

doi: 10.7690/bgzd.2025.05.003

生成式人工智能军事应用评估

侯立志, 李立伟, 陈鑫, 赵擎天

(军事科学院系统工程研究院, 北京 100101)

摘要: 为促进生成式人工智能在军事领域高效发展, 对其进行探析。提出关于研究生成式人工智能军事应用评估的重要意义, 分析评估工作应遵循的基本原则, 建立评估的指标体系, 研究评估应采取的基本方法, 并就生成式人工智能军事应用评估问题提出几点思考。结果表明, 该研究可促进生成式人工智能与军事应用紧密结合, 加快技术落地并高效发展。

关键词: 生成式人工智能; 军事应用; 评估

中图分类号: TP18 **文献标志码:** A

Evaluation of Generative Artificial Intelligence for Military Application

Hou Lizhi, Li Liwei, Chen Xin, Zhao Qingtian

(Institute of Systems Engineering, Academy of Military Sciences, Beijing 100101, China)

Abstract: In order to promote the efficient development of generative artificial intelligence in the military field, this paper analyzes it. This paper significance of advanced research on the military application evaluation of generative AI, analyzes the basic principles that should be followed in the evaluation work, establishes the evaluation index system, studies the basic methods that should be adopted in the evaluation, and puts forward some thoughts on the military application evaluation of generative AI. The results show that this research can promote the close integration of generative AI and military applications, accelerate the technology landing and efficient development.

Keywords: generative artificial intelligence; military application; evaluation

0 引言

生成式人工智能可以在海量数据上进行无监督预训练, 并自己寻找和发现规律, 进而分类、优选、积累和生成答案, 其强大的网络渗透能力、舆情认知能力和分析与处理能力具有颠覆战场应用的巨大潜力, 已成为当前军事智能化发展关注的热点^[1]。自哈马斯对以色列发动突袭以来, 双方的冲突已引发前所未有的信息浪潮, 生成式人工智能技术生成的虚假图像已经产生了微妙且复杂的作用, 影响战争外围。由此可见, 生成式人工智能的军事应用已初见端倪, 必将在未来战场大展拳脚。笔者关于思考生成式人工智能军事应用评估问题, 既有利于生成式人工智能与军事应用的紧密结合, 又有助于其在军事领域高效发展。

1 生成式人工智能军事应用评估的重要意义

传统意义上的人工智能更多的依赖于先前确定的规则和算法来处理和解决问题。与传统人工智能不同, 生成式人工智能可以从机器学习过程中模仿现有数据的模式和特征来创建新的内容, 甚至可以

通过自身的创造力来生成前所未有的内容, 如基于用户的输入自动生成图像、音乐或视频等多模态的输出, 更具灵活性和创造性^[2]。作为新生技术, 将其应用于实践具有两面性, 带来机遇的同时也伴随风险挑战。其向军事应用转化是一把双刃剑, 既有可能颠覆传统作战样式, 又存在极大的不确定性, 甚至产生事与愿违的结果, 效果究竟如何, 需要针对性的试错与检验。关于研究生成式人工智能军事应用评估问题, 实现以评促用、以评增效、以评助管, 对于推进作战体系能力生成和提高具有重要意义。

1.1 以评促用

当前关于生成式人工智能军事应用的研究仅停留在讨论层面, 生成式人工智能究竟如何与军事应用相结合, 具体应用于军事领域的哪些方面, 以及如何应用都尚不清晰。关于谋划生成式人工智能军事应用评估问题, 通过评估判断生成式人工智能应用于军事领域的可行性、是否应用在正确的方向和方式上、以何种方式应用以及应用的代价有多大, 能够促进其与军事应用相结合, 加快技术落地, 更好地应用于军事领域。

收稿日期: 2024-08-05; 修回日期: 2024-09-06

第一作者: 侯立志(1987—), 男, 天津人, 硕士。

1.2 以评增效

生成式人工智能应用于军事领域后究竟能迸发出多大能量,能为作战体系带来多大作用,目前尚未可知。通过评估,能够在评估过程中衡量相应指标在生成式人工智能加持下的优劣程度,判断生成式人工智能落地军事领域的发展方向和方案计划是否具有可行性、科学性和长效性,进而明确是否按照既定方案持续发力还是需要选择不同的发展方向和结合方式。同时,能够及时发现应用环节存在的问题,通过改进技术,不断试错,生成式人工智能可适用于不同的军事场景并最大限度发挥作用。

1.3 以评助管

生成式人工智能应用于军事领域不是一蹴而就的,涉及到人力、装备、财务、资源、政策、法规等各方面的管理,任何方面出现问题,都会导致军事应用受阻。进行专业评估,对于管理生成式人工智能在军事领域的应用有着关键性的意义。通过评估能够规范生成式人工智能装设备研发、定型采购以及日常管理制度和机制,防止重复性研究、低水平研究泛滥,节约经费,促进经费保障更加合理规范。确保生成式人工智能在军事领域应用的过程中向作战聚焦、向使用聚焦、向体系聚焦。

2 生成式人工智能军事应用评估的基本原则

对生成式人工智能军事应用进行评估是一项系统而复杂过程,它是由评估得以运行所必须具备的基本要素及其之间的相互关系构成的^[3]。评估过程中遵循基本的原则,才能确保评估工作的科学性和评估结果的准确性,需贯穿于评估工作的全过程^[4]。

2.1 根据不同评估对象分类评估

生成式人工智能军事应用评估对象具有多样性、动态性和差异性的特点。因此,开展评估工作时必须从评估对象的实际情况出发,遵循其内在规律,研究和把握其特殊性,针对不同的评估对象,灵活采用不同的评估指标和评估方法,做到科学准确地评估^[5]。

2.2 突出对抗环境下的动态评估

动态评估是在构建近似实战的对抗环境,根据对抗进程及作战行动效果对生成式人工智能进行评估。未来智能化作战更加强调信息的融合和力量的联合,只有双方自主对抗,才能全面体现作战过程中的各种智能要素,全面反映生成式人工智能应用

于军事领域的作战能力^[6]。因此,要突出对抗环境下的动态评估。

2.3 坚持定性定量相结合评估

生成式人工智能应用于军事领域是一个复杂过程,其评估指标种类繁多、来源多样,既包括可准确量化的指标,也包括一些难以量化的指标。因此,单一的评估方式不能满足实际的评估需求,而应将定性评估与定量评估相结合,对生成式人工智能军事应用进行系统且全面的综合性评估,才能够得出科学、准确的评估结果^[7]。

2.4 采取全面又突出重点的评估

生成式人工智能有望赋能态势感知、指挥决策、精准打击、机动防护、后勤保障等各方面,形成具有倍增效应的体系作战能力^[8]。因此,对生成式人工智能军事应用进行评估需全面考虑体系作战能力的各组成要素,建立相对完整的评估内容体系和对应的指标体系。全面评估的同时还应重点突出生成式人工智能在军事领域的适用性、可靠性、稳定性等问题,采取全面而又有重点的评估。

3 生成式人工智能军事应用评估指标及方法

建立评估指标体系是评估工作的基础,指标选取不同,评估的结果可能大相径庭。建立科学、合理的评估指标体系,选择合适的评估方法,才能获得科学有效的评估结果。

3.1 评估指标体系

预先研究生成式人工智能军事应用评估问题,构建适应智能化作战条件下新质战斗力生成的指标体系,要着眼于未来智能化作战体系应具备的基本能力和基本特点,着眼于其与军事应用的契合程度、能力提升程度以及安全性等问题,在遵循指标选取原则下,通过专家咨询法确定指标集^[9]。笔者构建的生成式人工智能军事应用评估指标体系如图 1 所示。

3.2 评估方法

生成式人工智能军事应用评估方法是指对生成式人工智能技术应用于军事领域的适用性、不确定性、有效性等进行评估所采用的评估模式和运算程序等手段的总称。评估对象和评估目的不同,评估方法也不尽相同,主要采用以下 5 类评估方法。

3.2.1 人工评估

通过军事专家对生成式人工智能的输出内容进

行评价，如就对话体验、自然度、专业度、准确性等进行打分，对打分结果进行统计分析，准确评价出生成式人工智能系统是否适合该军事场景并就军事问题做出符合目的的、有效的、高质量的输出。

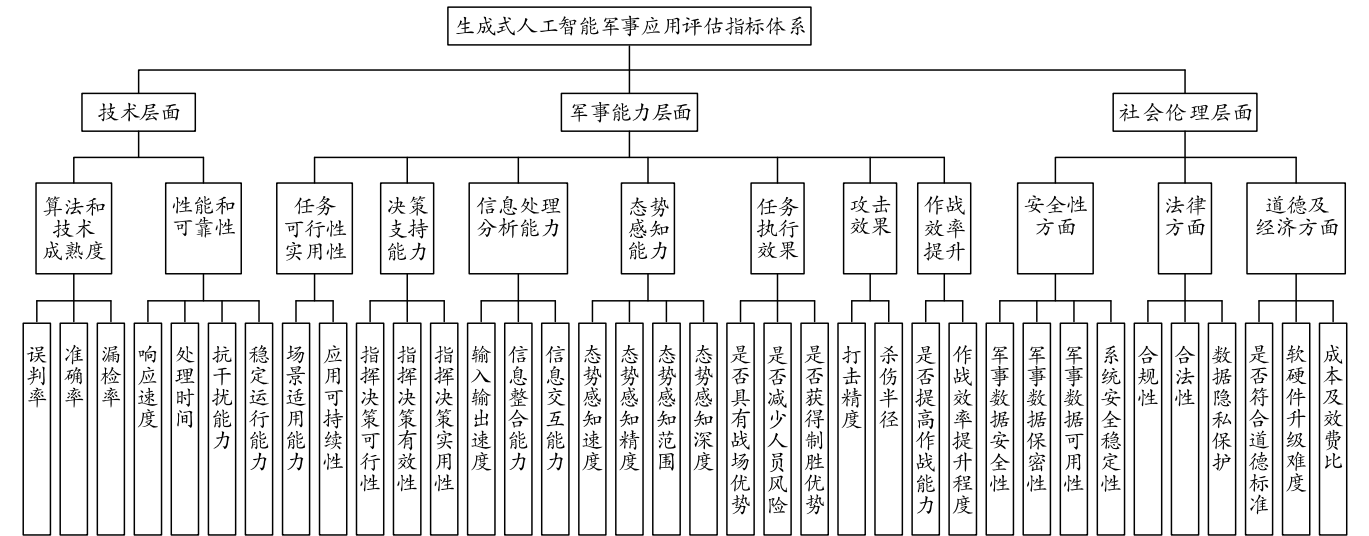


图 1 生成式人工智能军事应用评估指标体系

3.2.2 自动评估

在特定的军事场景下设定评价指标，如 BLEU、ROUGE、METEOR、Perplexity 等，评估系统自动将生成式人工智能应用于军事领域后对应的指标值进行输出，根据指标结果进一步判断生成式人工智能在该军事场景下是否符合需求以及作用效果。

3.2.3 对比评估

包括横向对比评估和纵向对比评估。横向对比评估是在生成式人工智能技术加持下，通过与潜在对手各指标的比较来评估自身作战能力。纵向对比评估是通过与过去的的能力状态和未来预期的能力状态进行对比来评估应用生成式人工智能技术后的能力变化，查找发现能力缺陷，进一步调整优化完善^[10]。

3.2.4 实验评估

通过构造逼真的作战环境进行自主对抗，如在实兵演习演练过程中将生成式人工智能技术应用于对抗过程，对关键指标进行测试和验证，评估其在实际军事场景中的适用性和效果。

3.2.5 推演评估

利用大型软件建模技术进行作战推演仿真，通过构造典型智能作战场景进行红蓝双方对抗，由推演仿真得到作战数据，再经过数据统计、分析和处理得出生成式人工智能军事应用评估所需的各项指标值。

生成式人工智能军事应用评估有其自身的特点，合理选择评估方法至关重要。但无论采用何种

评估方法，适用性是关键。只要能够在评估目的的指引下，完成评估任务，发现能力缺陷和存在问题，提出解决、优化和完善的措施与建议即可。同时，生成式人工智能应用于军事领域是一个全新的课题，其评估方法有待进一步探索、发展和完善。

4 生成式人工智能军事应用评估问题的思考

精准把握生成式人工智能与军事领域的结合点和发力点，洞悉我军军事智能化所处方位，是推进生成式人工智能全面应用于军事领域的重要基础。研究生成式人工智能军事应用评估问题，需要着力做好以下工作。

4.1 明确评估作用，厘清评估思路

预先研究生成式人工智能军事应用评估问题是推动军事智能化发展的应有之义，也是构建智能化作战体系的重要一步。由于评估对象具有特殊性、评估数据具有保密性和敏感性，且评估结论对未来生成式人工智能与军事应用的结合点和释能方式起到关键作用，因此，必须做到正确认识生成式人工智能军事应用评估的重要作用，精准把握评估过程，客观分析评估的结果。在此基础上进一步优化评估流程，为生成式人工智能与军事应用结合指引正确的方向和发展路径。

4.2 重视评估工作，建立评估试点

生成式人工智能军事应用通常涉及到高度复杂的军事行动和决策，可能会对人类安全和世界秩序产生重大影响。如果应用的算法和模型不正确或不

完善,可能导致错误的决策或操作,甚至引发意外的冲突和人员伤亡。评估这些应用可以帮助防止或减少此类问题的发生和减少由此带来的负面影响。因此,必须重视生成式人工智能军事应用评估工作,并尽早成立评估机构,建立评估试点,使我军在生成式人工智能军事应用领域走在前列。

4.3 确立评估机制,构建评估团队

建立生成式人工智能评估机制是军事智能化发展从粗犷化向精细化发展的现实需要,是全面落实二十大关于掌握智能化战争特点规律要求的具体表现,有助于全面把握生成式人工智能在军事领域的适用性及作用发挥情况。且生成式人工智能军事应用评估具有专业要求高、体系层次多、方法要求杂、结果处理难等各种问题,要求评估者具有丰富的专业知识储备和实践经验,同时还要具备高屋建瓴的战略思维,因此,亟需建立一支优势互补、梯次合理的专业化评估团队。

4.4 探索评估方法,统一数据标准

探索生成式人工智能军事应用评估方法,是全面准确掌握生成式人工智能在军事领域的适用程度以及作用发挥情况的基础。所提评估方法能实现评估目的,但也存在一定程度的局限性,要继续探索体系化的生成式人工智能军事应用评估方法。同时,要解决不同军种或部门数据格式和存储方式不统一的问题,建立统一的数据格式标准,搭建共享数据平台,为不同军种或部门提供共享数据的接口和服务,提高数据资源利用率,便于评估工作高效开展。

5 结束语

笔者思考生成式人工智能军事应用评估的重要意义、评估要遵守的基本原则、评估的指标体系以及评估方法问题,有利于生成式人工智能与军事应用紧密结合,加快技术落地,进而促进生成式人工智能在军事领域高效发展。

参考文献:

- [1] 瞿崇晓,郑寄平,张永晋,等. GPT 技术原理及其潜在军事应用研究[J]. 中国电子科学研究院学报, 2023, 18(7): 624-633.
- [2] 周中元. ChatGPT 对军事安全的挑战与应对策略[J]. 国防科技工业, 2023(7): 46-48.
- [3] 胡晓峰,杨镜宇,张明智,等. 战争复杂体系能力分析与评估研究[M]. 北京: 科学出版社, 2017: 3-4.
- [4] 毛翔,褚睿,邢蓬宇. 美军作战评估理论与实践[M]. 北京: 知识产权出版社, 2019: 29-31.
- [5] 杨满喜. 体系作战能力评估基本问题研究[M]. 北京: 国防大学出版社, 2016: 30-32.
- [6] 何榕,常壮,于皓. 作战体系评估模型构建研究[J]. 指挥控制与仿真, 2021, 43(5): 29-33.
- [7] 闫雪飞,李新明,刘东. 武器装备体系评估技术与研究[J]. 火力与指挥控制, 2016, 41(1): 7-10.
- [8] 何榕,王大旌,崔帅豪. 基于复杂网络理论的网络化作战体系结构效能评估[J]. 兵工自动化, 2021, 40(1): 43-49.
- [9] 马宝林,刘德胜. 作战体系评估指标体系结构构建方法[J]. 指挥控制与仿真, 2021, 43(2): 50-56.
- [10] 刘德胜,付东. 作战体系评估及评估方法研究[J]. 军事运筹与系统工程, 2018, 32(3): 14-17.