国际安全背景下合成数据的治理影响

技术与安全研讨会报告

Federico Mantellassi

致谢 2
关于裁研所 2
关于安全与技术计划 2
注 2
引文 3
关于作者 3
缩略语 3
1. 导言 4
1.1 活动简介 5
2. 合成数据与国际安全：确定问题的框架 6
2.1 什么是合成数据 6
2.2 军事领域的合成数据 7
3. 治理方面的挑战和影响 8
3.1 合成数据和民用数据治理 9
3.2 标准的作用 10
3.3 合成数据与军事人工智能的国际管理 11
a. 挑战的新颖性和现有框架的适用性 11
b.多利益攸关方方法的重要性 12
c.关于指导方针和具体情况 13
d.关于军事领域的治理机会 13
4. 结论 14
附件：活动议程和与会者 15

致谢

裁研所核心资助者的支持为裁研所的所有活动奠定了基础。本出版物由欧盟资助，是裁研所安全与技术计划的一部分，该计划得到了捷克、法国、德国、意大利、荷兰、挪威和瑞士政府以及微软公司的支持。作者衷心感谢何雯婷主持和组织了本次活动的第一场小组讨论，并感谢 Jessica Espinosa Azcarraga 协助组织本次活动。此外，作者还要感谢所有发言人的参与，以及 Giacomo Persi Paoli 博士、Sarah Grand Clément、何雯婷、Calum Inverarity、Ana Beduschi 博士和 Aldo Lamberti 对本报告提出的宝贵意见。

关于裁研所

联合国裁军研究所（裁研所）是联合国内部一个由自愿捐款供资的自主机构。裁研所是世界上为数不多专注于裁军的政策研究所之一，提供裁军与安全方面的知识，并促进这方面的对话和行动。裁研所总部设在日内瓦，协助国际社会提出切实可行的创新想法，以找到解决重大安全问题的办法。

关于安全与技术计划

当代科学技术的发展为国际安全与裁军带来了新的机遇和挑战。裁研所的安全与技术计划（SecTec）力求建立对特定技术创新的国际安全影响和风险的了解和认识。该计划召集利益攸关方探讨各种想法，并就如何应对这些想法形成新思维。

注

本出版物所使用的名称和材料的编排方式并不意味着联合国秘书处对任何国家、领土、城市或地区或其当局的法律地位，或对其边界或界线的划分表示任何意见。出版物中表达的观点仅由作者本人负责，不一定反映联合国、裁研所、其工作人员或赞助者的观点或意见。

引文

Mantellassi, Federico.”国际安全背景下合成数据的治理影响：技术与安全研讨会报告》。瑞士日内瓦：联合国裁军研究所，2024 年。

关于作者

Federico Mantellassi 是联合国裁军研究所安全与技术项目的研究员。他的工作重点是新兴科学、技术发展和创新对国际安全的影响、风险和机遇。此前，Federico 是日内瓦安全政策中心的研究和项目官员，研究新兴技术、国际安全和战争之间的交叉问题。他拥有伦敦国王学院情报与国际安全硕士学位和莱顿大学国际研究学士学位。

缩写与缩略语

AI 人工智能

EU 欧洲联盟

GDPR 一般数据保护条例

IEEE 电气与电子工程师学会

ISO 国际标准化组织

ODI 开放数据研究所

PET 隐私增强技术

1. 导言

数据对于人工智能（AI）系统的训练和开发至关重要。然而，与数据相关的三个关键问题可能成为开发和部署人工智能能力和系统的障碍。首先，人工智能技术的发展——至少在一定程度上——依赖于用于训练人工智能模型的大型数据集的可用性。其次，数据是一种资源，其可用性、收集、清理、使用和共享受到各种因素的影响，如收集成本、某些领域缺乏真实世界的数据，以及监管、法律和道德限制。第三，数据质量、代表性和多样性与人工智能模型的性能、偏差程度、准确性和可靠性直接相关。**合成数据——在数字世界中人工生成的数据，其属性通常来源于原始数据集——已被提出作为解决其中一些数据相关问题的解决方案，尤其是在人工智能模型训练方面。**[1] 事实上，合成数据有助于解决数据集的偏差等问题，同时还能实现数据集的扩展、创建、多样化和微调。合成数据也常被称为隐私增强技术（PET），可促进敏感数据集的使用和共享。[2]在军事等领域，合成数据尤其大有可为。[3]在这一敏感领域，对人工智能能力的需求与日俱增，但高质量、多样化的数据集却供不应求，而错误算法可能会造成严重后果。合成数据可以帮助开发先进的人工智能能力，而不需要大量的真实世界数据。[4]然而，合成数据并不是万能的。事实证明，它可能会加剧许多它试图遏制的问题，从而引发治理和监管方面的讨论。[5]

---

[1]郝爽等，2024 年。《人工智能中的合成数据：挑战、应用与伦理意义》。华中科技大学软件工程学院。 https://arxiv.org/pdf/2401.01629 ; Lee, Peter.2024.《合成数据与人工智能的未来》。康奈尔法律评论》。 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4722162 ; Deng, Harry.2023.《探索人工智能和自主系统的合成数据：入门》联合国裁军研究所。联合国裁军研究所。 https://unidir.org/wp-content/uploads/2023/11/UNIDIR_Exploring_Synthetic_Data_for_Artificial_Intelligence_and_Autonomous_Systems_A_Primer.pdf 。

[2]Naughton、Mitchell 等人，2023 年。《合成数据作为解决体育科学中数据隐私和保密问题的策略：实例和 R Shiny 应用》。《国际运动生理学和表现杂志》。第 18 卷（10）：1213-1218. doi: 10.1123/ijspp.2023-0007; Syntheticus 公司。《合成数据 101： 它是什么，如何运作以及用途》 Syntheticus 公司.Web. n.d. https://syntheticus.ai/guide-everything-you-need-to-know-about-synthetic-data#chapter-8 。

[3]Chahal, Husanjot et al.《比石油更混乱：评估军事人工智能的数据优势》。安全与新兴技术中心。 https://cset.georgetown.edu/wp-content/uploads/Messier-than-Oil-Brief-1.pdf 。

[4]同上。

[5]Deng, Harry.2023.《探索人工智能和自主系统的合成数据：入门》联合国裁军研究所。 https://unidir.org/wp-

合成数据在监管和治理方面处于相对 “灰色地带”。主要的数据治理和人工智能监管框架，如欧盟的《人工智能法》和《通用数据保护条例》（GDPR），即使提到合成数据，也只是一带而过。在一些人看来，这意味着合成数据作为一种隐私增强技术，可以绕过严格的监管框架，或成为一种有用的合规工具。[6]另一些人则指出，合成数据与真实世界的数据存在许多相同的风险，并可能对人工智能模型的准确性、安全性、公平性和代表性产生类似的下游影响，因此他们坚持认为有必要制定新的监管框架和方法，以避免出现治理漏洞和盲点。[7]有鉴于此，最重要的是要了解当前的治理（民用和军用）和监管框架是如何涵盖合成数据的，它们是否适合解决潜在风险，以及是否需要调整。在军事领域快速采用人工智能能力的背景下，监管和治理差距尤其重要。因此，了解合成数据对新出现的军事人工智能治理讨论的影响至关重要。

为探讨国际安全背景下合成数据的治理挑战，裁研所安全与技术计划举办了题为 “合成数据技术与安全研讨会：探讨治理影响”的活动。

本报告概述了活动讨论的关键主题和要点。报告分为两部分，反映了活动的组成结构。第一部分简要概述了该技术及其在军事领域的应用。第二部分介绍了合成数据在国际安全背景下对治理提出的各种观点、问题和潜在挑战。

1.1 活动简介

技术与安全研讨会包括裁研所安全和技术计划组织的一系列活动，重点是各种使能技术。该系列活动的主要目标有三个方面：

1. 使外交界接触到各种新兴的关键使能技术；
2. 提醒外交界注意开发和使用此类技术可能对国际安全产生的影响；
3. 通过多个利益攸关方的对话和参与，探索治理的可能性。

---

content/uploads/2023/11/UNIDIR_Exploring_Synthetic_Data_for_Artificial_Intelligence_and_Autonomous_Systems_A_Primer.pdf 。

[6]Zojer, Alexander. 《合成数据：欧盟人工智能法案下人工智能合规的关键工具》。Mostly.AI.30 October 2023. https://mostly.ai/blog/ai-compliance-with-eu-ai-act-using-synthetic-data 。

[7]Gal、Michal、Lynskey、Orla.2023.《合成数据：数据生成革命的法律意义》。《爱荷华法律评论》.109. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4414385 ; Whitney, Cedric Deslandes, Norman, Justin.2024.《假数据的真实风险：合成数据、多样性清洗和同意规避》。 https://doi.org/10.1145/3630106.3659002 。

2024 年 10 月 29 日，以合成数据治理为主题的技术与安全研讨会如期举办。这次为期半天的活动包括**技术早餐会（为决策者介绍技术）**和**合成数据多利益攸关方对话**，来自行业、国际组织和学术界的专家在**对话**中就国际安全背景下的具体治理挑战分享了各种观点。这次活动在 2024 年联合国大会第一委员会第七十九届会议期间以虚拟方式举行。

活动日程安排见本报告附件。

2. 合成数据与国际安全：确定问题的框架

要点
• 生成式人工智能领域的进步和人工智能在各行各业的应用日益广泛，扩大了合成数据的普及范围，提高了生成数据的规模和便捷程度以及数据的种类和质量。 • 合成数据为解决民用和军用领域与数据有关的挑战（偏差、稀缺性、质量、代表性、隐私）带来了希望和潜在的解决方案。 • 武装部队越来越多地采用人工智能功能来训练军事人工智能模型，用于识别、瞄准系统、作战和战术规划，以及开发场景和合成环境。在此背景下，武装部队越来越多地转向合成数据。 • 尽管合成数据有很多好处，但它可能会延续现有的数据相关风险，产生新的风险，或扩大其影响范围。

2.1 什么是合成数据[8]

合成数据可定义为 **“由计算机模拟或算法创建的信息，可再现真实世界数据的某些结构和统计特性”。**[9]合成数据有多种生成方法，生成的数据集既可以是完全合成的（所有数据都是人工生成的），也可以是部分合成的（用合成数据替换真实数据集的一小部分），还可以是混合的（真实世界数据与完全合成数据的混合）。[10]简而言之，合成数据主要用

[8]本节借鉴了裁研所安全与技术计划以前在合成数据与国际安全方面开展的工作。有关合成数据的详细、深入探讨，以及与合成数据相关的国际安全风险和机遇，特别是在人工智能和自主军事能力方面，请参见 https://unidir.org/wp-content/uploads/2023/11/UNIDIR_Exploring_Synthetic_Data_for_Artificial_Intelligence_and_Autonomous_Systems_A_Primer.pdf 。

[9]De Wilde、Philippe 等人，2024 年。《关于使用合成数据训练人工智能模型的建议》。联合国大学。 https://collections.unu.edu/eserv/UNU:9480/Use-of-Synthetic-Data-to-Train-AI-Models.pdf 。

[10]Syntheticus.《合成数据 101：它是什么、如何运作以及用途》。Syntheticus 公司。Web. n.d. https://syntheticus.ai/guide-everything-you-need-to-know-about-synthetic-data 。

于**补充数据集**（并寻求解决数据中的问题，如与偏差或代表性相关的问题）、**创建不存在的数据集**，或在敏感性要求时**删除个人身份信息**。因此，合成数据的价值在于其协助解决**关键数据问题**的能力，**即偏差、代表性、质量、稀缺性和隐私。**

虽然合成数据并不是一个新概念，而且已经使用了一段时间，但最近的技术进步——尤其是在生成式人工智能方面——极大地增加了合成数据的**规模和**制作难度，提高了数据类型的**多样性和质量**。这些进步降低了获取合成数据的门槛，使更多没有丰富专业技术知识的个人和组织能够利用合成数据。一些评估估计，到 2024 年，60%的人工智能训练数据将是合成数据。[11]合成数据的日益普及也是人工智能模型训练对更多数据需求不断增长的结果。

然而，合成数据并非灵丹妙药。事实证明，它有可能延续，有时甚至加剧其使用所要解决的问题。事实上，合成数据本身并不具有私密性、安全性、代表性或公正性，因此必须经过深思熟虑和精心策划才能使其具有私密性、安全性、代表性或公正性。此外，研究表明，对人工智能模型进行重复训练，使其使用以前版本生成的合成数据，可能会导致“模型崩溃”，即模型忘记其基础数据分布，从而导致输出质量和准确性急剧下降。[12]此外，合成数据越来越普遍，可能会扩大数据相关问题的风险面，增加负面影响的程度。[13]。

2.2 军事领域的合成数据

合成数据在军事领域越来越普遍，因为在军事领域，围绕数据稀缺性、偏差和敏感性的问题尤为突出。[14]与民用领域一样，合成数据在这一领域的使用增加与武装部队转向人工智能解决方案有关。在这种情况下，合成数据主要用于**训练军事人工智能模型**，以进行**识别、瞄准、作战和战术规划**，以及开发场景和**合成环境。**

主要而言，合成数据可帮助武装部队**填补空白并提高数据集的质量**，例如从不同角度和不同条件下创建物体图像，以提高人工智能模型的性能。此外，合成数据还可协助**数据管理**，帮助降低与标注和收集相关的成本，**加快**人工智能产品的**开发**。此外，合成数据还可用于创建各种军事行动的逼真模拟，包括对抗性攻击。这些模拟可使各国测试其人工

---

[11]Gartner。"Gartner 确定塑造数据科学和机器学习未来的主要趋势"。2023 年 8 月 1 日。https://www.gartner.com/en/newsroom/press-releases/2023-08-01-gartner-identifies-top-trends-shaping-future-of-data-science-and-machine-learning 。

[12]Shumailov, Ilia, et al."人工智能模型在递归生成的数据上训练时会崩溃"。Nature. https://www.nature.com/articles/s41586-024-07566-y 。

[13]Deng, Harry.2023."探索人工智能和自主系统的合成数据：入门."联合国裁军研究所。https://unidir.org/wp-content/uploads/2023/11/UNIDIR_Exploring_Synthetic_Data_for_Artificial_Intelligence_and_Autonomous_Systems_A_Primer.pdf 。

[14]同上。

智能系统的有效性，制定新的战略和战术，并在受控和安全的环境中为应对更广泛的潜在威胁做好准备。

然而，在军事领域使用合成数据存在与此类数据使用相关的固有风险。事实上，尽管合成数据旨在代表现实状况，但它可能会延续甚至重新解释原始数据中的现有偏见。这种可能性带来了巨大的风险，尤其是在敏感的军事环境中，有偏差的决策可能会造成严重后果。此外，在数据集中重新识别个人或敏感信息的风险依然存在，有可能导致敏感军事数据的泄露，而恶意行为者的“数据中毒”攻击可能会扭曲人工智能系统的学习过程。[15]

3. 治理方面的挑战和影响

要点

- 在民用和军用领域，合成数据的总体管理状况都不成熟。我们需要做更多的工作，以明确现有的管理框架和法规如何适用于合成数据，以及如何对其进行调整，以更好地弥补可能存在的差距。
- 对于是否需要专门针对合成数据制定新的专门法规和框架，目前尚无共识。
- 国际标准是技术管理工具箱中的一个重要工具。虽然目前还没有关于合成数据的国际标准，但制定这些标准的工作正在进行中，这将有助于促进负责任的创新和技术的采用。
- 由于合成数据越来越多地用于军事领域的人工智能系统训练，因此它与军事人工智能治理讨论高度相关。应开展更多工作，应用、调整或借鉴与军事领域数据相关的既定实践和治理概念。
- 合成数据为军事人工智能的管理提供了机遇，因为它有可能促进更多的数据共享、人工智能应用的联合开发以及负责任的合成数据生成和使用指南的共同开发，从而推动全球负责任的军事人工智能目标的实现。
- 军事领域合成数据的管理需要更多利益攸关方的参与。这需要国家之间的合作，也需要与私营部门的合作，私营部门应密切参与治理讨论和努力。促进政府与企业之间的信任将是这项工作的基础。

[15]同上。

3.1 合成数据和民用数据治理

要平衡合成数据的风险和机遇，就必须了解其治理方面的挑战。虽然合成数据并不一定是新事物，但**有关其生成和使用的管理讨论直到现在才在民用和军用领域出现。**有关合成数据的法律地位、监管需求和潜在治理方法的问题尚处于萌芽阶段，总体治理状况仍不成熟。目前还没有专门针对合成数据的立法或框架。欧盟人工智能法案》等一些监管框架顺便提到了合成数据，一些国家的政府也发布了合成数据生成指南。[16]

在民用领域，对于合成数据是否对数据监管和治理框架构成挑战，以及如果构成挑战，以何种方式构成挑战，目前尚无共识。例如，有人指出，合成数据可能会挑战个人/非个人数据的类别，而这些类别是欧盟 GDPR 等数据管理法规和框架的基础。**有人认为，这些法规不足以应对合成数据的复杂性，因为合成数据会模糊这些类别之间的界限。**根据合成数据的类型——完全合成、部分合成或混合——存在的个人信息水平和重新识别的风险，以及数据保护法的适用性会有很大不同。[17]这种模糊性给合成数据的开发者和使用者都带来了法律上的不确定性。因此，**随着合成数据的使用越来越多，有必要扩大数据保护法规中传统的个人/非个人数据范式的范围。**

对其他人来说，适当生成的合成数据是一种有用的隐私增强技术，可用作遵守各种数据监管框架的工具。此外，合成数据还可用于实现更广泛的数据治理目标，在**保护隐私的**同时实现有价值数据获取的民主化，建立透明的数据目录和审计跟踪以促进**问责制，**通过提供一致可控的数据源**提高数据质量**，以及**促进**国家和国际层面的**安全数据共享**。对于合成数据是一种需要激励的有用工具，还是一种可能破坏为防范各种数据相关风险而制定的法律机制的创新，一直存在分歧。[18]

---

[16]新加坡个人资料保护委员会。2024."Privacy Enhancing Technology Proposed Guidance on Synthetic Data Generation"。新加坡个人数据保护委员会。 https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/other-guides/proposed-guide-on-synthetic-data-generation.pdf 。

[17]完全合成数据集包含完全由人工智能模型生成的数据，不包含真实世界的数据。该模型能识别数据集的统计特性和模式，并生成一个全新的数据集。部分合成数据会替换数据集中某些选定的敏感特征，并用合成值取而代之，同时保留一些真实数据。混合合成数据结合了真实世界和完全合成数据，将真实数据集中的随机记录与合成记录配对。更多详情，请参阅 Syntheticus。《合成数据 101：它是什么、如何运作以及用途》。Syntheticus 公司。Web. n.d. https://syntheticus.ai/guide-everything-you-need-to-know-about-synthetic-data 。和 IBM。《什么是合成数据》。 https://www.ibm.com/topics/synthetic-data 。

[18]Gal、Michal、Lynskey、Orla.2023."合成数据：数据生成革命的法律意义"。Iowa Law Review.109.即将出版。 https://papers.ssrn.com/sol3/papers.cfm? 。

因此，合成数据的处理缺乏法律和规范的明确性。因此，一些人认为有必要制定明确的指导方针，以确保各类合成数据处理的**透明度**、**公平性**和**问责制**，并加强对用于生成合成数据的基础模型所使用的数据的明确性和指导方针。[19]建议包括：

- **透明度**：合成数据应明确标注为合成数据，并提供其生成过程的信息；
- **问责制**：应制定明确的程序，对负责生成和处理合成数据的人员进行问责；
- **公平性**：合成数据应保证其生成和使用方式不会带来不利影响，如让偏见永久化或产生新的偏见。

由于缺乏明确的法律规定，合成数据有可能不在监管范围之内，同时也有可能带来一些与这些治理框架试图解决的真实世界数据相同的问题。**因此，民用领域的一个关键治理挑战将是确保合成数据的开发和使用方式，如果它不在当前数据法规的范围内，则不会延续或造成新的危害。**因此，需要开展进一步的研究和工作，以更清晰地阐明合成数据的法律地位，并找出现有法规和框架中潜在的治理差距，为合成数据的开发者和使用者提供清晰度。

3.2 标准的作用

标准（包括技术标准和非技术标准）是民用技术管理的一个重要方面。在合成数据方面，这些标准既是急需的，也是当前工作的一个重要领域。**事实上，目前还没有生成合成数据的国际标准，也**没有普遍认同的定义或基准来评估合成数据的质量和可信度。**通过提供明确的定义、方法和评估标准，标准可以建立评估合成数据的共识和基准**。标准可以向组织保证，他们使用的合成数据符合特定的质量和隐私阈值。此外，合成数据的标签和文档标准，以及跟踪来源的审计和机制，可以成为确保合成数据生成和处理过程中的透明度、公平性和问责制的关键基石。

这方面的工作正在各条战线上展开。例如，电气和电子工程师学会（IEEE）正在牵头制定隐私安全合成数据的全球标准和最佳实践。[20]国际标准化组织 (ISO) 也在开展类似的工作。[21]开放数据研究所（ODI）为开发一种名为 “Croissant”的工具做出了贡献，这是一

---

[19]Beduschi, Ana.2024."合成数据保护：迈向数据监管的范式变革？"大数据与社会》。第 11 卷（1）。 https://doi.org/10.1177/20539517241231277 。

[20]电气与电子工程师学会。《合成数据》。电气与电子工程师学会。n.d. https://standards.ieee.org/industry-connections/activities/synthetic-data/ 。

[21]国际标准化组织。《信息技术——人工智能——人工智能系统背景下的合成数据概述》。国际标准化组织。无日期. https://www.iso.org/standard/86899.html#lifecycle 。

种社区标准，可为数据集提供机器可读的元数据，有助于实现机器学习数据集文档的标准化。[22]标准将为负责任的创新提供框架和参数，激励私营部门在合成数据的生成、使用和创新方面采取良好做法。

3.3 合成数据与军事人工智能的国际管理

由于人工智能在军事领域的许多方面日益占据中心地位，数据在军事上的重要性与日俱增。有鉴于此，而且由于合成数据在军事领域主要用于训练和开发各种人工智能能力，因此有关合成数据对国际安全影响的治理讨论应在军事人工智能治理的背景下进行。

数据及其相关问题已被确定为军事领域负责任人工智能的优先工作领域之一。[23]**然而，这些问题并未成为当前治理工作的核心，**在很大程度上仍然是高层次的，缺乏细化。因此，**围绕合成数据、其治理以及重要的是其对当前军事人工智能治理工作的潜在影响所展开的**讨论**本身也是萌芽状态。**一些区域层面的讨论虽然同样处于萌芽阶段，但已经展开，特别是在世界上仍在数字化的地区，这些地区的数据缺口是进入军事人工智能的一个重大障碍，而合成数据等替代品可以在那里发挥替代作用。

a. 关于挑战的新颖性和现有框架的适用性

军事领域的合成数据会带来与真实数据类似的挑战和影响，如偏差（包括偏差的延续、加剧或产生）、可靠性和代表性问题、问责制、可追溯性或缺乏可解释性等。因此，**在军事人工智能的背景下，合成数据不应被排除在外，也不应该仍是一个未得到探索的问题。**重要的是，这可能导致某些类型的数据仍处于治理讨论范围之外，同时可能使风险长期存在，从而加剧军事人工智能中与数据相关的风险。

例如，确保数据*问责*——责任的一个关键原则——在合成数据的情况下可能会更加复杂。事实上，合成数据的使用增加了一层负责数据生成的人员，有时是外部行为者，从而在出现错误时更难追查直接责任。与此相关的是，**由于在合成数据集的生成、使用和标记方面缺乏国际公认的标准，**合成数据可能会加剧数据*可解释性*问题。由于可追溯性有限，数据来源不明确，这可能会阻碍审计能力，从而无法解决数据集中的偏差问题。此外，通过合成数据实现数据访问的民主化，可以为人工智能和其他数字能力的发展提供更多机会。一方面，这有助于解决数字鸿沟问题。另一方面，在国际安全方面，通过降低开发先进人

---

[22]Simperl、Elena 和 Thomas Carey-Wilson。《开放数据研究所帮助制定机器学习数据开放元数据标准》。开放数据研究所。 https://theodi.org/news-and-events/blog/the-odi-to-help-develop-an-open-metadata-standard-for-machine-learning-data/ 。

[23]Afina, Yasmin, Persi Paoli, Giacomo.2024.《军事领域的人工智能治理：关于优先领域的多个利益攸关方观点》。联合国裁军研究所。 https://unidir.org/wp-content/uploads/2024/09/UNIDIR_Governance_of_Artificial_Intelligence_in_the_Military_Domain_A_Multi-stakeholder_Perspective_on_Priority_Areas.pdf 。

工智能模型的准入门槛，这也可能成为**人工智能军事能力进一步扩散的**助推器。[24]然而，这些并不代表新的挑战，而是军事领域原有数据挑战的进一步复杂化，需要在合成数据的背景下予以特别关注和澄清。

因此，围绕新技术对国际安全影响的治理讨论最好首先关注法律和规范框架是否适用以及如何适用。在合成数据方面，重要的是分析这些治理挑战是否是新的，是否只是使已有的数据挑战复杂化，以及现有的数据治理概念在多大程度上适用于合成数据。因此，与一定要设计新的治理框架或方法相反，国际社会应着眼于既有的做法和概念，如公平性、责任、可追溯性和可靠性，并努力将其应用、调整或发展到合成数据中。在这一问题上，国际社会并非从零开始，而是可以利用关于何为“良好数据 ”的大量工作，以及关于军事领域良好数据做法的新知识。

插文 1.
未来研究领域：合成数据集的国际贸易

需要进一步研究的一个重要领域是合成数据集国际贸易的潜在影响，以及在某些情况下是否应该以及如何控制、监督或限制这种贸易。事实上，合成数据集交易市场可能会发展起来，恶意行为者可能会利用这些数据集开发破坏性人工智能能力。因此，国际社会应考虑合成数据集贸易如何与防扩散努力和军备控制相衔接。国际社会还应探讨是否应通过出口管制清单等工具来控制某些合成数据集或合成数据类型。

b.多利益攸关方方法的重要性

能够最有效地利用合成数据为军事领域带来的益处的管理框架是**多利益攸关方性质**的框架。这不仅需要国家之间的合作，**还需要与私营部门行为体密切合作，后者应参与此类治理讨论。**私营部门在人工智能军事领域发挥着重要作用，主要负责核心技术的研发。在合成数据方面也是如此，合成数据集的生成、测试和评估能力大多由行业参与者掌握。**这就造成了对私营技术公司的额外依赖**，特别是对资源较少的国家而言，这些国家可能不具备独立的合成数据测试和评估能力，在合成数据集的质量方面依赖于私营部门。这种依赖性要求建立公私合作框架，优先考虑**建立政府与企业之间的信任**。这种信任对于确保私营部门参与治理讨论，并在为军事人工智能系统开发、部署和测试合成数据时采取负

[24]Maas, Matthijs M. 2019.《军事人工智能的创新型全球治理？我是如何学会停止担忧并爱上机器人的》。《国际人道主义法律研究杂志》.10 (1). https://brill.com/view/journals/ihls/10/1/article-p129_129.xml?language=en 。

责任的做法至关重要。此外，多利益攸关方的方法有助于建立急需的共同语言和对合成数据的理解。

c.关于指导方针和具体情况

虽然在合成数据的生成和处理方面制定明确的准则应该是一个目标，但有人指出，特别是在军事背景下，制定这些准则可能为时过早。事实上，明确的指导原则通常以定义明确的最佳实践为基础。然而，**就合成数据而言，该领域可能仍处于起步阶段，无法确立明确的最佳实践。**在这种情况下，在透彻了解技术的能力、优势、局限性和潜在风险之前，就对测试程序进行标准化或制定严格的指导方针，可能会适得其反。

此外，**明确的指导方针可能无法适应在军事背景下评估合成数据使用的适当性和责任程度的高度特定性。**事实上，特定合成数据集在某种情况下可能被“负责任地”使用，而在另一种情况下使用则可能被视为“不负责任”。合成数据使用的这种环境依赖性使我们难以制定普遍适用的准则，以有效解决不同使用情况下的细微差别。

此外，质量指标本身也取决于具体情况。例如，合成数据集与现实的接近程度被用作其质量的关键指标。[25]然而，在某些情况下，特别是在军事领域，偏离这一点恰恰是目的所在。换句话说，**在军事行动中使用合成数据来表现前所未有的情景，以帮助进行创造性规划，这可能是合成数据的优势之一。**

此外，治理方法需要考虑到地区和国家背景。由于合成数据可能产生外部依赖性——如依赖外部行动者生成合成数据集并保证其质量——确保**在特定区域外创建的合成数据集反映预期使用环境中的当地现实尤为重要。**这就要求合成数据集的参数和假设具有透明度，说明为什么要创建合成数据、如何创建、创建的目的和创建者。

d.关于军事领域的治理机会

合成数据不仅带来了治理方面的挑战，也带来了**机遇，尤其是对军事人工智能的治理而言。**合成数据确实有可能促进武装部队之间的**数据共享能力**，并有助于**共同开发军事人工智能能力。**例如，合成数据的隐私保护潜力可为共享数据集提供机会，而这往往是人们所希望的，但却因军事数据的敏感性和机密性而受到阻碍。在军事领域，这将在政府组织和国家内部及之间产生巨大的价值。

---

[25]Deng, Harry.2023.《探索人工智能和自主系统的合成数据：入门》联合国裁军研究所。https://unidir.org/wp-content/uploads/2023/11/UNIDIR_Exploring_Synthetic_Data_for_Artificial_Intelligence_and_Autonomous_Systems_A_Primer.pdf 。

合成数据可作为国家间合作开展军事人工智能项目的“中立地带”。通过使用反映真实世界场景但不包含敏感信息的合成数据集，**各国可以合作开发和测试人工智能系统，增强互操作性，并分享最佳实践，而不会有与交换真实军事数据相关的风险。**此外，各国还可合作开发集体合成数据集，用于训练和测试人工智能系统，以增强互操作性——这是开发军事人工智能能力的一个关键问题。这种合作方法可以增强盟军之间的凝聚力，提高联合行动的效率，并有助于在军事人工智能方面建立一个更加稳定和安全的国际环境。

此外，合成数据为各国提供了一个机会，可借以制定有关军事领域生成和使用**合成数据**的**共同规范和准则**。由于合成数据的新生性质，围绕合成数据的讨论为国际社会提供了一个制定**共同责任框架的**机会。多个国家可就治理原则达成一致，并开始分享最佳实践，以系统的方式共同建立合成数据生成和使用的良好实践。例如，数据管理多边工作组可以共同解决一些问题，并为制定程序、流程和问责标准提供一个论坛。

4. 结论

合成数据在推进民用和军用领域的人工智能能力方面具有巨大潜力。其优势包括：解决数据稀缺问题、提高隐私保护、促进创建更具代表性和更少偏见的数据集，这些都使合成数据成为一个强大的工具。然而，合成数据并非万能的灵丹妙药，其使用也存在固有风险。要想从这项技术中获取最大利益，最重要的是要开始讨论治理问题。目前，这些工作在民用和军用领域都处于起步阶段，合成数据的生成、处理和使用在法律和规范上仍然存在模糊之处。为避免出现法律和规范真空，使合成数据风险得不到解决，应努力找出现有框架中的漏洞，并为合成数据的用户和生成者提供清晰的说明。为此，指导方针、国际技术标准的制定以及与产业界的合作将是关键所在。

在军事领域，与数据相关的问题仍处于人工智能治理讨论的边缘。在这种情况下，不仅要努力确保其在这些工作中的地位，还要特别考虑合成数据在军事领域使用的影响。由于合成数据主要是使现有的治理挑战复杂化，而不是创造一个全新的环境，因此不一定需要制定新的框架或法规。这可能需要将军事数据最佳实践和概念应用于合成数据的生成和使用。因此，需要就这一问题开展更多工作，将新出现的军事人工智能治理框架扩展到合成数据，明确如何应用这些实践和概念。

随着合成数据的日益普及，它不仅带来了治理方面的挑战，也带来了国际合作的机遇，这可能会对全球军事人工智能治理产生积极的下游影响。为此，多个利益攸关方的努力将有助于汇聚各国，尤其是私营部门的力量。

展望未来，继合成数据之后，数据科学肯定还会有创新。这凸显了创建适应未来发展的管理框架的重要性。建立以灵活性为核心的框架对其可持续性至关重要，可确保在新技术和新用例出现时，这些框架仍能保持相关性。

附件：活动议程和与会者

介绍性发言

- **Federico Mantellassi**，联合国裁军研究所研究员

合成数据与国际安全技术早餐会：

- **Eleonore Fournier-Tombs 博士**，联合国大学政策研究中心预测行动与创新负责人
- **Calum Inverarity**，开放数据研究所高级研究员

由联合国裁军研究所副研究员何雯婷主持

关于合成数据的多利益攸关方对话：国际治理的机遇与挑战

- **Jane Pinelis 博士**，约翰霍普金斯大学应用物理实验室应用信息科学处首席人工智能工程师
- **Aldo Lamberti**，Syntheticus 创始人兼首席执行官；欧洲主题专家
  委员会副主席；IEEE 工业连接合成数据委员会副主席；IEEE 生成式预训练人工智能 (AI) 模型安全和可信度要求标准工作组专家
  生成式预训练人工智能 (AI) 模型安全和可信度要求标准工作组专家，电气和电子工程师学会
- **Yasmin Afina**，联合国裁军研究所研究员；军事领域负责任人工智能全球委员会专家
- **Ana Beduschi 博士**，埃克塞特大学个人讲座法学正教授；埃克塞特大学法学院科学、文化与法律研究中心主任

由联合国裁军研究所研究员 Federico Mantellassi 主持