

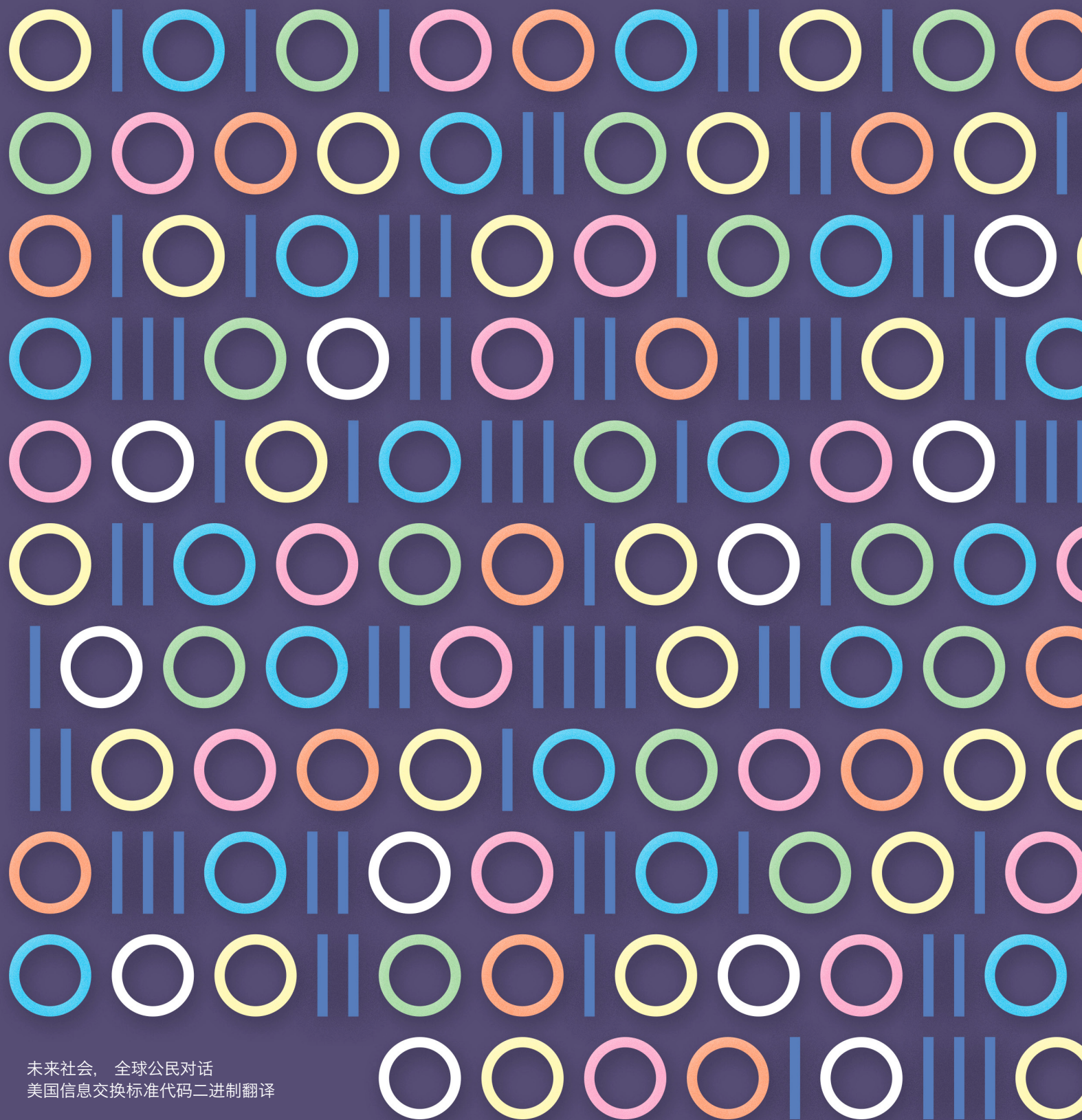
01

人工智能治理 的全球公民对话

THE
FUTURE
SOCIETY

行政总结

2018年9月



5 种语言：

英语，
法语，
中文，
日语，
俄语

参与：

70 390 页面浏览

2 525 学习时间

2 200 参与者

2 074 投票

1 291 书面提议

702 作者

17 个全球伙伴：

H5, Bluenove, IEEE, CNIL,
Les Echos, JSAI Ethics Committee,
Usbek & Rica, Cap Digital,
CNNum, World Future Society,
Tencent Research Institute,
The Millennium Project
Future of Life Institute, Centre
for the Study of Existential Risk,
CFI, Nesta, GCSP

21 则通讯：

让参与者随时了解最新的讨论进展

社交网络：

#AlforGood(人工智能向善) 和

#ImaginariesofAI(人工智能想象)

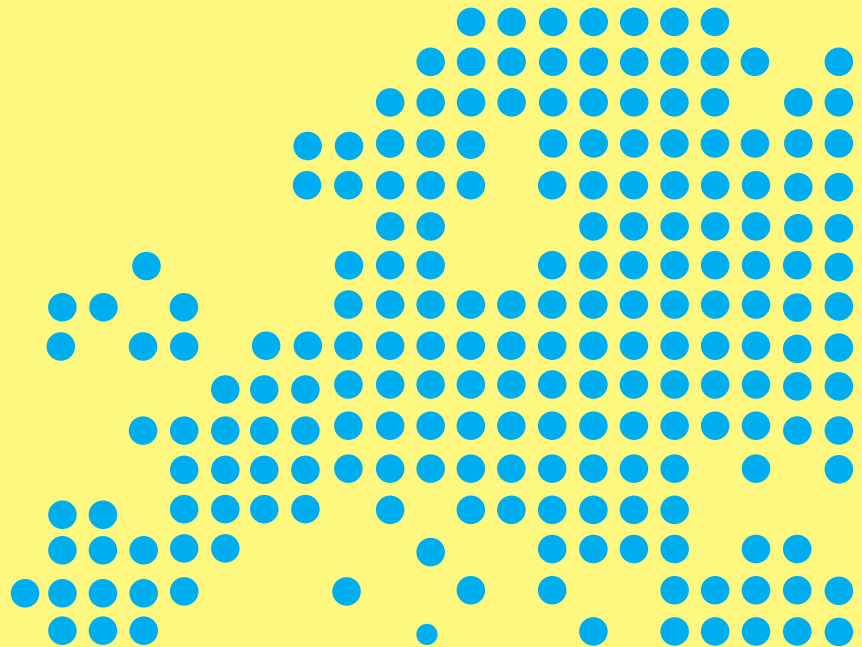
有超过**550**万关注

21 场线上 和线下活动

范围：

中国，
美国，
欧洲，
俄罗斯

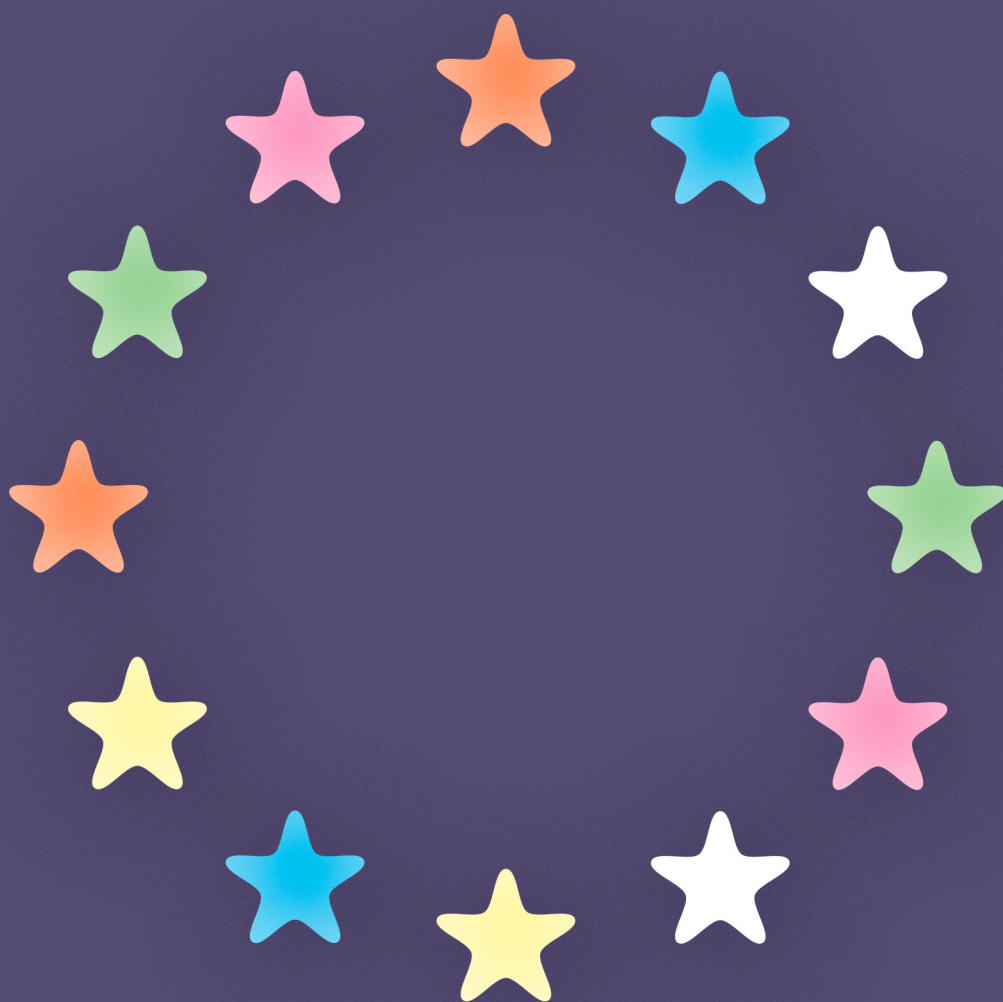




人工智能和机器学习已经不可逆转地嵌入在我们的生活中。从医疗到法律分析到金融贷款和虚拟个人助理，人工智能有各种通过扩展我们取得商品和服务的途径，给我们的社会创造利益的可能。它是人类发展历程中的重要一步。随着不断的创新，人工智能给我们带来的正面影响会越来越大。但是，和之前的科技革命一样，人工智能也给我们带来了许多重大的风险和威胁，包括个人隐私的瓦解，个人丧失能动性和控制力，互联网安全风险的增加，持续的社会不稳定，等等。

人工智能革命所带来的机会和挑战是紧密相关的。设计和应用智慧人工智能治理意味着在追求正面影响和减少负面影响之间取得一个平衡。智慧治理也需要区分哪些要根据具体的文化环境而定，哪些需要成为全球化的标准。这在瞬息万变的全球背景下，在各个国家，民族，和经济体日益相互依赖和联通，但在文化和价值上仍相去甚远的环境下，更是具有挑战。比如说，发展中的经济体更加关注当下的发展，现代化，人口问题，和各种差距；然而发达国家希望通过重塑和创新他们的社会经济体系来保持它们现有的权力。在这个大环境中，协调各种价值体系将会非常困难。

人工智能将如何发展并不是既定的命运或者随机发展。各个公民，政府，商业领袖，和学者都有能力决定发展的方向。这也是为何我们要组织“全球公民论坛”这个为期7个月的跨学科的社会讨论。我们的目的是从多个角度对人工智能和它的治理进行分析。通过收集超过2000个参与者，近700位作者，用5种语言所书写的超过3300想法，我们了解了全球各地公民对人工智能影响的见解。我们汇总了这些见解，并以此为依据提出了一系列对人工智能未来发展具有积极意义的政策推荐。



对人工智能的普世概念是不存在的。

1

这样的理解不能代表人工智能的现实，或者内部动态。对人工智能缺乏普世理解是一个历史的产物。一直以来人工智能并没有一个通用的定义，它总是充满了模糊性。它的概念总是宽泛的，流动的，具有争议的。人工智能指代了一系列的科学技术学科，科技，和方法。在众多领域的实践中，对人工智能的定义都比较随意，这也导致了进一步的概念模糊性。随着人工智能技术的发展，人们所牢记的却是电影中对机器智慧科幻化的描绘，比如它们对常识，意识和智能的复制。为了更好地治理，我们必须重新对人工智能进行定义，用对人工智能组成部分和发展动力的正确理解来纠正它概念中的不准确，不统一的方面。我们对它现行的定义是：“由大数据主导，以机器学习算法为中心，由超级计算能力支持的复杂的社会科技体系。”这个定义可以用来衡量目前人工智能内部动态和治理情况。

对人工智能在期待在不同社会之间以及社会内部都会产生分化，但建立对人工智能的信任是大家共同的诉求。

2

虽然我们并没有为人工智能普世观念奠基的道德，伦理和社会习俗，但是建立对人工智能科技的信任对我们的论坛成员来说都很有价值。如果通过在论坛和活动不断地和公众交流，影响科技标准和认证，把更多资源放在增加解释性上，公众对人工智能会越来越建立起信任感。这是一个需要跨越的重要障碍，因为如果要把人工智能的利益最大化，我们必须要让公众对这个科技建立起信任感，从而在日常生活中与它接触互动。

区块链和其它自我治理标准有解决人工智能治理问题的潜力。

3

区块链和分散账本技术可以被运用以解决人工智能的一些重要挑战，比如说保护个人数据安全，让更多开发商能够取得数据，治理的伦理系统。区块链和其他自我治理标准运用分散的，去中心化的数字账本来记录，存储，密锁让多方交易更安全。这些基础设施会促进人工智能在信息安全和支持数据交流方面的应用。这也会激励更多的公民贡献他们的数据，让更多开发商有机会获取数据，增加数据来源的多样性，减少数据的偏向性。以区块链为基础的激励机制和智慧合同也可以为我们提供其他形式的人工智能治理模式。这些潜在的应用目前还处于初步阶段。很多领域我们还未探索，或许存在很多技术难度。但是它们为我们解决人工智能治理的一些重大挑战提出了理论上的可行框架。

5

人工智能预计将会加快社会进步的步伐。

4

人工智能拥有加速和支持新的科学发现的潜力，有产生大规模效应的能力，比如监控，促进以及达成联合国可持续发展目标。因此，人工智能可能会引来一个社会转折，转变现在趋向于不平等和排他的发展。如果我们采取健全的全球合作手段使用人工智能，这个科技可以帮助我们解决一些重大的挑战，比如说气候变化和贫困问题。它可以帮助我们寻找新的，更有效的资源配置方式，以此促进经济发展。如果我们采取恰当的平衡和制约手段来治理人工智能。人类社会将会更繁荣地发展。

智慧治理是把握人工智能风险的关键，我们必须灵活和与时俱进。

5

从数据隐私的风险，到大规模的劳动力被取代，我们的讨论探索了很多随着人工智能越来越普及而带来的威胁与风险。讨论的参与者都明确地表示，治理在风险管理和确立价值观中起到了至关重要的作用，比如说我们要保证“包容，平等，代表性，和透明化”等价值得以保存。治理也能避免让任何一个行业或者国家“探索伦理底线”以取得市场份额，从而取得先行者机遇。因为科技发展无法比拟的节奏，治理必须要在不同程度上进步，比如说全球和本地层面的融合，政治与科技目标的一致化。智慧治理要赋能政府，让它们不断加强在这些复杂问题方面的知识。

全球人工智能群体都在为失去人身安全，失去控制，以及人工智能系统的重大失误感到担忧。

6

随着人类社会越来越依赖人工智能做一些关键的决定，不管是个人决定（像是安排时间）还是超越国界的决定（像是运用自动化武器），大家对人类失去控制权而可能产生的后果感到深深的担忧。这和对人工智能决策过程的不信任，成为了公众接受和使用人工智能的重大绊脚石。

人工智能让人类重新审视他们的身份和在社会中的位置。

7

随着人类社会越来越依赖人工智能做一些关键的决定，不管是个人决定（像是安排时间）还是超越国界的决定（像是运用自动化武器），大家对人类失去控制权而可能产生的后果感到深深的担忧。这和对人工智能决策过程的不信任，成为了公众接受和使用人工智能的重大绊脚石。

6

The Future Society
人工智能治理的全球公民对话

.0. 00000 88 00000 000 080
888. `888' .8' `888' `8' `””
”888. 888 .8' 888 8 000. .00. 0000 0000 000 .00000. 0000 d8b .0000.0 .000
`888. 888 .8' 888 8 `888P”Y88b `888 `88. .8' d88' `88b `888””8P d88(“8 `P
008888. 888 .8' 888 8 888 888 888 `88..8' 888000888 888 `”Y88b. .oP
`888. 888 .8' `88. .8' 888 888 888 888 `888' 888 .o 888 o.)88b d8(
o8888o o888o 88 `YbodP' o888o o888o o888o `8' `Y8bod8P' d888b 8””888P' `Y88

000. .o8 0000 . 88 0.000000..0
`Y8b “888 `888 .o8 d8P' `Y8
888 8880000. .0000.0 .00000. 888 .00000. .o88800 .00000. Y88bo. .00000
888 d88' `88b d88( “8 d88' `88b 888 d88' `88b 888 d88' `88b `”Y8888o. d88' `88b
888 888 888 `”Y88b. 888 888 888 888000888 888 888000888 `”Y88b 888000888
d88' 888 888 o.)88b 888 888 888 888 .o 888 . 888 .o oo .d8P 888 .o
od8P' `Y8bod8P' 8””888P' `Y8bod8P' o888o `Y8bod8P' “888” `Y8bod8P' 88 8””88888P' `Y8bod8P'

000. .o8 . 88 0.000000..0
`Y88. “888 `888 .o8 d8P' `Y8
.d88' .00000. 0000 d8b .0000.0 .00000. 000. .00. .0000. 888 .0000888 .0000. .o88800 .0
o88P' d88' `88b `888””8P d88(“8 d88' `88b `888P”Y88b `P )88b 888 d88' `888 `P)88b 888 888
888000888 888 `”Y88b. 888 888 888 888 .oP”888 888 888 888 .oP”888 888 888
888 .o 888 o.)88b 888 888 888 d8(888 888 888 888 888 d8(888 888 . d8(
`Y8bod8P' d888b 8””888P' `Y8bod8P' o888o o888o `Y888””8o o888o `Y8bod88P” `Y888””8o “888” `Y8

0000. 0000 0000 0000 080 88
`Y8b `888 `888 .00000. .00000. 888 0000 .00000. 888 .00. .0000. 0000 000. .00. 88
888 888 d88' `88b d88' `”Y8 888 .8P' d88' `”Y8 888P”Y88b `P)88b `888 `888P”Y88b .8'
`88b 888 888 888 888 888888. 888 888 888 .oP”888 888 888 888 .8'
.88P 888 888 888 888 .o8 888 `88b. 888 .o8 888 888 d8(888 888 888 888 .8'
od8P' o888o `Y8bod8P' `Y8bod8P' o888o o888o `Y8bod8P' o888o o888o `Y888””8o o888o o888o o888o 88

000. 88 0.000000..0
`Y88. .8' d8P' `Y8
.d88' 0000 d8b .00000. .00000000 0000 d8b .00000. .0000.0 .0000.0 .8'
o88P' `888””8P d88' `88b 888' `88b `888””8P d88' `88b d88(“8 d88(“8 Y88bo. 000. .
888 888 888 888 888 888 888000888 `”Y88b. `”Y88b. `”Y8888o. `888P”
888 888 888 `88bod8P' 888 888 888 .o o.)88b o.)88b .8'
d888b `Y8bod8P' `8000000. d888b `Y8bod8P' 8””888P' 8””888P' 88
d” YD
“Y88888P'

000. 88 0.000000..0
`Y8b .8' d8P' `Y8
.00000 888 888 `88..8' 888000888 888 888 888 .oP”888 888 888 888 .00000. .00000. .8'
.88' 888 888 `888' 888 .o 888 888 888 d8(888 888 888 888 .o8 888 .o 7 .8
od8P' `Y8bod8P' `8' `Y8bod8P' d888b o888o o888o `Y888””8o o888o o888o `Y8bod8P' `Y8bod8P' 88

.0. 080 0000 .o8
888. `”” `888 “888
”888. .00000000 0000 888 .00000. .0000. 000. .00. .0000888 888 00000000000 `888' `8
`888. 888' `88b `888 888 d88' `88b `P )88b `888P”Y88b d88' `888 88800008 `88. .8'
008888. 888 888 888 888 888 888000888 .oP”888 888 888 888 888 888 888 “ `888' `888
`888. `88bod8P' 888 888 888 d8(888 888 888 888 888 888 888 o `888' `888
o8888o `8000000. o888o o888o `Y8bod8P' `Y888””8o o888o o888o `Y8bod88P” o88800000d8

.o8 080 . 88 00000 00000
“888 .o8 `888' `888' 888 888 0000 0000
.0000888 .00000. 000. .00. .o88800 0000 .o88800 0000 000 888 888 8880000888 888 888 888
d88' `888 d88' `88b `888P”Y88b 888 888 888 `88. .8' 888 888 888 888 888 888 888 888 888 888
888 888 888000888 888 888 888 888 `88..8' 888 888 888 888 888 888 888 888 888 888
888 888 888 .o 888 888 888 . 888 888 . `888' 888 888 888 888 888 888 888 888 888
`Y8bod88P” `Y8bod8P' o888o o888o “888” o888o “888” .8' 88 o888o o888o `Y88V”V8P'

重要建议

为了解决以上提出的一系列问题，
参与的讨论成员提出了以下一些纲领性建议。

将“决定智能操纵”的能力融入到治理的过程中，以取得政策制定过程中对社会目标的平衡。

1

决定智能可以建立我们对行为如何导致结果的不断迭代的理解。操作的概念是理解谁有掌控某一具体问题的能力。在人工智能领域，决定智能操纵可以让合适的人才聚合在一起解决问题，从而达到预期的效果。解构人工智能科技的复杂性然后理解它的影响，这个重复操作的机制是有效治理的关键。我们的决定智能能力可以在以下方面提高：

- 一个跨政府的人工智能委员会(IPAI)可以成为政策制定的基石。一个庞大的跨学科人工智能科学家和专家群体，一个不断在科学过程中做出评估的群体，在统一人工智能全球意见中，将会起到关键的作用。
- 一个“人工智能为可持续发展中心”，它可以运用人工智能的力量去监督，刺激，预测，和促进联合国可持续发展目标的达成。作为一个集合商界，政府，学界和公民社会的行动智库，这个中心将会成为把IPAI的政策建议运动到现实世界中的重要引擎。它可以把治理和伦理嵌入到人工智能的解决方案中，并且追求包容的，多样的人工智能发展。
- 把“美国中国相对论”变成“美国中国相联通”，从而让人工智能商业模式，科技标准，和内在框架和观念更一致。美国和中国正在成为“人工智能超级大国，成为领先全球的两强。我们需要探索竞争和合作中的动态平衡，这非常重要。

传播有效的人工智能治理方式。

2

- 在政策制定和监督过程中，融入具有多个利益相关群体，跨学科的，集体智能收集过程。比如说我们可以把艺术家结合到治理问题的讨论中来，他们可以提供很有价值的建议，也对公民有很大的影响力。
- 倾向于支持正在发展的国际科技标准制定过程和相关组织，这可以让人工智能利用到“软性治理”的力量。
- 与已经成型的组织和社区合作，建立普世的，可以实践的，但也可以适应各种文化的伦理，规范和操作标准。
- 探索是否需要“硬性治理”，例如国际条约，操作和监督机制等。这可以包括治理数据的收集，存储，处理，和流动，也可以包括治理无人机的试验使用，自动化杀伤性武器的使用。

保证人工智能拥有安全， 有代表性的数据。

3

- 建立一个独立的，市场主导的，规范的全球数据交流平台。
- 运用区块链科技，从源头保证数据安全。
- 教育公民人工智能会促使数据更具有代表性，包容性，和公民参与度。
- 建立一个人工智能可以探测和改正其自身的偏见和单一程度的机制。

降低失业的风险。

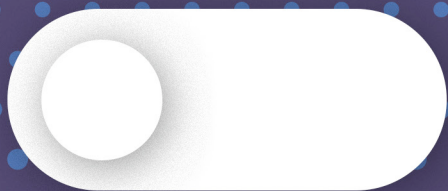
4

- 用具体的资金流动模型来估计有没有可能在不同国家和不同时间重新建立社会保障体系。
- 让受雇者，工会，雇佣者和政府建议合作机制，投资发展终生培训项目。
- 鼓励社区工作运动，以此作为工作的补充，并借此促使大家重新审视“工作”的定义。
- 发展全国性的教育项目，以此提高人们在人工智能革命对工作造成不确定影响环境中的适应性和学习能力。

全球公民论坛以后

全球公民论坛开始了一场关于人工智能兴起，发展，和影响的话。虽然这个报告提供了许多见解和政策建议，我们意识到这只是一个开始。未来社会将会继续在以下方面作出努力：

- 通过研究，指导，会议，教育项目，和“特殊项目”等与政府和政策制定者紧密合作，来保证我们共同体的想法可以转化为可操作的政策和影响。
- 继续扶植这个讨论，通过草根运动让这个共同体更加多元和包容。我们计划通过视频内容把更多讨论中得出的见解传播给世界各地的公民，从而把让更多人知道这个讨论，让关于人工智能的讨论更包容。



- 利用人工智能想象中的创造性。把它作为一个工具, 用来与政府, 商业领袖和公民交流合作。我们希望与剧本创作人, 制作人和其他娱乐行业的人员合作, 了解如何利用动态的想象影响公众的认知, 从而影响治理的过程。我们希望能创造更准确以及平衡的人工智能叙事, 从而大规模地影响观众群体。
- 继续利用集体智能平台来支持政府, 学术界, 商界和公民社会之间关于人工智能的对话。受到全球公民论坛“人工智能的想象”这个主题的启发, 我们正在为更大的组织 (包括政府和商界) 开发方法和策略, 从而让他们更好地了解和利用人工智能, 在他们的实施有效的人工智能手段, 以及促使变革。
- 通过与全球人工智能有关的研究, 会议和教育影响决定智能。我们将全力以赴, 用动态的, 迭代的方式来更好地理解人工智能的影响和治理。我们将特别拥护一个跨政府的人工智能委员会(IPAI) 和试验人工智能为可持续发展中心的想法。这些努力会基于全球人工智能治理迪拜圆桌会议的框架 (这是一个我们与阿联酋政府合作的项目)。这个全桌会议为我们提供了一个不断演进的共同体和人工智能治理各个方面的意见领袖。我们会继续扶植这些联系, 因为我们希望创造长期的理解 and 政策影响。
- 与人工智能的本质相同, 我们相信我们对人工智能革命影响的思索也会不断演进。我们把公民论坛看做是一个坚固的基石。它会支持我们继续建造一个对社会各方都有益的未来。