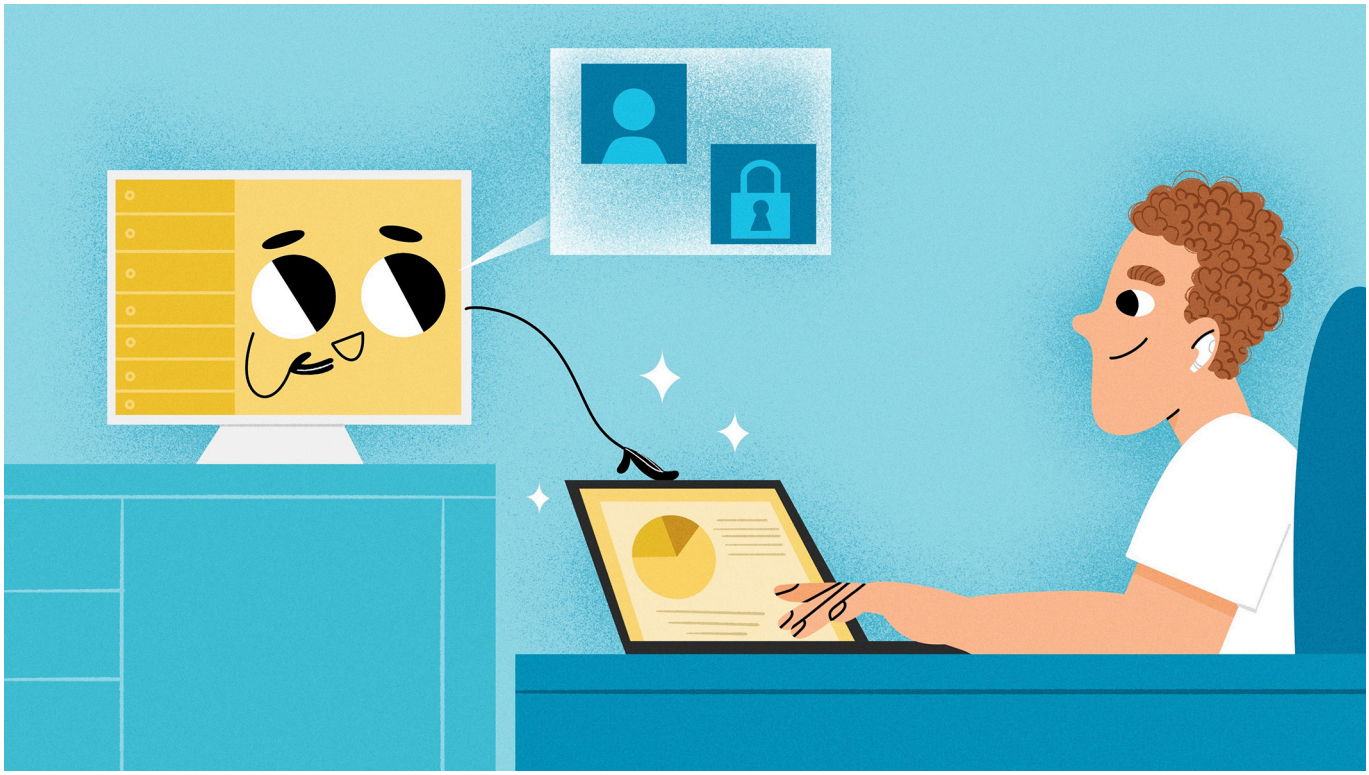
工程与技术
线上发布: 2025 年 9 月 30 日
doi: 10.3389/frym.2025.1575140-zh



如何让研究人员更安全地共享敏感数据

Olga Fink[1], *Lisette van Gemert-Pijnen*[2*], *Dongwon Lee*[3], *Andrew Maynard*[4] 和 *Bastiaan van Schijndel*[5]

[1]洛桑联邦理工学院, 智能维护与运营系统实验室（瑞士, 洛桑）
[2]特文特大学, 劝导健康技术（荷兰, 恩斯赫德）
[3]宾夕法尼亚州立大学, 信息科技学院（美国, 宾夕法尼亚州, 大学园）
[4]亚利桑那州立大学, 社会创新未来学院（美国, 亚利桑那州, 坦佩）
[5]*ZorgTTP* 基金会（荷兰, 豪滕）

少年审稿人

ANSAR

年龄: 14

为解决疾病、环保等复杂问题, 数据共享与分析至关重要。然而, 医疗记录、财务信息等敏感数据必须保持私密安全。新兴技术可创建不包含个人可识别私密或敏感信息的仿真数据, 提升共享或使用敏感数据时的安全性。目前, 合成数据与加密技术已被用于研究疾病、检测欺诈和预测自然灾害等罕见事件。尽管仍存在合成数据精确度不足、数据创建能耗过高等问题, 这些技术终将为全球科研协作开辟更安全高效的数据共享路径。

我们需要更安全的数据共享方式

我们每天都身处数据的包围中。数据包含数字、测量结果和图像等事实与信息, 它们帮助我们认知世界。从医疗记录到卫星图像, 数据让我们得以了解周遭世界的状况、解决难题并取得新的发现。例如, 科学家利用数据研究疾病传播规律、预测自然灾害, 并改进可再生能源等技术。

人工智能（AI）通过分析海量数据、发现人类可能忽略的规律，提升科研发现效率（点击此处 深入了解 AI 及其在科研发现中的作用）。

运行 AI 需要获取大量数据，但共享某些类型的数据存在一些大问题。像"某城市年均晴天数"这类普通数据可以自由共享，而包含个人详细信息的数据则被视为敏感数据，需要特殊保护，这主要涉及隐私和安全。隐私要求保护个人信息，未经授权不可共享，例如患者的医疗记录必须保持私密；安全则要确保无关人员无法接触敏感数据，例如政府档案或商业机密等敏感信息需要防范黑客攻击等威胁（更多数据隐私和安全知识可参阅这篇《Frontiers for Young Minds》 文章）。

敏感数据不限于医疗记录，像银行账户信息这类财务数据同样敏感，可能被用于诈骗；姓名、住址等个人身份信息则可能导致身份盗用；甚至连人们的网购记录、社交媒体观看偏好等数据也具有敏感性，可能被滥用，诱导用户做出不理智的选择。此外，政府和科研机构处理的许多信息也属于敏感数据，例如军事禁区地图关系国家安全，濒危动物栖息地数据涉及生态保护，这些数据都需要严格的安全防护措施。

另一大挑战在于数据主权，即确保数据始终处于其合法所有者（个人、企业或国家）的掌控之下。例如，某国的隐私保护法律可能禁止将医疗数据共享给境外研究人员，即便相关研究能够挽救生命。这些法规对于保护个人与组织至关重要，但也给科学家合作解决复杂问题设置了障碍。那么，是否存在一种既能共享敏感数据或训练 AI 系统，又不会危及隐私、安全或主权的方法？

前沿技术：合成数据

合成数据与其他隐私增强技术使研究人员能够安全地共享信息，既为科学发现开辟新途径，又确保敏感数据安全无忧 [1]。这些技术有望重塑全球科研机构与企业的协作模式，助力攻克人类面临的一些重大挑战。

要理解合成数据，不妨将它们视为真实数据的"高仿真副本"——虽然不是原件，却能精准模拟原始数据的规律与趋势。以医院的患者康复记录为例：合成数据会呈现真实的康复规律（如患者的平均痊愈时长），但绝不会包含任何真实患者的个人信息。这种特性使得研究人员在分析时完全无需担忧隐私泄露风险。合成数据的生成依赖于名为"生成对抗网络"（**GAN**）的 AI 技术。GAN 如同数字艺术家，可深度学习真实数据，进而创造出足可以假乱真的人工数据。例如，经过人脸图像训练的 GAN 能生成栩栩如生的虚拟面容，而这些面容在现实中并不存在。GAN 的运行流程分为两个阶段：首先利用真实数据对 AI 系统进行"教学"；待学习完成后，AI 系统便会与数据合成系统协同生成高度逼真的合成数据。由此产生的数据集既有实用价值，又杜绝了敏感信息暴露的风险。

人工智能 (Artificial Intelligence)

一种计算机技术，可帮助机器像人类那样思考、学习并解决问题，比如识别面容、预测天气或设计新药。

敏感数据 (Sensitive Data)

需要保护的信息，如医疗记录、财务数据或个人信息，共享此类信息可能造成伤害或侵犯隐私。

隐私 (Privacy)

保护医疗记录等个人信息的安全，未经个人允许不得向他人披露。

安全 (Security)

保护数据免遭盗窃或滥用，例如保护银行数据库免遭黑客入侵。

数据主权 (Data Sovereignty)

确保数据即便在跨境共享过程中也始终处于其所有者（个人、企业或国家）的掌控之下。

合成数据 (Synthetic Data)

由计算机生成的、用于模拟真实数据的人工数据，可安全地用于科学研究，不会泄露隐私细节。

生成对抗网络 (GAN) (Generative Adversarial Networks (GANs))

一种采用两个计算机系统的 AI 技术：一个从现实数据中学习规律，另一个则生成符合这些规律的合成数据。

同态加密 (Homomorphic Encryption)

一种数据保护手段，可将数据转换为加密代码，加密后的数据可直接用于分析，无需解密原始信息。

另一项关键技术是同态加密 [2]。此技术并非创建新的数据副本，而是将原始数据转换为一种特殊的"密码"，可直接用于分析，无需解密。合成数据与同态加密的关键区别在于对原始数据的处理方式：合成数据会生成一个全新的模拟数据集，完全替代原始数据，但遵循相同的规律；同态加密则保留原始数据，将它"锁"在加密系统中，只有密钥持有者能直接访问原始数据，同时又允许在不解密的情况下进行数据运算。简单地说，合成数据就像博物馆文物的逼真仿品，外观相同但不是原件；同态加密就像将文物锁在箱子中，无需开箱就能进行称重等操作，但始终看不到实物。

技术强化数据保护

合成数据与同态加密技术正以富有创意的新方式帮助科学家和企业解决敏感数据难题（图 1）。

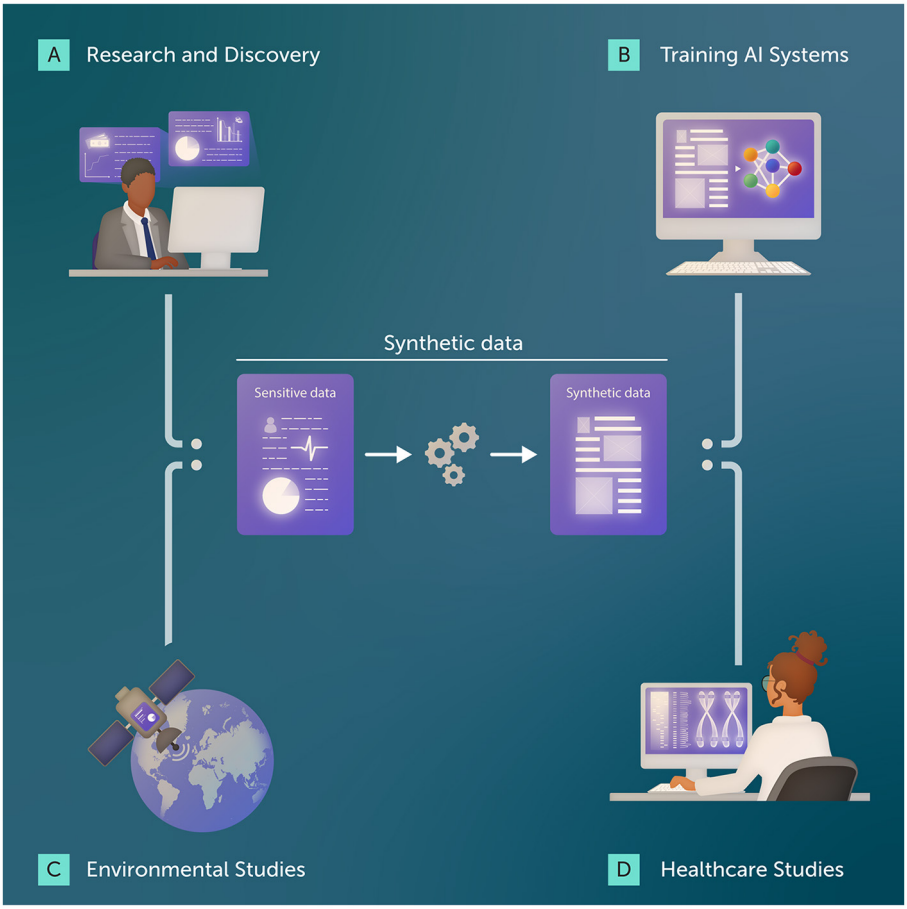


图 1

图 1

合成数据既能帮助研究人员通过数据分析获得重要发现，又能确保敏感信息得到保护：**(A)** 银行家与经济学家可利用合成数据研究"重大股灾如何影响投资者"等课题；**(B)** 训练 AI 系统需要海量数据，合成数据可满足此需求；**(C)** 加密卫星数据可在保护国家安全的前提下追踪环境变化；**(D)** 医生和科学家可借助合成数据研究疾病与治疗方案，不会泄露患者隐私。

在医疗领域，由于隐私法规通常限制研究人员使用真实病历，合成数据可支持疾病与治疗方案研究，同时保障患者隐私安全 [3]。例如，科学家可创建模拟真实病历规律的合成数据集，呈现不同患者群体对特定药物的反应差异。通过分析这些规律，研究人员可能会发现该疗法对携带特定基

因或处于某些健康状况的患者更有效，从而开发出更加个性化的精准医疗方案，并且整个过程完全不涉及真实患者信息。合成数据还有助于解决罕见病等数据稀缺难题。

合成数据可安全训练需要学习海量数据的 AI 系统，避免使用私密敏感信息。企业可借助 AI 发现趋势、预测需求并优化决策。基于合成数据训练的 AI 既可识别异常消费模式（表明受到欺诈），也能预测客户潜在需求（如购买登山包的消费者往往也需要水壶）。这种技术可提升产品推荐精准度，又不会泄露消费者的真实信息。此外，医疗领域的真实数据较为稀缺，且采集通常耗时费力，合成数据可在这方面显著降低研究成本。

当数据必须在受保护状态下分析时，这就是同态加密的用武之地。例如，政府可利用加密卫星图像追踪森林退化或冰川消融等环境变化。这些图像可能包含国家安全敏感资源或区域的位置信息，而加密技术能使研究人员分析森林缩减速度、气温对冰层影响等规律，同时杜绝数据滥用风险。同样，加密医疗数据可供科学家研究全球疾病趋势，帮助定位疫情热点并预测爆发风险，同时严格保护个人隐私。

最后，这些技术还能帮助研究人员预测经济危机或自然灾害等罕见事件。研究人员可通过合成数据模拟异常场景，然后安全测试应对方案：例如，银行可模拟股市崩盘对储蓄的影响，并据此制定更完善的客户保护方案。

重大挑战与更大机遇

合成数据和同态加密技术让信息共享与分析变得日益便捷和安全。这些工具既能帮助研究人员从数据中获取更多洞见，又能确保敏感信息始终安全，为曾经难以实现的重大突破铺平道路。

然而，这些技术仍面临一些挑战。合成数据并非完美无缺，如果原始数据存在偏差，基于其生成的合成数据也会继承这些缺陷，正如数据科学家所说的''输入垃圾，产出垃圾''。此外，如果合成数据过度简化或曲解现实数据，同样可能导致分析失真。同态加密虽然安全性极高，但加密过程耗时耗电，难以用于大型项目。另一个风险是黑客可能破解加密或合成数据，导致隐私泄露。信任同样是一大挑战：要让这些技术真正发挥作用，公众需要了解其原理并确信其安全性。科学家、医生和政府领导者必须共同制定明确规范，并向公众普及这些技术的工作原理。

尽管存在这些挑战，合成数据与其他安全数据共享工具在应对重大社会问题（如抗击疾病或气候变化）方面潜力巨大，同时能够保护敏感数据。通过构建更安全的数据共享环境，这些工具有望开创全球科学家与组织无障碍协作的新纪元。

致谢

由 SJD Consulting, LLC. 科学撰稿人/编辑、毕业于美国马萨诸塞大学陈氏医学院晨兴生物医学研究生院的 Susan Debad 博士参与撰写和编辑。图表制作方为 Somersault18:24。

AI 人工智能工具使用声明

本文中所有图表附带的替代文本（alt text）均由 Frontiers 出版社在人工智能支持下生成。我们已采取合理措施确保其准确性，包括在可行情况下经由作者审核。如发现任何问题，请随时联系我们。

参考文献

1. Jordon, J., Szpruch, L., Houssiau, F., Bottarelli, M., Cherubin, G., Maple, C., et al. 2022. Synthetic Data – what, why and how? *arXiv* [preprint] arXiv:2205.03257. doi: 10.48550/arXiv.2205.03257
2. Rivest, R. L., and Dertouzos, M. L. 1978. *On Data Banks and Privacy Homomorphisms*. Available online at: https://luca-giuzzi.unibs.it/corsi/Support/papers-cryptography/RAD78.pdf (accessed May 7, 2025).
3. Gonzales, A., Guruswamy, G., and Smith, S. R. 2023. Synthetic data in health care: a narrative review. *PLOS Digital Health* 2:e0000082. doi: 10.1371/journal.pdig.0000082

线上发布: 2025 年 9 月 30 日

编辑: Robert T. Knight

科学导师: Adiya Rakymzhan

引用: Fink O, van Gemert-Pijnen L, Lee D, Maynard A 和 van Schijndel B (2025) 如何让研究人员更安全地共享敏感数据. Front. Young Minds. doi: 10.3389/frym.2025.1575140-zh

英文原文: Fink O, van Gemert-Pijnen L, Lee D, Maynard A and van Schijndel B (2025) Safer Ways for Researchers to Share Sensitive Data. Front. Young Minds 13:1575140. doi: 10.3389/frym.2025.1575140

利益冲突声明: 作者声明本研究不涉及任何潜在商业或财务关系。

版权 © 2025 © 2025 Fink, van Gemert-Pijnen, Lee, Maynard 和 van Schijndel. 这是一篇依据 Creative Commons Attribution License (CC BY) 条款发布的开放获取文章。根据公认的学术惯例，在注明原作者和版权所有者，及在标明本刊为原始出处的前提下，允许使用、传播、复制至其他平台。如违反以上条款，则不得使用、传播或复制文章内容。

少年审稿人

ANSAR, 年龄: 14
我今年 14 岁，住在哈萨克斯坦，是一名狂热的职业游泳运动员，也是当地致力于打造环保可持续发展方案的城市规划团体的一份子。在课余时间，我热爱烘焙、舞蹈、读小说，还喜欢和朋友们一起玩。

作者

OLGA FINK
Olga Fink 现任瑞士洛桑联邦理工学院（EPFL）智能维护与运营系统实验室主任，运用人工智能技术改善机械设备与基础设施的性能与寿命。2022 年加入 EPFL 前，她曾任苏黎世联邦理工学院教授，并主导苏黎世应用科学大学的智能维护课题组。她拥有苏黎世联邦理工学院博士学位，曾荣获世界经济论坛"青年科学家"称号。

LISETTE VAN GEMERT-PIJNEN
Lisette van Gemert-Pijnen 是荷兰特文特大学的劝导健康技术教授，聚焦于通过劝导设计提升技术可信度与依从性，并开发实际应用方法。她创立了首个电子健康研究中心，主导制定 CEHRES 电子健康路线图，目前参与校级战略研究项目，以加速健康技术应用落地。她还是《Frontiers》期刊健康技术实施版块主编。(https://www.utwente.nl/en/techmed/research/research-programmes/sht/). *j.vangemert-pijnen@utwente.nl

DONGWON LEE
Dongwon Lee 现任宾夕法尼亚州立大学信息科技学院教授，研究领域涵盖数据科学、机器学习及网络安全，重点关注虚假新闻、网络欺诈等问题。他持有加州大学洛杉矶分校"计算机科学"博士学位，曾在 AT&T 贝尔实验室工作。作为美国国家科学基金会项目主管，他参与管理网络安全教育与研究项目。现为 ACM 杰出会员及富布赖特网络安全学者，带领宾夕法尼亚州 PIKE 研究组开发应对虚假新闻的计算与社会技术解决方案。

ANDREW MAYNARD
Andrew Maynard 是一位研究新兴技术如何影响社会的科学家兼作家，现任亚利桑那州立大学（ASU）社会创新未来学院教授，兼任风险创新实验室主任。他专注于纳米技术、人工智能等新兴技术的风险与效益评估，著有《Films from the Future》、《Future Rising》等书籍，探讨技术对大众生活的影响。他还通过播客和文章分享个人见解，帮助公众理解技术与社会之间的复杂关系。

BASTIAAN VAN SCHIJNDEL
Bastiaan van Schijndel 是荷兰 ZorgTTP 基金会的创新经理，该机构专注于数据安全与隐私解决方案。他主要从事医疗等领域的敏感信息保护技术开发工作，包括实施高级加密方法，以确保数据在接受处理和分析时的私密与安全。此外，他还参与世界经济论

坛''十大新兴技术''报告等国际倡议，在隐私增强技术等前沿领域的研讨中积极分享见解。

中文翻译由下列单位提供
Chinese version provided by