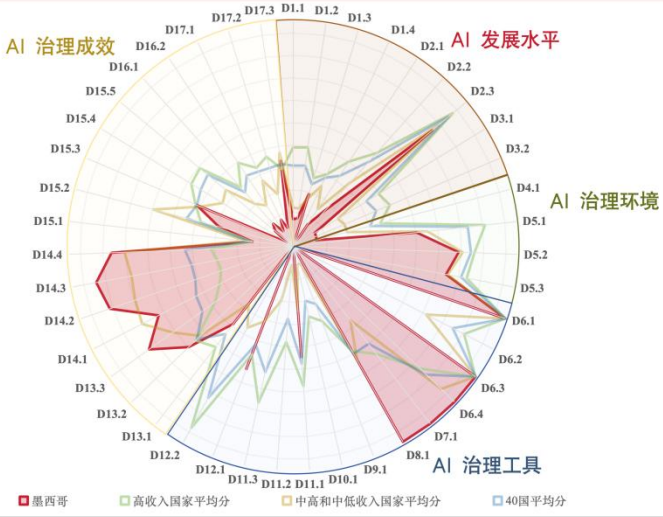


表现较好的部分：
D14 AI 社会接受度（排名：8/40）
D3 AI 产业活力（排名：13/40）

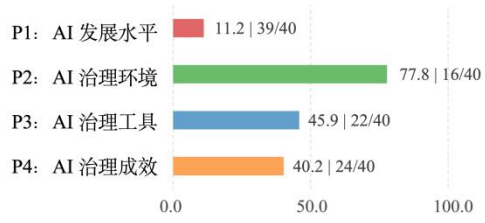
表现欠佳的部分：
D13 公众AI认知度（排名：37/40）
D16 数据算法开放度（排名：37/40）

墨西哥

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
31/40	1.31亿	13,482 $	高收入



各评估方面的分数和排名

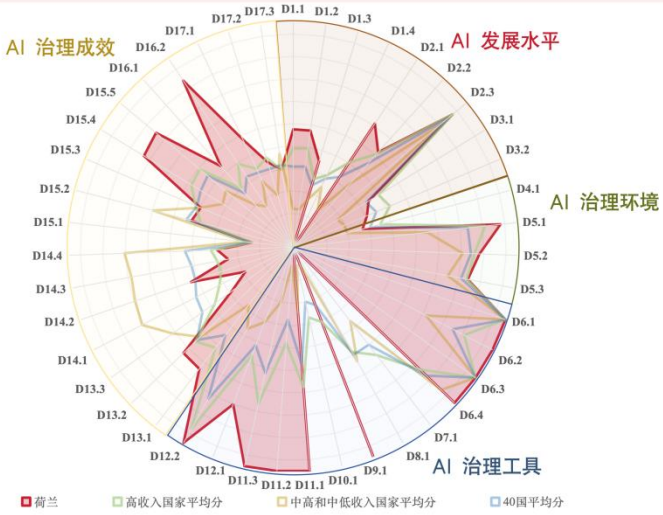


表现较好的部分：
D14 AI 社会接受度（排名：5/40）
D4 AI 风险暴露度（排名：6/40）

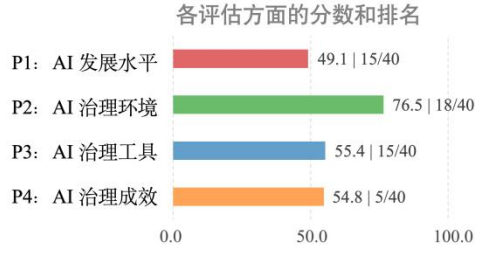
表现欠佳的部分：
D2 AI 基础设施（排名：40/40）
D3 AI 产业活力（排名：39/40）

荷兰

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
11/40	1800万	65,357 $	高收入



各评估方面的分数和排名

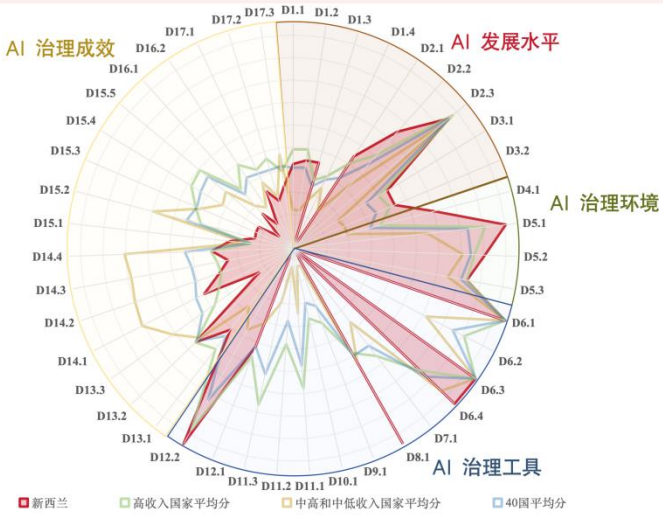


表现较好的部分：
D16 数据算法开放度（排名：5/40）
D2 AI 基础设施（排名：6/40）

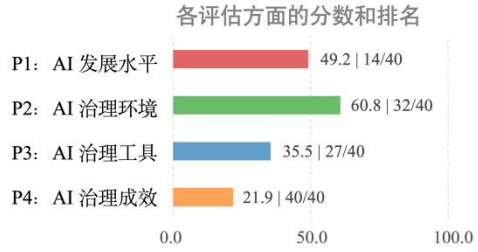
表现欠佳的部分：
D14 AI 社会接受度（排名：26/40）
D4 AI 风险暴露度（排名：21/40）

新西兰

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
33/40	521.4万	55,591 $	高收入



各评估方面的分数和排名

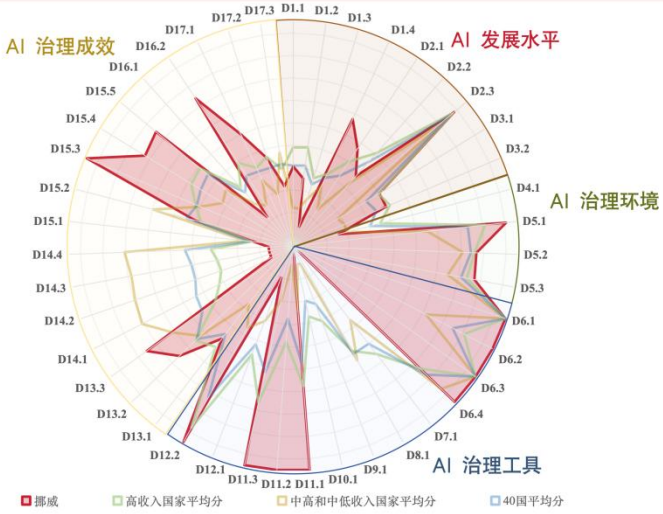


表现较好的部分：
D5 AI 治理准备度（排名：4/40）
D2 AI 基础设施（排名：8/40）

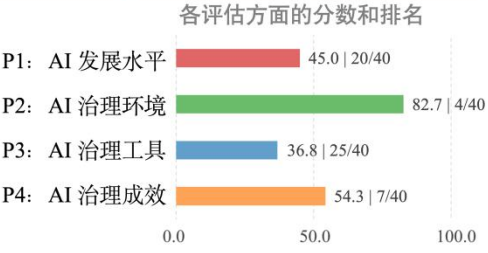
表现欠佳的部分：
D15 AI 发展包容度（排名：38/40）
D4 AI 风险暴露度（排名：36/40）

挪威

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
17/40	557.7万	103,915 $	高收入



各评估方面的分数和排名

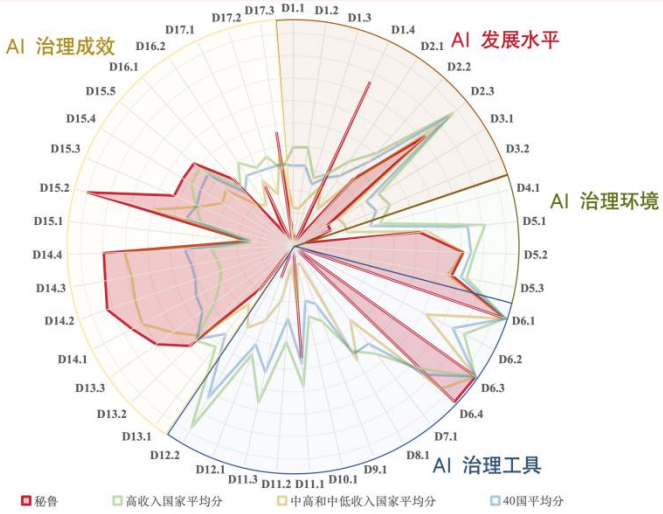


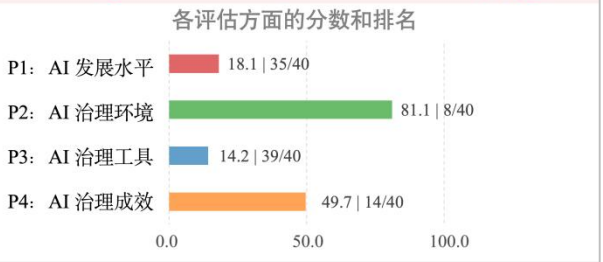
表现较好的部分：
D15 AI 发展包容度（排名：1/40）
D13 公众AI认知度（排名：5/40）

表现欠佳的部分：
D14 AI 社会接受度（排名：40/40）
P3 AI 治理工具（排名：25/40）

秘鲁

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
37/40	3400万	8,620 $	中高收入



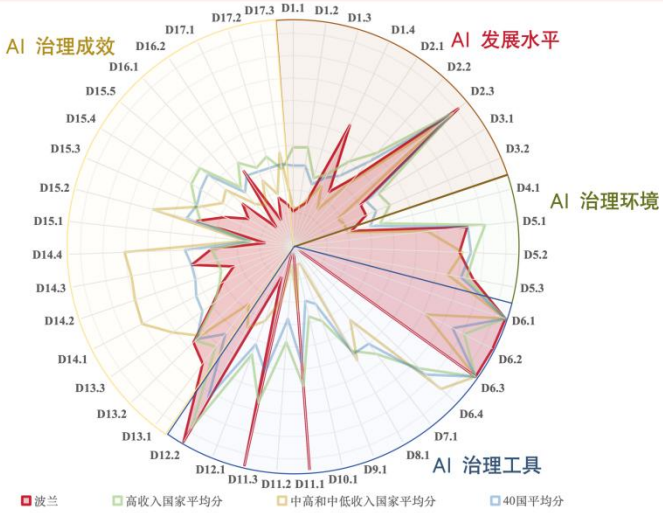


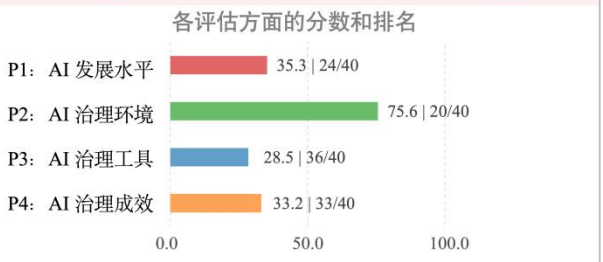
表现较好的部分：
D4 AI 风险暴露度 （排名：1/40）
D14 AI 社会接受度（排名：4/40）

表现欠佳的部分：
D16 数据算法开放度（排名：40/40）
P3 AI 治理工具（排名：39/40）

波兰

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
32/40	3900万	21,397 $	高收入



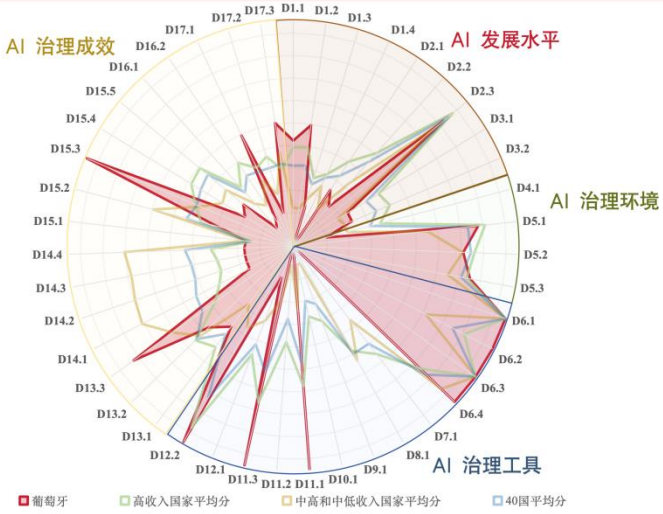


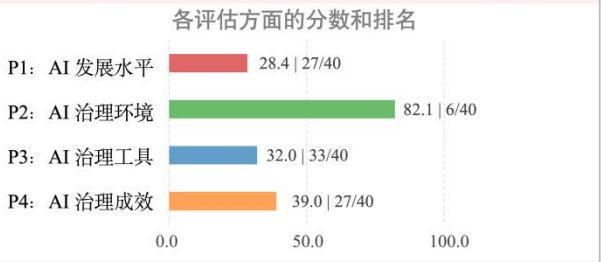
表现较好的部分：
D13 公众AI认知度（排名：10/40）
D4 AI 风险暴露度 （排名：16/40）

表现欠佳的部分：
D17 AI 治理研究活跃度 （排名：37/40）
P3 AI 治理工具（排名：36/40）

葡萄牙

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
26/40	1000万	30,048 $	高收入



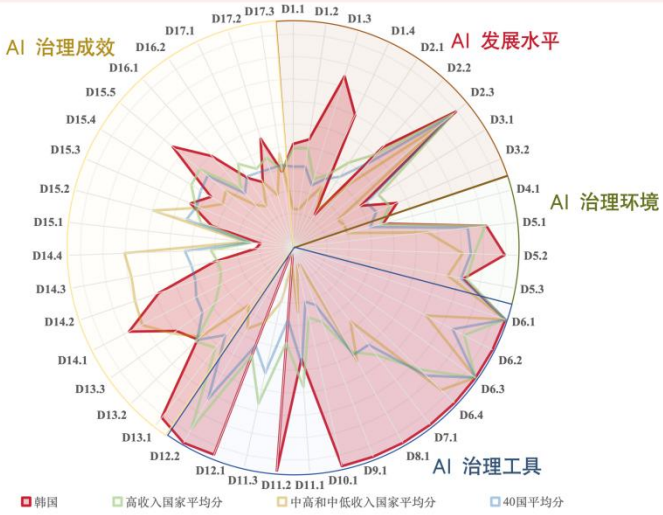


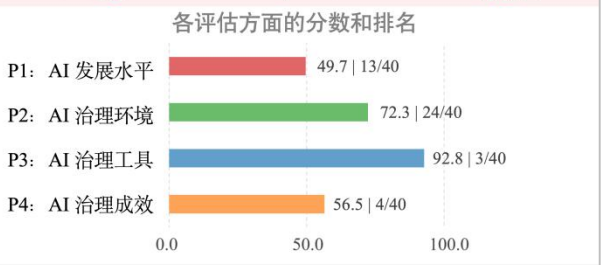
表现较好的部分：
D4 AI 风险暴露度 （排名：9/40）
D13 公众AI认知度（排名：9/40）

表现欠佳的部分：
D14 AI 社会接受度（排名：37/40）
D2 AI 基础设施（排名：33/40）

韩国

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
4/40	5200万	45,377 $	高收入



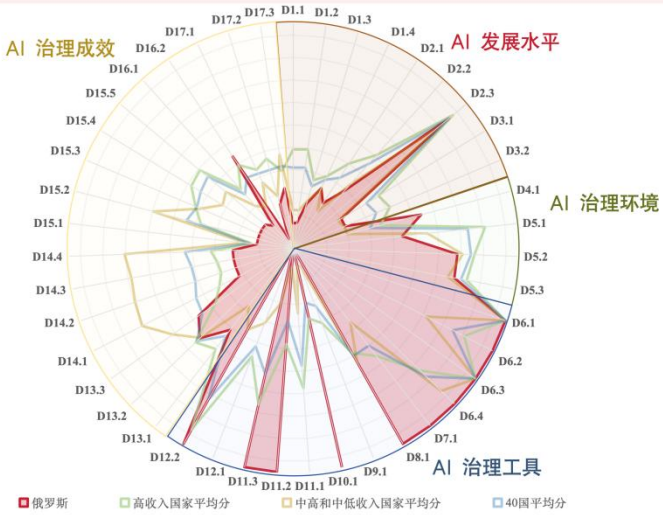


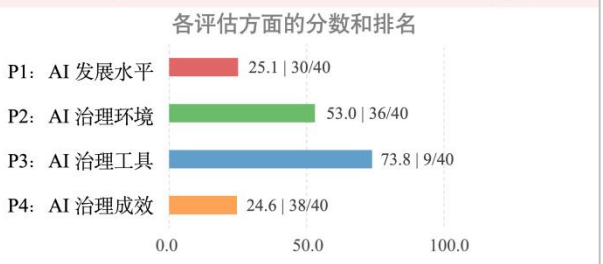
表现较好的部分：
D13 公众AI认知度（排名：2/40）
P3 AI 治理工具（排名：3/40）

表现欠佳的部分：
D4 AI 风险暴露度（排名：26/40）
D2 AI 基础设施（排名：24/40）

俄罗斯

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
30/40	14500万	14,001 $	高收入



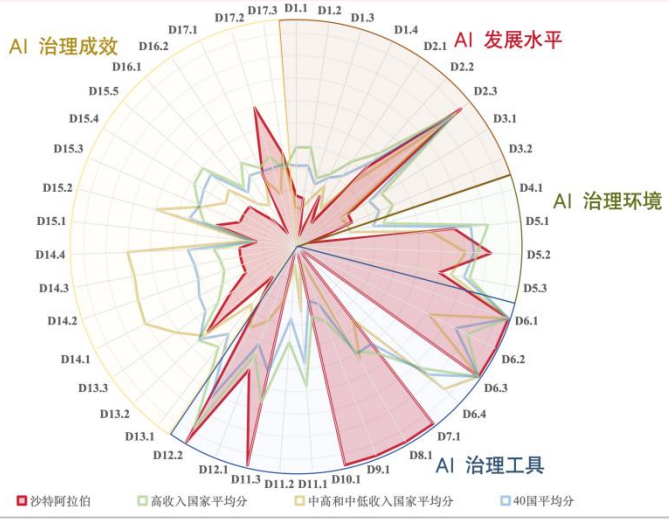


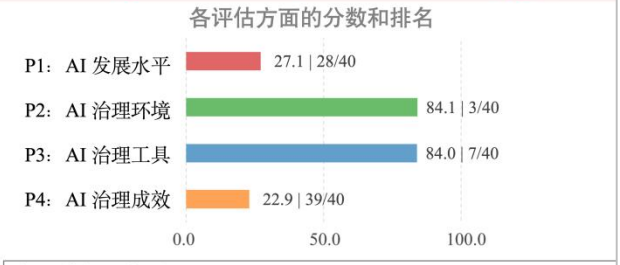
表现较好的部分：
P3 AI 治理工具（排名：9/40）
D16 数据算法开放度（排名：18/40）

表现欠佳的部分：
D15 AI 发展包容度（排名：39/40）
D17 AI 治理研究活跃度（排名：38/40）

沙特阿拉伯

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
18/40	3400万	30,393 $	高收入



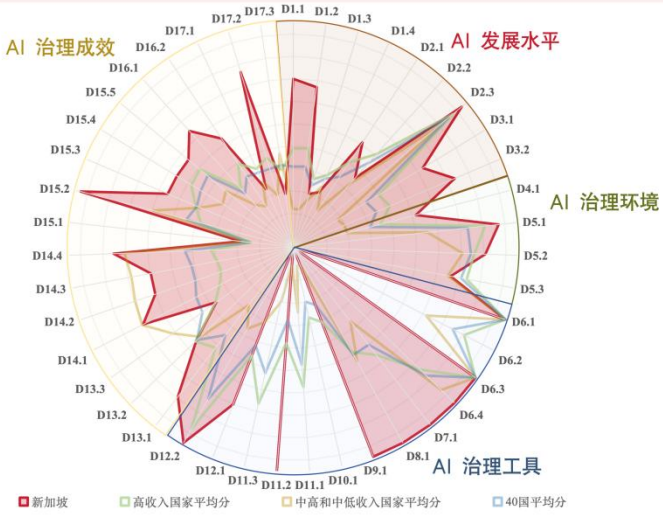


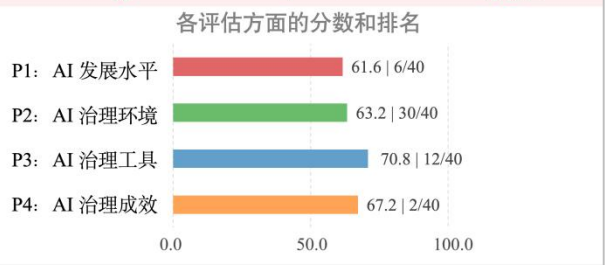
表现较好的部分：
D4 AI 风险暴露度（排名：1/40）
P3 AI 治理工具（排名：7/40）

表现欠佳的部分：
D13 公众AI认知度（排名：39/40）
D14 AI 社会接受度（排名：33/40）

新加坡

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
6/40	583.2万	88,285 $	高收入



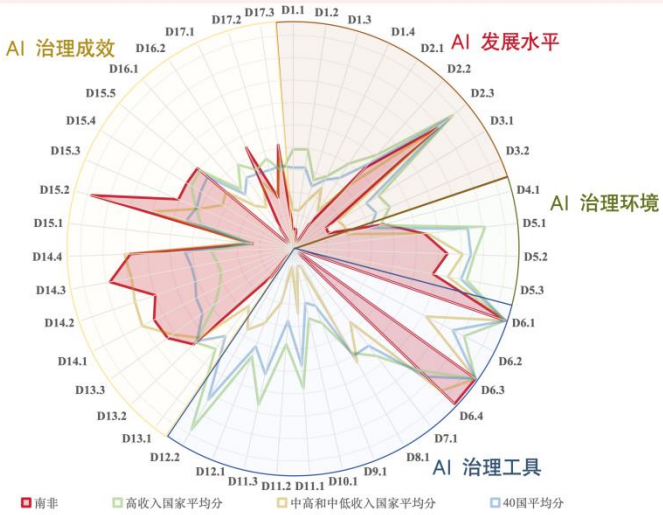


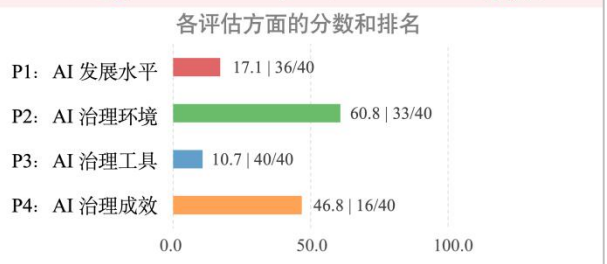
表现较好的部分：
D3 AI 产业活力（排名：3/40）
D15 AI 发展包容度（排名：4/40）

表现欠佳的部分：
D4 AI 风险暴露度 （排名：32/40）
D5 AI 治理准备度（排名：17/40）

南非

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
40/40	6400万	7,540 $	高收入



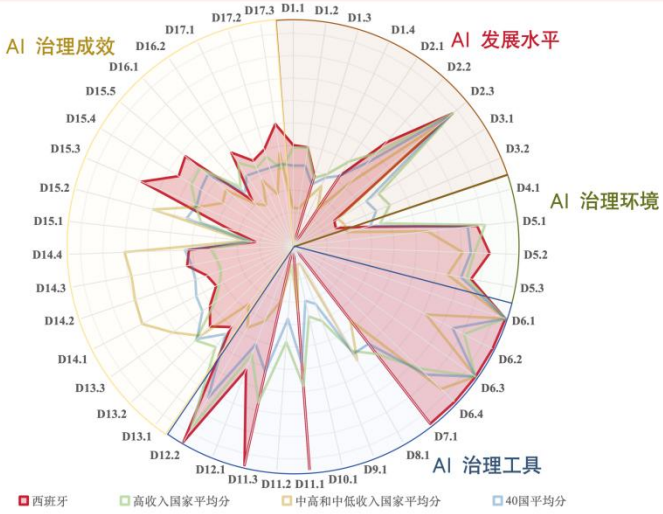


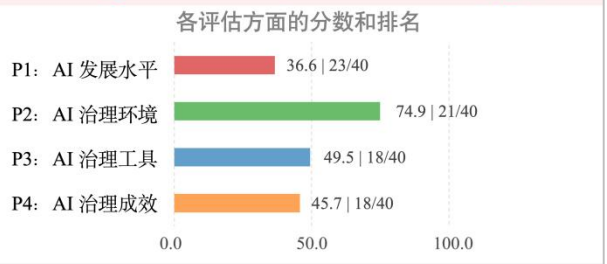
表现较好的部分：
D14 AI 社会接受度（排名：6/40）
D15 AI 发展包容度（排名：9/40）

表现欠佳的部分：
D1 AI 研发创新（排名：40/40）
D5 AI 治理准备度（排名：40/40）
P3 AI 治理工具（排名：40/40）

西班牙

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
20/40	4800万	37,663 $	高收入



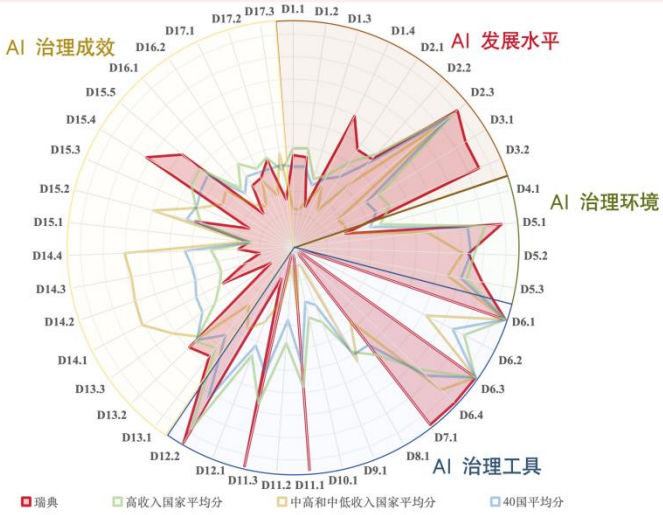


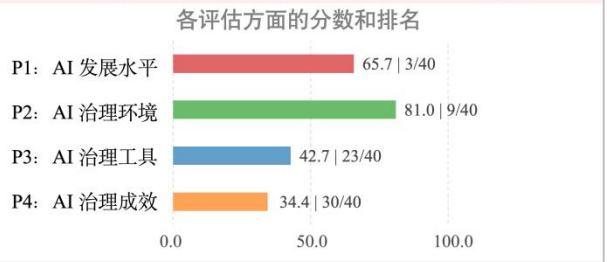
表现较好的部分：
D17 AI 治理研究活跃度 （排名：9/40）
D5 AI 治理准备度（排名：14/40）

表现欠佳的部分：
D3 AI 产业活力（排名：32/40）
D13 公众AI认知度（排名：28/40）

瑞典

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
13/40	1800万	73,353 $	高收入



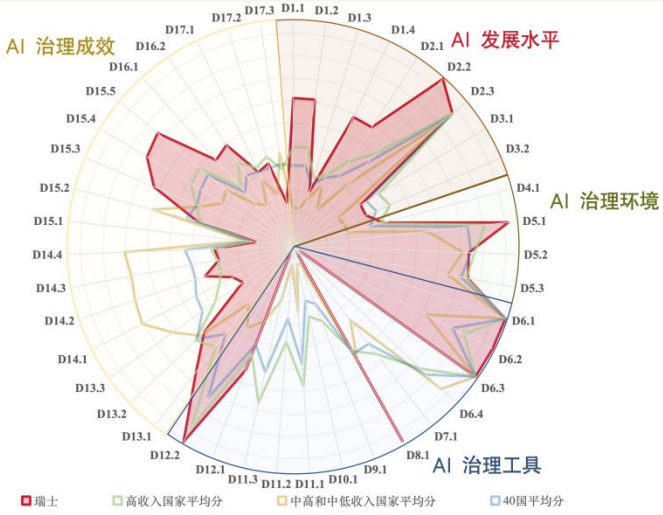


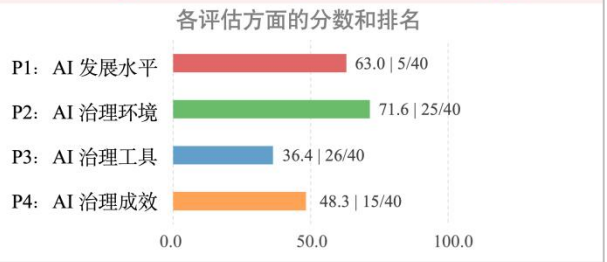
表现较好的部分：
D3 AI 产业活力（排名：2/40）
D5 AI 治理准备度（排名：8/40）

表现欠佳的部分：
D14 AI 社会接受度（排名：34/40）
D13 公众AI认知度（排名：30/40）

瑞士

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
16/40	892.2万	118,656 $	高收入



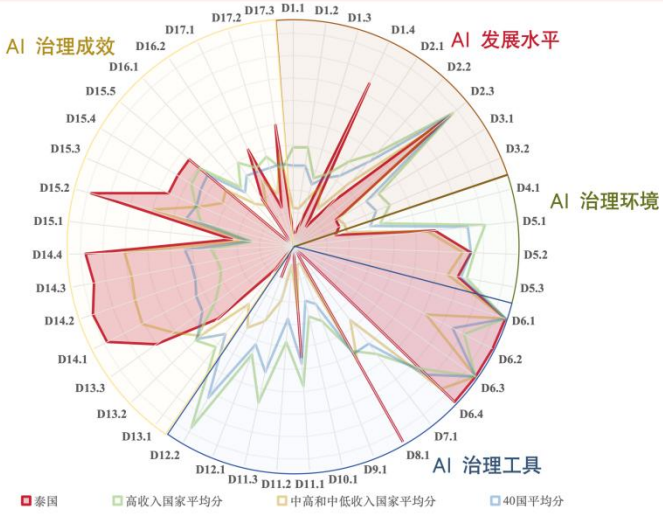


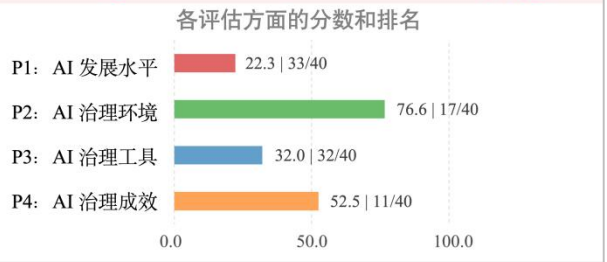
表现较好的部分：
D2 AI 基础设施（排名：2/40）
D1 AI 研发创新（排名：5/40）

表现欠佳的部分：
D4 AI 风险暴露度 （排名：26/40）
D17 AI 治理研究活跃度 （排名：26/40）

泰国

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
25/40	7200万	8,508 $	中高收入



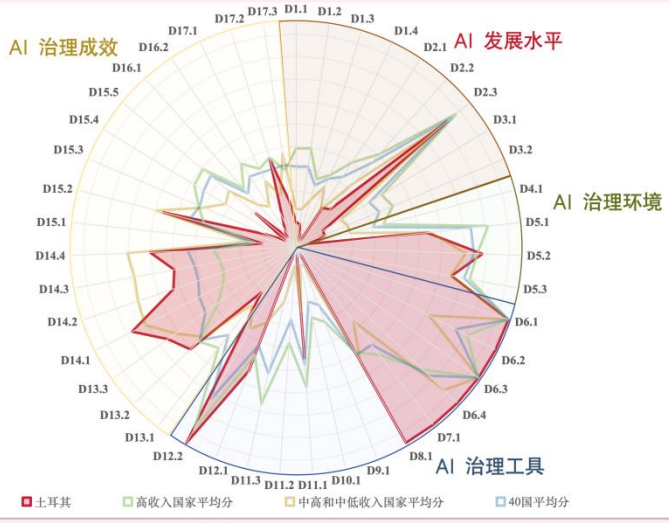


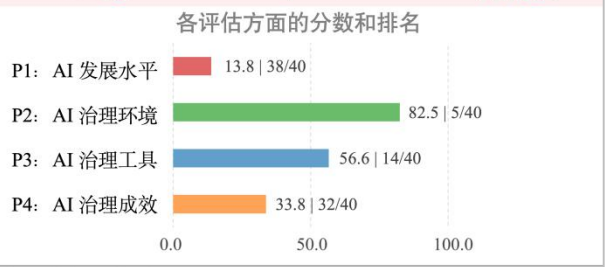
表现较好的部分：
D14 AI 社会接受度（排名：2/40）
D15 AI 发展包容度（排名：5/40）

表现欠佳的部分：
D2 AI 基础设施（排名：35/40）
D13 公众AI认知度（排名：32/40）

土耳其

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
23/40	8700万	18,981 $	中高收入



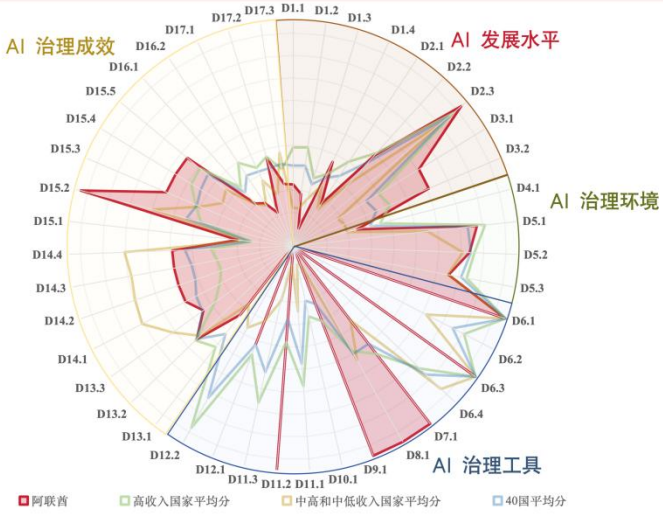


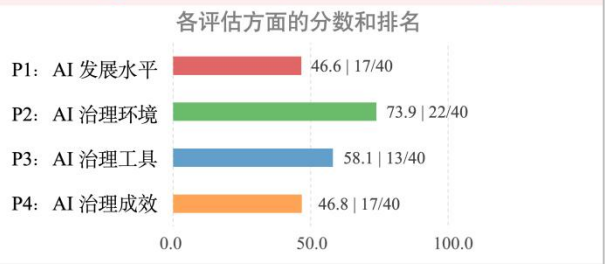
表现较好的部分：
D4 AI 风险暴露度（排名：1/40）
D14 AI 社会接受度（排名：11/40）

表现欠佳的部分：
D1 AI 研发创新（排名：39/40）
D16 数据算法开放度（排名：38/40）

阿联酋

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
12/40	1100万	53,808 $	高收入



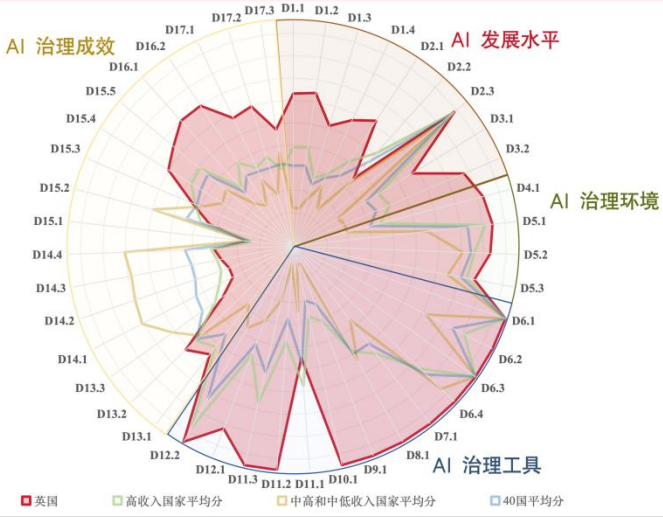


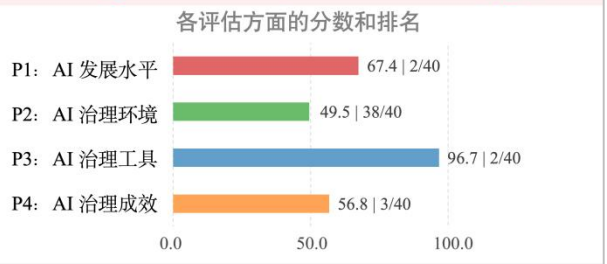
表现较好的部分：
D15 AI 发展包容度（排名：2/40）
D3 AI 产业活力（排名：6/40）

表现欠佳的部分：
D1 AI 研发创新（排名：28/40）
D17 AI 治理研究活跃度（排名：27/40）

英国

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
5/40	6900万	61,804 $	高收入



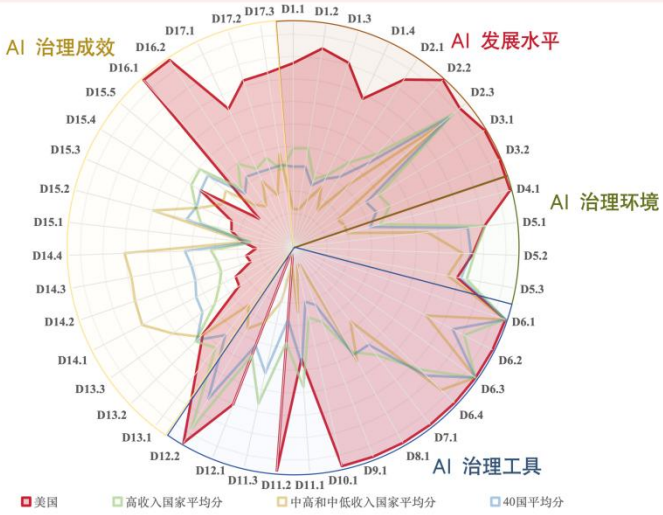


表现较好的部分：
P3 AI 治理工具（排名：2/40）
D1 AI 研发创新（排名：3/40）

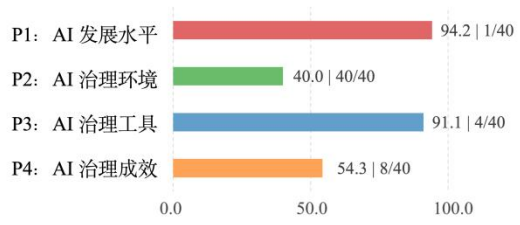
表现欠佳的部分：
D4 AI 风险暴露度（排名：39/40）
D14 AI 社会接受度（排名：27/40）

美国

AGILE 指数排名	人口（2024）	人均GDP（2024）	国家分类
2/40	3.45亿	83,845 $	高收入



各评估方面的分数和排名



表现较好的部分：

D1 AI 研发创新（排名：1/40）
D2 AI 基础设施（排名：1/40）
D3 AI 产业活力（排名：1/40）
D16 数据算法开放度（排名：1/40）
D17 AI 治理研究活跃度（排名：1/40）

表现欠佳的部分：

D4 AI 风险暴露度 （排名：40/40）
D14 AI 社会接受度（排名：38/40）

附录

1. 指标体系与数据来源

表 5 AGILE 指数 2025：评估维度和评估指标（详细）

评估方面	评估维度	评估内容	联合国教科文组织《人工智能伦理问题建议书》(UNESCO Rec.)与《准备状态评估方法》（UNESCO RAM）对应条款[2]	评估指标 （数据来源及统计周期详见表后说明）
P1. AI 发展水平	D1. AI 研发创新	评估各国在人工智能相关研发方面的活跃度	UNESCO Rec. A83, UNESCO RAM4.2.1	D1.1. AI 相关出版物数量&人均数量
				D1.2. AI 领域活跃研究者数量&人口占比
				D1.3. AI 专利授权数量&人均数量
				D1.4. 大型 AI 系统研发数量& 与 GDP 比例
	D2. AI 基础设施	评估各国人工智能技术及数字生态系统基础设施的部署/发展和访问/获取水平	UNESCO Rec. A59, A80, UNESCO RAM6.2.1, 6.2.3	D2.1. 托管数据中心数量&人均数量
				D2.2. 国家超级计算机总浮点运算能力&人均算力
				D2.3. 国家信息与通信技术发展水平评估表现
	D3. AI 产业活力	评估各国在人工智能产业方面的投融资活跃度	UNESCO Rec. A117, UNESCO RAM5.2.3	D3.1. AI 领域私人投资&与 GDP 比例
				D3.2. 新获融资的 AI 企业数量&与 GDP 比例
P2. AI 治理环境	D4. AI 风险暴露度	评估各国在人工智能相关的伦理安全风险暴露程度/治理压力	UNESCO Rec. A50	D4.1. AI 相关风险案例 / 事故数量&与 GDP 比例
	D5. AI 治理准备度	评估各国在人工智能治理方面的准备程度和治理基础	UNESCO Rec. A54, UNESCO RAM2.2.2, UNESCO RAM2.2.8	D5.1. 国家治理整体水平评估表现
				D5.2. 国家数字治理整体水平评估表现
				D5.3. 国家可持续发展整体水平评估表现
P3. AI 治理工具	D6. AI 战略规划	评估各国人工智能战略 / 规划 / 路线图的制定情况及完备性	UNESCO Rec. A56, A71, UNESCO RAM 2.2.1	D6.1. 是否发布了国家层面的 AI 战略
				D6.2. AI 战略是否包含实施计划
				D6.3. AI 战略是否提及（劳动力）培训或技能提升
				D6.4. AI 战略是否明确提及伦理影响
	D7. AI 治理机构	评估各国人工智能治理机构的建立情况	UNESCO Rec. A58, UNESCO RAM1.4	D7.1. 是否已建立或指定了国家 AI 治理机构

[2] 本列参考内容指向两个文件中与 AGILE 指数具体维度的评估内容密切对应的条款：一是联合国教科文组织《人工智能伦理问题建议书》（以下简称 UNESCO Rec.） “IV. 政策行动领域”部分的相关内容，二是联合国教科文组织《准备状态评估方法》（*Readiness Assessment Methodology*，以下简称 UNESCO RAM）中的相关指标。

	D8. AI 原则规范	评估各国人工智能治理原则和规范的制定情况	UNESCO Rec. A48	D8.1. 是否发布了国家层面的 AI 原则或规范
	D9. AI 影响评估	评估各国人工智能影响评估工具/框架的完善程度	UNESCO Rec. A50	D9.1. 是否出台了官方的 AI 影响评估机制/工具
	D10. AI 标准认证	评估各国人工智能标准/认证机制的设立情况	UNESCO Rec. A64	D10.1. 是否制定了国家层面的AI标准/认证机制
	D11. AI 立法现状	评估各国人工智能相关法律法规的制定情况	UNESCO Rec. A133, UNESCO RAM2.2.2	D11.1. 是否已制定或正在制定国家层面专门针对 AI 的综合性法律法规
				D11.2. 是否已制定国家层面专门针对 AI 垂直领域的法律法规
				D11.3. 是否已实施国家层面与 AI 相关的数据/信息保护法
	D12. AI 治理国际参与	评估各国参与人工智能国际治理的程度	UNESCO Rec. A80, UNESCO RAM6.2.2	D12.1. 参与国际 AI 治理行动/机制的程度
				D12.2. 参与 ISO AI 标准制定的程度
P4. AI 治理成效	D13. 公众 AI 认知度	评估各国公众的人工智能能力与对人工智能风险影响的认知水平	UNESCO Rec. A101, UNESCO RAM4.2.1, 4.2.2	D13.1. 公众 AI 素养的基础水平
				D13.2. 公众对 AI 议题的讨论程度
				D13.3. 公众对 AI 影响的认知程度
	D14. AI 社会接受度	评估各国对人工智能技术与应用的社会接受程度	UNESCO Rec. A39, UNESCO RAM3.2.2 & 3.3.4	D14.1. 公众对 AI 发展的积极态度
				D14.2. 公众对 AI 使用的积极或中性预期
				D14.3. 公众对 AI 应用的信任程度
				D14.4. 企业采用 AI 的意愿
	D15. AI 发展包容度	评估各国人工智能研发与应用相关的包容性水平	UNESCO Rec. A91, A105, UNESCO RAM3.2.1	D15.1. AI 领域活跃研究者的性别比例
				D15.2. 互联网使用的性别比例
				D15.3. 掌握编程技能的年轻女性比例
				D15.4. 老年人的互联网使用比例
				D15.5. 低收入人群的互联网使用比例
	D16. 数据算法开放度	评估各国在人工智能数据、模型、算法上的开源与开放水平	UNESCO Rec. A75, A76	D16.1. 有影响力的开放 AI 模型与数据集数量
				D16.2. 对 AI 开发者社区的贡献度
	D17. AI 治理研究活跃度	评估各国在人工智能治理研究方面的活跃度	UNESCO Rec. A131, UNESCO RAM3.2.3, UNESCO RAM4.2.1	D17.1. AI 治理相关出版物数量&在 AI 相关出版物总量中的占比
				D17.2. AI 安全相关出版物数量&在 AI 相关出版物总量中的占比
				D17.3. AI 与可持续发展相关出版物数量&在 AI 相关出版物总量中的占比

表 6 AGILE 指数 2025 新增指标

评估方面	评估维度	新增指标
P1. AI 发展水平	D2. AI 基础设施	D2.3. 国家信息与通信技术发展水平评估表现
P2. AI 治理环境	D5. AI 治理准备度	D5.2. 国家数字治理整体水平评估表现
P3. AI 治理工具	D6. AI 战略规划	D6.4. AI 战略是否明确提及伦理影响
P4. AI 治理成效	D13. 公众 AI 认知度	D13.3. 公众对 AI 影响的认知程度
	D14. AI 社会接受度	D14.2. 公众对 AI 使用的积极或中性预期
		D14.3. 公众对 AI 应用的信任程度
	D15. AI 发展包容度	D15.2. 互联网使用的性别比例
		D15.3. 掌握编程技能的年轻女性比例
	D17. AI 治理研究活跃度	D17.2. AI 安全相关出版物数量&在 AI 相关出版物总量中的占比

P1. AI 发展水平

D1. AI 研发创新

人工智能研发创新是指世界各国在人工智能相关研发创新领域的活跃程度。根据联合国教科文组织《人工智能伦理问题建议书》第 83 条，“会员国应鼓励在人工智能领域开展国际合作与协作，以弥合地缘技术差距”，该建议与 D1 维度相契合，D1 维度涉及对人工智能发展水平的评估，从而为不同国家和地区之间的技术差距提供比较分析的依据。

D1 维度目前涵盖四个指标：

- **D1.1. AI 相关出版物数量&人均数量**
  - 数据来源：基于 DBLP 计算机科学文献数据库的统计分析（数据时间范围为 2024 年 4 月至 2025 年 3 月）
- **D1.2.AI 领域活跃研究者数量&人口占比**
  - 数据来源：基于 DBLP 计算机科学文献数据库的统计分析；研究对象为过去一年（2024 年 4 月至 2025 年 3 月）在 DBLP 上发表过人工智能相关论文的研究人员。
- **D1.3.AI 专利授权数量&人均数量**
  - 数据来源：基于世界知识产权组织（WIPO）[3] PATENTSCOPE 数据库的统计分析（数据时间范围为 2024 年 4 月至 2025 年 3 月）；WIPO 2024 年《生成式人工智能专利态势报告》（数据集）
- **D1.4.大型 AI 系统研发数量&与 GDP 比例**
  - 数据来源：Epoch（2025）—— 经 Our World in Data[4]进行主要处理（按国家/地区划分的大型人工智能系统累计数量；此处的“国家/地区”指“大型人工智能系统作者所属的主要机构所在地”；

3 https://www.wipo.int/portal/en/index.html

4 https://ourworldindata.org/grapher/cumulative-number-of-large-scale-ai-systems-by-country.

网络数据获取时间截至 2025 年 3 月）

D2. AI 基础设施

人工智能基础设施指的是人工智能的基础技术和数字生态系统。根据联合国教科文组织《人工智能伦理问题建议书》第 59 条，“会员国应促进数字生态系统的发展和获取，以便在国家层面以合乎伦理和包容各方的方式发展人工智能系统……此类生态系统尤其包括数字技术和基础设施……”；以及第 80 条，“会员国应通过国际组织，努力为人工智能促进发展提供国际合作平台，包括……基础设施，以及促进多利益攸关方之间的合作……”这些建议与 D2 维度是一致的。

D2 维度目前涵盖三个指标：

- **D2.1.托管数据中心数量&人均数量**
  - 数据来源：Data Center Map 数据库[5]（网络数据获取时间截至 2025 年 3 月）
- **D2.2.国家超级计算机总浮点运算能力&人均算力**
  - 数据来源：基于超级计算机 TOP500 榜单[6]的统计分析（2024 年 11 月，第 64 届 TOP500 榜单；按国家/地区统计的 Rpeak 与 Rmax 总和）
- **D2.3.国家信息与通信技术发展水平评估表现**
  - 数据来源：国际电信联盟（ITU）2024 年信息与通信技术发展指数（IDI）

D3. AI 产业活力

人工智能产业活力指的是一个国家在人工智能相关产业中的活跃程度。根据联合国教科文组织《人工智能伦理问题建议书》第 117 条，“会员国应支持政府、学术机构、职业教育与培训机构、产业界、工人组织和民间社会之间的合作协议，以弥合技能要求方面的差距，让培训计划和战略与未来工作的影响和包括中小企业在内的产业界的需求保持一致”，该建议与 D3 维度是一致的。

D3 维度目前涵盖两个指标：

- **D3.1.AI 领域私人投资&与 GDP 比例**
  - 数据来源：《2025 年人工智能指数报告》——斯坦福大学&QUID[7]（2024 年按地理区域划分的全球人工智能私人投资情况）
- **D3.2.新获融资的 AI 企业数量&与 GDP 比例**
  - 数据来源：《2025 年人工智能指数报告》——斯坦福大学&QUID（2024 年按地理区域划分的新获融资的 AI 企业数量）

5 https://www.datacentermap.com/datacenters/

6 https://www.top500.org/lists/top500/list/2024/11/

7 https://40006059.fs1.hubspotusercontent-na1.net/hubfs/40006059/Stanford_HAI_2024_AI-Index-Report.pdf

P2. AI 治理环境

D4. AI 风险暴露度

人工智能风险暴露度指的是世界各国面临的与人工智能相关的伦理、安全风险及问题的暴露程度。相关问题越多，该国对人工智能治理的紧迫性就越高。因此，这一维度会产生负面影响。根据联合国教科文组织《人工智能伦理问题建议书》第 50 条，“会员国应出台影响评估（例如伦理影响评估）框架，以确定和评估人工智能系统的惠益、关切和风险，并酌情出台预防、减轻和监测风险的措施以及其他保障机制”。该建议与 D4 维度是一致的。

D4 维度目前涵盖一个指标：

- **D4.1. AI 相关风险案例 / 事故数量&与 GDP 比例**
  - 数据来源：人工智能事件数据来自多个渠道，包括经合组织人工智能事件监测站(OECD AI Incidents Monitor, AIM)[8]，AI Incident Database (AIID)[9]，AI, Algorithmic, and Automation Incidents and Controversies Repository (AIAAIC)[10]，以及人工智能治理公共服务平台案例库(AI Governance Observatory, AIGO)[11]。（网络数据获取时间截至 2024 年 12 月）

D5. AI 治理准备度

人工智能治理准备度指的是国家在治理人工智能以及利用人工智能实现联合国可持续发展目标方面所具备的有利条件。根据联合国教科文组织《人工智能伦理问题建议书》第 54 条，“会员国应确保人工智能治理机制具备包容、透明、多学科、多边（包括跨界减轻损害和作出补救的可能性）和多利益攸关方等特性。特别是，治理应包括预测、有效保护、监测影响、执行和补救等方面”。该建议与 D5 维度是一致的。

D5 维度目前涵盖三个指标：

- **D5.1.国家治理整体水平评估表现**
  - 数据来源：世界治理指标[12]（WGI）——世界银行;数据获取日期为 2024 年 10 月 30 日；2022 年人类发展指数[13]（HDI）——联合国开发计划署
- **D5.2.国家数字治理整体水平评估表现**
  - 数据来源：2024 年全球网络安全指数（第 5 版）[14]——国际电信联盟（国家得分）；全球数据晴雨表第一版（各国在治理方面的表现如何[15]）；政府技术成熟度指数（GTMI）数据仪表板[16]—

8 https://oecd.ai/en/incidents

9 https://incidentdatabase.ai/

10 https://www.aiaaic.org/aiaaic-repository

11 https://www.ai-governance.online/ai-governance-observatory

12 https://www.worldbank.org/en/publication/worldwide-governance-indicators/interactive-data-access

13 https://hdr.undp.org/data-center/human-development-index#/indicies/HDI

14 https://www.itu.int/en/ITU-D/Cybersecurity/Pages/Global-Cybersecurity-Index.aspx

15 https://globaldatabarometer.org/module/governance/

16 https://www.worldbank.org/en/programs/govtech/gtmi

—世界银行（数据评估截至 2025 年 3 月）；2024 年电子政务发展指数；2024 年电子参与指数[17]——联合国电子政务知识库

- **D5.3.国家可持续发展整体水平评估表现**
  - 数据来源：《2024 年可持续发展报告》[18]——可持续发展目标转型中心

P3. AI 治理工具

D6. AI 战略规划

人工智能战略与规划指的是世界各国政府为人工智能的发展和应用制定的总体规划。联合国教科文组织《人工智能伦理问题建议书》第 56 条指出：“鼓励会员国……制定国家和地区人工智能战略……”；第 71 条则提到：“会员国应致力于制定数据治理战略……”该建议与 D7 维度所评估方向是一致的，即是否建立了人工智能相关的战略。

D6 维度目前涵盖四个指标：

- **D6.1.是否发布了国家层面的 AI 战略**
  - 数据来源：基于公开可得数据的案头调研（网络数据获取时间截至 2025 年 3 月）
- **D6.2.AI 战略是否包含实施计划**
  - 数据来源：基于公开可得数据的案头调研（网络数据获取时间截至 2025 年 3 月）
- **D6.3. AI 战略是否提及（劳动力）培训或技能提升**
  - 数据来源：基于公开可得数据的案头调研（网络数据获取时间截至 2025 年 3 月）
- **D6.4.AI 战略是否明确提及伦理影响**
  - 数据来源：基于公开可得数据的案头调研（网络数据获取时间截至 2025 年 3 月）

D7. AI 治理机构

人工智能治理机构是指各国政府为负责人工智能治理事务而设立的专门机构。根据联合国教科文组织《人工智能伦理问题建议书》第 58 条，各国应“…… 考虑增设独立的人工智能伦理干事岗位或某种其他机制，负责监督伦理影响评估、审计和持续监测工作，确保对于人工智能系统的伦理指导”。该建议与 D7 维度所评估方向是一致的，即是否建立了专门负责人工智能治理的机构。

D7 维度目前涵盖一个指标：

- **D7.1.是否已建立或指定了国家 AI 治理机构**
  - 数据来源：基于公开可得数据的案头调研（网络数据获取时间截至 2025 年 3 月）

17 https://publicadministration.un.org/egovkb/Data-Center

18 https://unstats.un.org/sdgs/report/2024/

D8. AI 原则规范

人工智能原则与规范是指各国政府为指导人工智能的开发、应用和治理而制定的原则和规范。根据联合国教科文组织《人工智能伦理问题建议书》，其中强调“……确保各国人工智能战略以伦理原则为指导”，且第 48 条指出，“主要行动是会员国出台有效措施，包括政策框架或机制等”。该建议与 D 维度是一致的，用于评估是否已确立指导人工智能发展的原则和规范。

D8 维度目前涵盖一个指标：

- **D8.1.是否发布了国家层面的 AI 原则或规范**
  - 数据来源：基于公开可得数据的案头调研（网络数据获取时间截至 2025 年 3 月）

D9. AI 影响评估

人工智能影响评估是指对人工智能系统潜在影响的评估，包括对个人、社会和环境产生的影响。根据联合国教科文组织《人工智能伦理问题建议书》第 50 条，各国应“出台影响评估（例如伦理影响评估）框架，以确定和评估人工智能系统的惠益、关切和风险”。该建议与 D9 维度是一致的，即是否已开发出评估人工智能影响的工具/框架。

D9 维度目前涵盖一个指标：

- **D9.1.是否出台了官方的 AI 影响评估机制/工具**
  - 数据来源：基于公开可得数据的案头调研（网络数据获取时间截至 2025 年 3 月）

D10. AI 标准认证

人工智能标准认证是指对人工智能系统进行评估，确保其符合相关伦理和安全标准，并颁发合规认证标识的机制。根据联合国教科文组织《人工智能伦理问题建议书》第 64 条，“会员国、国际组织和其他相关机构应制定国际标准，列出可衡量及可检测的安全和透明度等级，以便能够客观评估人工智能系统并确定合规水平”。该建议与 D10 是一致的，即是否已建立依据标准对人工智能系统进行评估的机制。

D10 维度目前涵盖一个指标：

- **D10.1.是否制定了国家层面的 AI 标准/认证机制**
  - 数据来源：基于公开可得数据的案头调研（网络数据获取时间截至 2025 年 3 月）

D11. AI 立法现状

人工智能立法是指各国为人工智能的发展和治理而制定的具有法律约束力的国家级文件。在立法维度，AGILE 重点关注三个关键领域：国家层面人工智能相关的通用法律法规、国家层面人工智能相关垂直领域法律法规，以及包含人工智能相关条款或修正案的数据保护法律法规。联合国教科文组织《人工智能伦理问题建议书》中“建议会员国在自愿基础上适用本建议书的各项规定，.....包括必要的立法或其他措施”，其第 133 条指出：“数据收集和处理工作应遵守国际法、关于数据保护和数据隐私的国家立法以及本建议书

概述的价值观和原则。” 该建议与 D11 维度是一致的，用于评估与人工智能相关的法律框架。

D11 维度目前涵盖三个指标：

- **D11.1.是否已制定或正在制定国家层面专门针对 AI 的综合性法律法规**
  - 数据来源：基于公开可得数据的案头调研（网络数据获取时间截至 2025 年 3 月）
- **D11.2.是否已制定国家层面专门针对 AI 垂直领域的法律法规**
  - 数据来源：基于公开可得数据的案头调研（网络数据获取时间截至 2025 年 3 月）
- **D11.3.是否已实施国家层面与 AI 相关的数据 / 信息保护法**
  - 数据来源：基于公开可得数据的案头调研（网络数据获取时间截至 2025 年 3 月）

D12.AI 治理国际参与

人工智能治理国际参与是指国家通过国际机制参与国际人工智能治理事务。根据联合国教科文组织《人工智能伦理问题建议书》第 80 条，各国应“通过国际组织，努力为人工智能促进发展提供国际合作平台”。该建议与 D12 维度所评估方向是一致的，即各国在人工智能治理领域的国际参与程度。

D12 维度目前涵盖两个指标：

- **D12.1.参与国际 AI 治理行动 / 机制的程度**
  - 数据来源：基于公开可得数据的案头调研，包括 2021 年《人工智能伦理问题建议书》、2023 年《布莱切利宣言》、2024 年《首尔部长声明》、2024 年《全球数字契约》、2025 年《关于发展包容、可持续的人工智能造福人类与地球的声明》、2019 年《经合组织人工智能原则》/《二十国集团人工智能原则》、2023 年《REAIM 行动呼吁》、2024 年《REAIM 行动蓝图》以及 2024 年国际人工智能安全研究所网络首次会议（数据截至 2025 年 3 月）
- **D12.2.参与 ISO AI 标准制定的程度**
  - 数据来源：基于 ISO/IEC JTC 1/SC 42（人工智能）参与成员名单；网络数据评估截至 2024 年 12 月

P4. AI 治理成效

D13.公众 AI 认知度

根据联合国教科文组织《人工智能伦理问题建议书》第 101 条，会员国应“会员国应与国际组织、教育机构、私营实体和非政府实体合作，在各个层面向所有国家的公众提供充分的人工智能素养教育，以增强人们的权能，减少因广泛采用人工智能系统而造成的数字鸿沟和数字获取方面的不平等”该建议与 D13 维度所评估方向是一致的，即是否有助于促进公众 AI 认知度。

D13 维度目前涵盖三个指标：

- **D13.1.公众 AI 素养的基础水平**
  - 数据来源：经合组织 2018 年及 2022 年国际学生评估项目（PISA）数学成绩；2024 年 Coursera

技能报告（Coursera 学习者数量）

- **D13.2.公众对 AI 议题的讨论程度**
  - 数据来源：基于谷歌趋势中“人工智能作为一个研究领域”的长期关注度数据。（数据截至 2025 年 3 月）
- **D13.3.公众对 AI 影响的认知程度**
  - 数据来源：益普索（IPSOS）2024 年人工智能监测报告[19] （包含问题：对于 “我对人工智能有清晰的理解；我知道哪些类型的产品和服务在使用人工智能”，您的同意或不同意程度如何）；经合组织数字化工具箱（GDT）[20] （成年人识别生成式人工智能制造的网络虚假信息的能力；网络数据评估时间为 2025 年 3 月）

D14.AI 社会接受度

根据联合国教科文组织《人工智能伦理问题建议书》第 39 条，“……公众监督，这可以减少腐败和歧视，还有助于发现和防止对人权产生的负面影响。透明度的目的是为相关对象提供适当的信息，以便他们理解和增进信任。” 该价值与 D13 维度相符，该维度用于评估是否开展了有关公众对人工智能态度的调查。

D14 维度目前涵盖四个指标：

- **D14.1.公众对 AI 发展的积极态度**
  - 数据来源：IPSOS 2024 年人工智能监测报告（对于“使用人工智能的产品和服务利大于弊；使用人工智能的产品和服务让我感到兴奋”，您的同意或不同意程度如何）；OECD GDT（认为人工智能将对自身生活产生积极影响的成年人比例；网络数据评估时间为 2025 年 3 月）。
- **D14.2.公众对 AI 使用的积极或中性预期**
  - 数据来源：IPSOS 2024 年人工智能监测报告（对于“未来 3-5 年，使用人工智能的产品和服务将深刻改变我的日常生活”，您的同意或不同意程度如何；您认为未来 3-5 年，人工智能使用的增加会让以下方面变得更好、更糟还是保持不变？——互联网上的虚假信息数量、我的娱乐选择、我完成事情所需的时间、我的健康状况、我的工作、就业市场、经济状况……）
- **D14.3.公众对 AI 应用的信任程度**
  - 数据来源：IPSOS 2024 年人工智能监测报告（对于“我相信使用人工智能的公司会保护我的个人数据；我相信人工智能不会对任何群体产生歧视或偏见”，您的同意或不同意程度如何）
- **D14.4.企业采用 AI 的意愿**
  - 数据来源：IBM 全球人工智能采用指数 2022[21] （全球人工智能采用率：已部署人工智能及正在探索人工智能的代表性公司数量）

19 https://www.ipsos.com/en-us/ipsos-ai-monitor-2024

20 https://goingdigital.oecd.org/indicator/81

21 https://www.ibm.com/downloads/documents/us-en/107a02e94a48f5c1

D15. AI 发展包容度

根据联合国教科文组织《人工智能伦理问题建议书》第 91 条，“会员国应鼓励女性创业、参与并介入人工智能系统生命周期的各个阶段”；第 105 条则指出，“会员国应提升女童和妇女、不同族裔和文化、残障人士、边缘化和弱势群体或处境脆弱群体、少数群体以及没能充分得益于数字包容的所有人在各级人工智能教育计划中的参与度和领导作用”。这些建议与 D15 维度所评估方向是一致的，即 AI 的发展是否对不同群体和性别的人们是否包容。

D15 目前涵盖五个指标：

- **D15.1. AI 领域活跃研究者的性别比例**
  - 数据来源：基于 DBLP 计算机科学文献数据库的统计分析；研究对象为过去一年（2024 年 4 月至 2025 年 3 月）在 DBLP 上发表过人工智能相关论文的研究人员。
- **D15.2.互联网使用的性别比例**
  - 数据来源：数字性别差距[22]—牛津大学（脸书性别差距：修复与移动指标；网络数据评估时间为 2024 年 10 月）； OECD GDT（男性与女性在互联网使用上的差异：男女性别差异；网络数据评估时间为 2025 年 3 月）
- **D15.3.掌握编程技能的年轻女性比例**
  - 数据来源：OECD GDT（在所有 16-24 岁会编程的人群中，女性所占比例；网络数据评估时间为 2025 年 3 月）
- **D15.4.老年人的互联网使用比例**
  - 数据来源：OECD GDT（55-74 岁的互联网用户；网络数据评估时间为 2025 年 3 月）
- **D15.5.低收入人群的互联网使用比例**
  - 数据来源：OECD GDT（低收入互联网用户；网络数据评估时间为 2025 年 3 月）

D16.数据算法开放度

根据联合国教科文组织《人工智能伦理问题建议书》第 75 条规定：“会员国应促进开放数据”；第 76 条规定：“会员国应推动和促进将优质和稳健的数据集用于训练、开发和使用人工智能系统，并在监督数据集的收集和使用方面保持警惕。”该建议与 D16 维度所评估方向是一致的，即数据、算法、模型是否对外开放。

D16 目前涵盖两个指标：

- **D16.1.有影响力的开放 AI 模型与数据集数量**
  - 数据来源：基于对 Hugging Face 平台排名前 1000 的热门数据集/模型的统计分析，并通过查看作者的 GitHub 用户资料来确定其所属国家（数据截至 2025 年 3 月）
- **D16.2.对 AI 开发者社区的贡献度**

22 https://www.digitalgendergaps.org/

- 数据来源：基于对 GitHub 上星级排名前 1000 的人工智能项目提交历史的统计分析（数据截至 2025 年 3 月）

D17.AI 治理研究活跃度

根据联合国教科文组织《人工智能伦理问题建议书》第 131 条规定：“会员国应根据本国具体国情、治理结构和宪法规定，采用定量和定性相结合的方法，以可信和透明的方式监测和评估与人工智能伦理问题有关的政策、计划和机制……（d）加强关于人工智能伦理政策的基于研究和证据的分析和报告；（e）收集和传播关于人工智能伦理政策的进展、创新、研究报告、科学出版物、数据和统计资料……” 该建议与 D17 维度所评估方向是一致的，即在 AI 治理主题方面的相关研究的量化分析。

D17 目前涵盖三个指标：

- **D17.1. AI 治理相关出版物数量&在 AI 相关出版物总量中的占比**
  - 数据来源：基于 DBLP 计算机科学文献数据库的统计分析（数据时间范围为 2024 年 4 月至 2025 年 3 月）
- **D17.2. AI 安全相关出版物数量&在 AI 相关出版物总量中的占比**
  - 数据来源：基于 DBLP 计算机科学文献数据库的统计分析（数据时间范围为 2024 年 4 月至 2025 年 3 月）
- **D17.3. AI 与可持续发展相关出版物数量&在 AI 相关出版物总量中的占比**
  - 数据来源：基于 DBLP 计算机科学文献数据库的统计分析（数据时间范围为 2024 年 4 月至 2025 年 3 月）

2. 数据收集与指数评估方法论

2.1 文献分析的数据收集方法

为了分析《计算机科学文献数据库》（DBLP）中的国籍和性别信息，我们采用了多步骤的方法。首先，利用出版物中提供的机构地址来确定作者的国籍。如果未提供地址，则通过作者的合作网络推断其国籍。为了识别作者的性别，我们使用了基于《世界性别姓名词典》（第二版）的 global_gender_predictor 包。在必要时，我们还会收集补充元数据，例如文章标题、摘要、作者姓名、出版日期、机构隶属关系、出版类别和链接等。为了确定一篇科学出版物是否与人工智能相关，我们应用了基于出版商和关键词数据的双重过滤方法。

首先，任何在人工智能期刊或会议上发表的文献都被归类为与人工智能相关。这些场所的名称和缩写是从 AMiner 人工智能期刊排名中获取的。其次，我们为各种人工智能子领域（例如，机器学习、神经网络、强化学习、贝叶斯模型、马尔可夫过程等）构建了一个全面的关键词列表。如果这些关键词出现在标题中，该出版物也被视为与人工智能相关。为了对与人工智能治理相关的出版物进行分类，我们开发了一个单独的关键词列表，其中包括“以人为本”“透明度”和“隐私”等术语。如果这些关键词出现在标题中，则认为该出版物与治理相关。

2.2 人工智能战略与规划（D6）的评分方法

对于指标 6.1“是否发布了国家层面的人工智能战略”，如果一个国家正式发布了国家级的人工智能战略或计划，则该指标得分为 100 分；如果尚未发布此类战略，则得分为 0 分。在这一维度中，我们对“国家级整体战略”的判断主要包括战略、方针、路线图和计划。值得注意的是，一些国家可能有多个版本的人工智能战略，但这并不会使指标 6.1 的得分更高。在对指标 6.2–6.4 进行评分时，我们将使用每个国家最新或最具代表性的战略作为评估目标，并在必要时考虑该国发布的其他人工智能战略。

指标 6.2“人工智能战略是否包含实施计划”得分为 0 分或 100 分。该指标要求人工智能战略在内容中（而非结构）包含可量化、可验证的目标指标，或者提出具有实际价值的具体、可操作的计划。如果一个国家的人工智能战略包含符合这些要求的目标或措施，则得分为 100 分；否则得分为 0 分。如果一个国家没有人工智能战略，则该指标得分为 0 分。我们认为，作为国家级整体规划，人工智能战略应该是清晰且可行的。如果人工智能战略仅提供模糊且宽泛的观点，则缺乏可执行性，也难以评估其目标是否已经实现。

指标 6.3“人工智能战略是否提及培训或技能提升”和指标 6.4“人工智能战略是否涵盖伦理内容”采用相似的评估方法：如果提及，则得分为 100 分；如果未提及，则得分为 0 分。如果一个国家没有人工智能战略，这两个指标都将获得 0 分。需要注意的是，指标 6.3 并不完全等同于人才或教育；它要求强调劳动力技能培训、公众意识提升以及相关学科招生计划的增强。指标 6.4 要求具体解决人工智能的伦理风险和价值追求。

2.3 人工智能立法现状（D11）的评分方法

不同国家的法律体系存在差异。例如，英国和美国等国家遵循普通法体系，主要以判例法为基础，而德国和法国等国家遵循大陆法体系，依赖成文法。在一些国家，行政命令或最高法院判决可能与立法机构通过的法律具有同等的法定效力。此外，不同国家在立法的各个阶段，如提案、审议、批准、签署、颁布、生效时间和执行时间等，具体程序和术语使用也存在差异。为了考虑到这些区别，在人工智能立法维度的评分中，我们承认所有具有法律约束力的法律文件，包括但不限于法律、法案或议案、法令、法典、条例、修正案、判令、行政令和判例等。本文一般将其简称为法律法规。

如果一个国家具有已实施的国家级的综合性人工智能法律法规，那么它将在指标 11.1“是否制定或正在制定专门针对人工智能的综合性国家法律或法规”中获得 100 分。如果一个国家只有正在制定中的人工智能法律和法规，那么它将获得 50 分。如果两者都不具备，该国将获得 0 分。

符合指标 11.1 要求法律和法规必须满足以下条件：首先，该法律法规应专门针对人工智能的治理或发展而设立。因此，除非在特殊情况下，一般的网络法或信息技术法等法律，虽然可能与人工智能有一定关联，或对人工智能有监管权力，但不包括在本类别中。其次，该法律或法规必须是全国性而非地区性的，因此，由省或州制定的法律不包括在内。最后，该法律或法规必须涉及人工智能的一般性内容，而不是专注于人工智能的某个具体子领域。关于这些子领域的立法将在指标 11.2 中进行评估。

需要注意的是，“正在制定中”要求至少有一个可供审查的草案或提案。一些国家（例如印度尼西亚）可能有立法计划，但尚无可查询的草案或提案，在这种情况下，该国将获得 0 分。如果一个国家的人工智能综合性法律已经获得批准但尚未正式实施（例如韩国），该国将获得 50 分。如果一个国家（例如美国）既实施了人工智能综合性法律，又有正在制定中的人工智能法律，该国将获得 100 分。如果一个国家遵守并实施了相应的区域性法律（如欧盟法律），则视为具有已实施的国家级的人工智能综合性法律，该国将获得 100 分。然而，如果区域性法律仍在制定中，相应的国家将不会获得 50 分，而是 0 分。

指标 11.2“是否制定了专门针对人工智能的国家层面垂直性法律或法规”主要考察各国在人工智能垂直领域的立法情况。人工智能在技术和应用的不同领域被认为是垂直领域，典型代表包括自动驾驶、生成式人工智能等。在指标 11.2 中，除了新颁布的法律外，我们还接受对现有法律增添具有针对性的条款、制定具有针对性的修正案或补充具有针对性的实施细则。这些法律和法规往往更注重细节，而不是广泛的综合性法律，其制定和实施相对模糊。因此，指标 11.2 采用更宽泛的区分标准，不再考虑“制定中”阶段，而是以“已颁布”作为评分标准。如果一个国家已经制定了针对人工智能垂直领域的国家法律和法规，该国将获得 100 分。如果没有这样的法律，该国将获得 0 分。

指标 11.3“是否实施了适用于人工智能的数据 / 信息保护相关国家法律”遵循与指标 11.2 类似的评分规则。我们承认专门制定的人工智能数据或信息法律，也认可对现有数据保护或个人信息法增加的具体条款或修正案。拥有被认可的人工智能数据或信息保护法律的国家将获得 100 分，那些遵守并实施相应区域性法律的国家也包括在内。没有这样法律的国家将获得 0 分。不具有针对人工智能的具体条款的一般数据

保护法不在本指标的范围内，但如果存在明确证据表明这些法律对人工智能相关的侵权案件具有管辖权，它们可以被视为人工智能特定的数据保护法律。

2.4 各层级的分数计算

在处理原始分数时，我们整合了多个来源的数据条目和统计结果，以实现三角验证，从而提高可靠性。这在数据稀缺且小的波动可能对分数产生重大影响时尤其有用，而使用多个数据来源可以减少偏差。不同数据元素之间的强相关性允许在缺失数据的情况下进行相互补充。在适当的情况下，我们考虑使用比例分数，以确保在不同基线统计数据（例如人口和国内生产总值）的国家之间进行公平比较。最后，我们使用分位适应标准化（见下文）对各种数据进行标准化并取平均值。在识别性别时，我们结合了平均水平推断，将 22.9%的未识别性别分配为女性，然后将识别出的比例进行分位适应标准化并取平均值。

在适当的情况下，我们汇总比例分数以确保在不同基准因素（如人口和国内生产总值）的国家之间进行公平比较。为了计算指标得分，我们将使用总数和比例的平均标准化得分。例如，如果一个国家在总数上获得了标准化得分为 5，在人均数量上获得了标准化得分为 3，那么该国在该指标上的得分将为 4。

在获得指标分数后，我们在每个维度内对得分进行了平均，并进行了标准化以获得维度分数。简单标准化被用于将均值调整为 50；由于基于调查指标和工具的得分分散，平均值在没有进一步标准化的情况下使用。然后，我们对维度得分进行了平均以获得评估方面得分，并对评估方面得分进行了平均以获得指数得分。在这里，维度 D4 人工智能风险暴露是评估方面 P2 的一个负面因素，因此使用[100 - 维度分数]进行平均。

2.5 分数标准化和数据插补

对于简单标准化，我们使用以下公式：

$$25 \times \frac{x_n - \mu}{\sigma} + 50$$

其中，μ 为所有国家的统计均值，σ 为统计标准差，x_n为某一国家的原始数据。标准化处理后，得分若超出 0 和 100，将被截断至 0-100 区间范围内。对于“百分位拟合”标准化方法，在完成每一次基础标准化之后，提取并剔除得分的一个百分位区间，再对剩余数据重复进行标准化和提取操作，直至获得四个四分位区间。由于原始数据存在显著聚集现象，且数值差异较大，因此需要对标准差进行调整，以更好地在较小量级上比较数据。

在指标存在缺失数据的情况下，插补顺序如下：第一，对于前一年有数据、但本年度缺失的指标，当前值通过以下方法估算：先计算该指标在所有国家中的平均增长率，再用该增长率乘以前一年该国的指标数据。第二，若无可参考的历史数据，则进行“逐级插补”。对于同一指标项下的缺失数据，首先根据可用数据计算该指标得分，再用计算结果插补缺失项，随后重新计算指标得分。若该指标得分本身缺失，则在维度层级采用相同方法：从可用的指标得分出发，计算维度得分，再用其插补指标得分，最终再行计算。若维度得分也缺失，则使用相同方法，以评估方面得分插补维度得分。

3. 其他相关指数工具

近年来，人工智能治理相关的多项国际指数报告持续更新，已成为各国政策制定、能力评估和参与全球对话的重要参考依据。为全面呈现全球人工智能治理相关评估工具的发展，本附录简要梳理了一些主要指数与报告的更新情况。

主要包括：由 Oxford Insights 发布的政府人工智能准备指数（Government AI Readiness Index，GRI）；由人工智能与数字政策中心（Center for AI and Digital Policy，CAIDP）发布的人工智能与民主价值观指数（Artificial Intelligence and Democratic Values Index，AIDVI）；由斯坦福大学以人为本人工智能研究院（Stanford University Human-Centered Artificial Intelligence，Stanford HAI）发布的全球 AI 活力工具（Global AI Vibrancy Tool，GAIVT）与 AI 指数报告（AI Index Report，HAI AI Index）；以及由 Tortoise Media 发布的全球人工智能指数（Global Artificial Intelligence Index，GAII）。

表 7 全球 AI 治理相关指数的发展状况

指数简称	基本描述	指标更新	数据来源更新	方法调整
GRI	评估各国政府有效应用人工智能的准备程度。	新增指标，替换了若干过时指标	包括责任人工智能在内的数据源更新	优化偏态数据处理方式，保留离群值国家
AIDVI	对各国人工智能政策在民主与伦理标准下进行评估。	纳入 Council of Europe AI Treaty 相关的指标	各国评分依据治理发展情况进行修订	评估重点转向法律—制度型治理
GAIVT	基于八大评估方面结构衡量国家 AI 生态系统活跃度。	延续原有指标体系	在以往版本基础上扩展数据更新	框架稳定，优化数据处理与操作方式
HAI AI Index	对全球人工智能发展趋势的年度综合评估。	新增 AI 硬件、推理成本、出版趋势等分析内容	包含企业采用责任人工智能实践的数据	“多样性”部分不再作为独立章节呈现
GAII	从创新、实施与投资三个维度跟踪国家 AI 发展水平。	增加半导体制造、AI 伦理、AI 初创企业收购等相关指标	提升数据时效性、全面性与来源多样性	调整权重分配，细化子评估方面定义

注：表中统一使用各报告的英文简称。

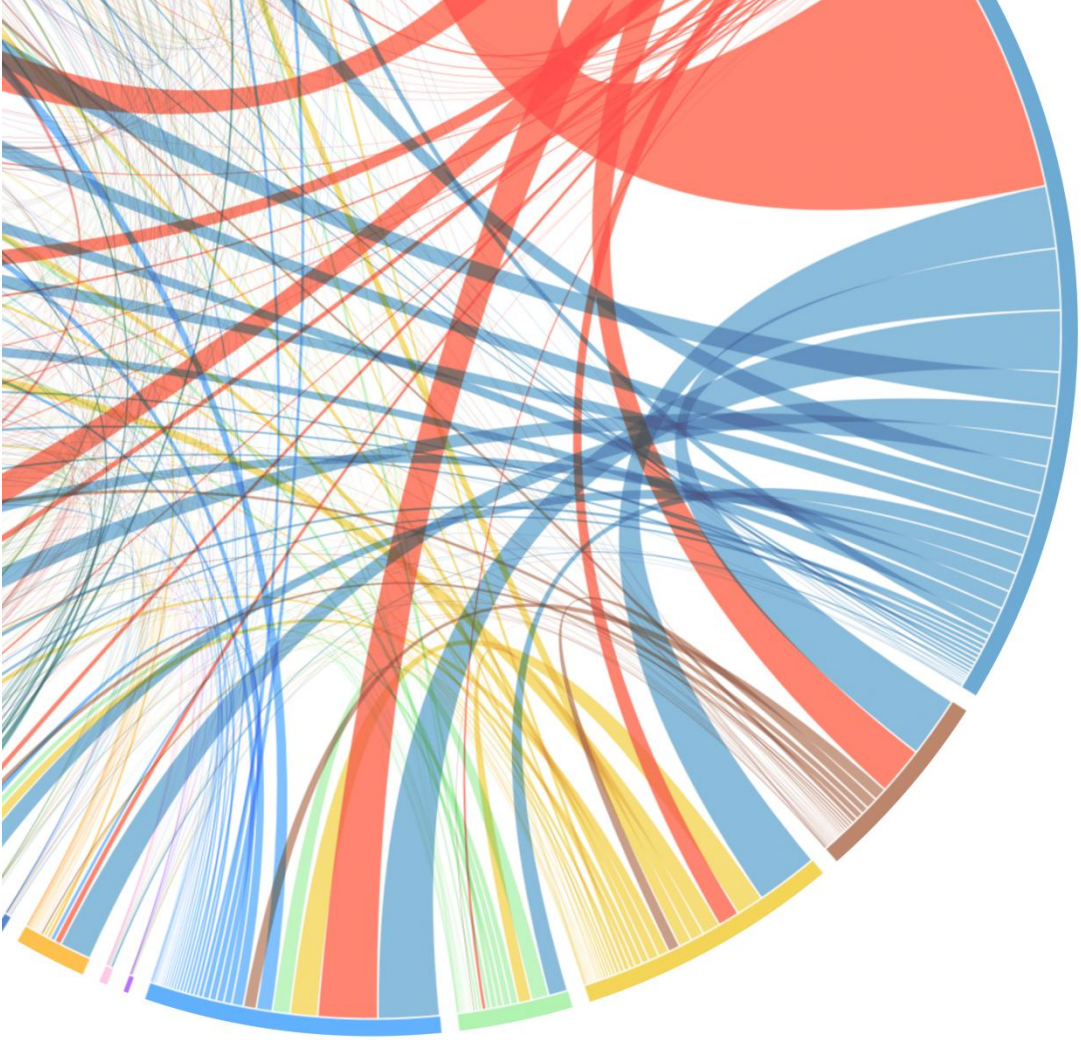
这些更新主要体现在指标设计、数据来源、方法框架和报告结构等方面，反映出相关机构持续在优化评估重点、提升工具适用性的努力。整体来看，人工智能治理指数的发展趋势是在总体结构保持稳定的同时，在治理重点、数据基础和评估维度等方面持续进行稳健调整。

- 指标体系扩展：保留原有结构框架，部分指数引入了新维度，涵盖政府执行能力、法律机制、环境影响及人工智能技术的实际部署等内容。
- 数据来源多元：多项更新替换了过时的数据集，采用国际数据库、行业调研及平台生成数据等方

面的新成果，注重时效性与情境相关性。

- 方法设计保持稳定：各指数在评估方面结构、评分模型和归一化处理等核心方法上基本保持一致，努力保持着跨年度的延续性与可比性。
- 国际治理参照：相关指数正逐步从静态排名工具演化为面向政策的测量工具，为国家治理能力改进提供参考，并助力国际合作框架的构建。

4. 图表索引



AGILE 指数 2025

网站: https://agile-index.ai/

邮箱: contact@long-term-ai.cn