

因果推断及其在国防领域的应用研究^{*}

——以美国DARPA因果推断智能项目为例

任前前¹ 白如江¹ 陈鑫¹ 刘睿琳²

(1. 山东理工大学信息管理学院, 淄博 255049; 2. 多伦多大学信息学院, 多伦多 M5S3G6)

摘要: 受科技迅猛发展、作战环境变化等因素的影响, 国防建设过程面临前所未有的复杂性, 如何辅助决策者更准确地理解复杂系统要素之间的因果逻辑, 进一步巩固和提升国防建设水平是国防领域的重要议题之一。在阐述鲁宾因果模型和结构因果模型等因果知识的基础上, 以美国国防高级研究计划局 (Defense Advanced Research Projects Agency, DARPA) 实施的一系列因果推断智能项目为例, 分析因果推断在国防建设中的重要价值与具体应用场景。结合因果推断智能项目的趋势分析, 探讨对我国国防建设的启示与建议。研究表明, 因果推断在提升国防决策准确性与可解释性方面具有独特优势, 未来应正视并积极引入因果推断, 重视数据资源建设, 提升技术智能化水平, 加强监督管理, 加快因果推断在国防建设中的应用。

关键词: 因果推断; 国防建设; 美国; 国防高级研究计划局; 鲁宾因果模型; 结构因果模型

中图分类号: G353.1 **DOI:** 10.3772/j.issn.1673-2286.2024.10.003

引文格式: 任前前, 白如江, 陈鑫, 等. 因果推断及其在国防领域的应用研究: 以美国DARPA因果推断智能项目为例[J]. 数字图书馆论坛, 2024, 20(10): 22-32.

国防建设关乎国家安全、社会稳定与经济发展, 其决策的正确性和有效性是国家生存与发展的保障。大数据、人工智能等智能化要素逐渐向国防决策领域渗透, 为决策者处理国防事务提供了全面、准确、及时的数据和信息支持, 从而帮助作出更明智、更高效的判断。在未来的智能化战争中, 智能化决策技术可通过构建场景模型规则, 结合机器学习、神经网络等技术谋局布势, 以机器智能拓展决策者智慧, 使决策人员从繁琐的情报信息甄别与处理的工作中解脱出来, 形成决策优势, 成为提升未来智能化战斗力的新质增长点^[1]。

国防智能决策过程面临着无可比拟的复杂性, 主要体现在两个方面: 一是国防智能决策技术涉及多部门、多领域、多层次合作及多技术集成应用, 本身具有复杂性; 二是随着未来战场新型武器系统、航空航天新技术、高精度打击武器等高端、复杂技术的应用, 智

能决策系统需要不断提高正确应对复杂安全威胁的能力。因此, 提升国防智能决策的正确性、可信任度和可解释性是未来一段时间内的关键任务。

因果推断的发展为应对上述复杂性带来了契机。这种基于溯因的决策支持不仅有助于揭示复杂系统中各种要素之间的因果关系, 还能够深入了解其影响原理和机制, 找出影响资源配置和技术性能的关键因素, 从而有助于制定科学合理的资源配置方案, 提高国防智能决策的准确性和可解释性。本研究在阐明因果推断概念内涵和框架的基础上, 选取美国国防高级研究计划局 (Defense Advanced Research Projects Agency, DARPA) 实施的融入因果推断思想或技术方法的项目案例, 分析因果推断在国防建设过程中的具体应用, 归纳因果推断智能项目的发展趋势, 并提出对我国国防建设领域开展因果推断研究、应用因果推断技术的启示与

收稿日期: 2024-07-04

^{*}本研究得到山东省哲学社会科学青年人才团队“面向国家安全的多源数据智慧情报感知研究”(编号: 2023-zkzd-061)资助。

建议, 以期为巩固和推进我国国防建设提供参考。

1 因果推断的理论基础

1.1 因果推断相关概念

因果关系往往容易与相关关系混淆, 但实际上两者有本质区别。相关关系是指总是共同发生的两事件之间的关系, 而因果关系是“因”事件和“果”事件之间客观存在的关系, 其中“因”事件是“果”事件发生的原因, 且不以人的意志为转移^[2]。二者的一个显著区别就是因果关系具有方向性, 即有了“前因”才会产生“后果”。

因果推断也称为因果推理, Savage^[3]发表在《Nature》中的“Why artificial intelligence needs to understand consequences”(《为什么人工智能需要理解因果》)将因果推断定义为用数学的形式化方法来确定一个变量是否影响另一个变量。运用因果推断的最终目的在于去除混杂的伪因果关系, 确定研究变量间真正的因果关系。其主要解决两个方面的问题: 一是因果关系发现, 即发现或识别研究变量之间是否存在因果关系; 二是因果效应评估, 即对原因未施加和施加干预产生的结果之间差异的评估, 评估因果效应可以提升“因事件”与“果事件”之间内在运行机制的透明度。

作为机器学习领域的重要分支, 可解释机器学习目的在于解决模型的“黑盒”问题, 增强用户对模型的信任^[4]。然而, 尽管其在提升模型透明度方面表现出色, 但在挖掘数据背后的深层次机制和根本原因时显得无能为力。相比之下, 因果推断更侧重理解因果机制, 深层次揭示系统变量之间的因果关系。因此, 因果推断为科学研究提供了一个不同于传统变量关联分析的新视角, 借助因果推断的理论框架及方法等, 可以为决策者作出更稳健的结论和预测提供支撑。

1.2 因果推断理论框架

因果推断的实践核心在于敏锐地发现并捕捉到数据中因果关系的线索, 并探索出能够模拟这种捕捉能力的计算模型, 较为著名的因果推断理论框架主要有潜在结果框架(Potential Outcome Framework, POF)和结构因果模型(Structural Causal Model, SCM)两种。

(1) 潜在结果框架。潜在结果框架由美国国家科学院院士Donald Rubin提出, 因此又称为鲁宾因果模型(Rubin Causal Model, RCM)。此模型旨在解决因果效应评估难题, 通过深入学习和理解因果效应, 设想与现有观测相悖的情境, 估计不同干预措施下的潜在结果, 以帮助决策者作出最为恰当的决策。Yao等^[5]在关于鲁宾因果模型因果推断方法和机器学习问题的综述中指出, 鲁宾因果模型的基础是因果关系作用于一个个体, 通过比较各个体产生的观察结果(事实结果)与潜在结果(反事实的结果)获得最终效果。

(2) 结构因果模型。图灵奖得主Judea Pearl认为当下正在进行一场以科学为中心的涉及科学归因与公平性的改变数据科学的新革命, 即因果革命。对此, Judea Pearl根据其在因果领域多年探索的经验, 提出了一个新颖的科学方法框架, 即结构因果模型^[6]。该模型结合了经济学和社会科学中结构方程模型的特征和鲁宾因果模型, 以及为概率推理和因果分析开发的图形模型^[7]。结构因果模型开发一种数学语言来捕捉因果关系, 通过构建因果图直观地表达出研究变量间的因果关系及因果路径, 使数据有了更加合理的特征表示。

鲁宾因果模型和结构因果模型框架对比如表1所示。虽然鲁宾因果模型与结构因果模型的研究视角不同, 主要解决问题以及表现形式也不相同, 但是二者在本质上是等价的。Judea Pearl在其著作《因果论: 模型、推理和推断》中对两个框架的等价性进行了介绍, 但也强调这并不意味着概念上或实际应用中的等价性^[8]。

表1 鲁宾因果模型与结构因果模型框架对比

要素	鲁宾因果模型	结构因果模型
提出者	Donald Rubin	Judea Pearl
主要解决问题	因果效应评估问题	因果关系发现问题
表述问题的语言	定义为反事实独立性的假设(可忽略性假设) ^[9]	定义为因果图
框架展现形式	形式化数学公式	因果图
优点	淡化因果关系, 聚焦因果效应评估	透明度、功效性、可检验性
缺点	将假设表述为包含反事实变量的条件独立关系, 较难理解	存在多研究变量和复杂关系时, 模型构建具有挑战性

2 因果推断在情报分析中的应用

因果推断涉及海量的数据和信息, 包括经济、军

事、政治等多方面的情报。这些数据和信息往往具有高度的异构性和不确定性，对情报分析提出了更高的要求。因果推断作为一种能够揭示变量之间因果关系的科学方法，已经成为推动情报分析工作现代化、智能化发展的重要力量，具有广泛的应用前景，其应用主要可归纳为以下几类。

(1) 异常检测与预警。异常检测是情报分析中的一个重要环节。情报分析人员的工作任务之一就是找出存在的风险或威胁点，及时拉响警报，为决策者提供具有前瞻性的建议或意见。借助因果推断方法，特别是通过构建结构因果模型，可识别出偏离正常关系模式的点，这些异常点往往代表着潜在的风险或威胁。目前，因果推断已被应用于多个领域的安全性检测与诊断中，如Zhou等^[10]开发了名为CausalKGPT的工业结构因果知识图谱增强的大语言模型，该模型相较于经典商业模型，在航空航天产品缺陷原因分析方面表现出更高的专业性和可靠性。Pu等^[11]提出基于结构因果模型的决策级融合方法，利用因果图保证融合的因果方向，利用结构方程实现融合目标、提高故障检出率，该方法在各单一系统状态的故障检出率和平均故障检出率上均有较好的表现。Li等^[12]在结构因果模型框架下，提出一种基于梯度提升决策树的结构化因果模型，利用GBDT机器学习方法提高模型在结构安全异常检测任务中的鲁棒性，实验表明该模型构建过程合理，能够在不同异常场景下保持良好的异常检测性能。

(2) 事件关联与模式识别。在情报分析中，如何将多个变量进行关联，进一步探究变量之间的因果关系是亟待研究的问题。因果推断能够从复杂的数据中抽取出事件的因果关系链，进而识别出特定事件或行为模式，揭示隐藏在数据背后的深层次信息，为未来情报预测提供基础。例如，Che等^[13]通过双向双样本孟德尔随机化分析发现胃癌与恶性贫血之间存在反向因果关系，而恶性贫血与前列腺癌之间则存在隐含的反向因果关系。Lee等^[14]基于鲁宾因果模型和贝叶斯算法开发了C-search算法，可以从各种遗传数据集中发现致病基因（因果基因），基于致病基因治疗患者可产生较好的效果。在军事决策过程中，战场态势的不完备性和模糊性对经典意图识别技术的有效性和因果关系分析提出了挑战。Zhang等^[15]将鲁宾因果模型与不确定性因果推理相结合，提出了一种目标意图因果分析范式，完成了证据结构下的空中目标意图识别，并进一步分析了样本属性对目标意图的因果影响。

(3) 因果效应评估与策略优化。作为数据分析和决策制定中的重要理论方法，因果推断通过量化数据评估因果效应，为进一步优化干预策略提供了证据。

McJames等^[16]采用因果推断概率框架，发现每日作业对数学成绩的益处最大，而每周3~4次的家庭作业对学习成绩提升最有效。依据该信息，教育工作者和政策制定者可科学把控并优化各学科家庭作业布置的频率，发挥家庭作业对学科成绩提升的正向作用。Zhang等^[17]采用固定效应模型逆方差加权方法，发现较高水平的超氧化物歧化酶和锌是出生体重增加的原因，怀孕期间补充抗氧化剂可以防止低出生体重婴儿出现。Li等^[18]利用孟德尔随机化分析方法发现，遗传驱动的阻塞性睡眠呼吸暂停是心房颤动产生的原因之一，该发现为采取有效措施预防和治疗阻塞性睡眠呼吸暂停以降低心房颤动产生风险提供了新的思路。

值得注意的是，除上述因果框架方法，还有许多学者也对因果推断进行了前沿性的探索。例如：Du等^[19]提出了因果引导的主动学习框架，能够自动且自主地识别信息性偏差样本并引导偏差模式；Liu等^[20]设计了一个迭代学习与识别框架，用于识别事件的因果性；Ma等^[21]提出了基于Vicinal图的学习算法，显著提高了ML4S模型的预测精度，使其能够构建准确的因果骨架。这些因果框架和方法对于我国国防领域因果推断研究工作具有重要指导和参考意义。

3 DARPA因果推断项目介绍

DARPA隶属于美国国防部，是美国国防技术优势的发源地，主要负责引领并推动美国军事领域高科技的研发工作。DARPA的科研重点涵盖多个领域，包括但不限于先进武器系统、通信技术、人工智能等，前瞻性的科研探索和灵活高效的运作模式使其在维持美国军事技术领先方面发挥了关键作用，进一步巩固了美国在全球军事科技领域的领军地位。本研究选取2014—2024年DARPA实施的因果推断智能项目作为案例，分析因果推断在其中的具体应用与潜在价值，为将因果推断应用于我国国防建设提供借鉴。

3.1 大机制项目

大机制项目（Big Mechanism Program, BMP）旨在构

建揭示复杂系统(如生态系统、大脑、经济和社会系统等)内部交互作用因果关系的解释模型^[22]。其目标是通过阅读大量文献, 赋予机器常识推理能力, 提取文章及其摘要中的因果片段从而自动组装成更完整的因果模型, 并对这些模型进行推理以产生解释。尽管该项目围绕癌症生物学领域, 重点关注癌症信号传递, 但其首要目标是为美国国防部中重要的复杂系统构建完整的解释模型, 使数据和知识自动或半自动地集成到因果解释模型中, 癌症通路仅是

因果解释模型的一个应用方面。该项目主要划分为3个阶段: ①利用机器阅读相关领域期刊文章, 提取有用信息即因果关系片段, 转化为机器能够理解的形式; ②将知识整合至癌症通路的计算模型中, 即装配大机制, 并不断完善该机制, 在模拟环境中进行潜在结果评估, 这体现了鲁宾因果模型的思想; ③利用较为完备的大机制系统, 实现可通过实验检验的解释、推理、预测等。具体大机制系统如图1^[22]所示。

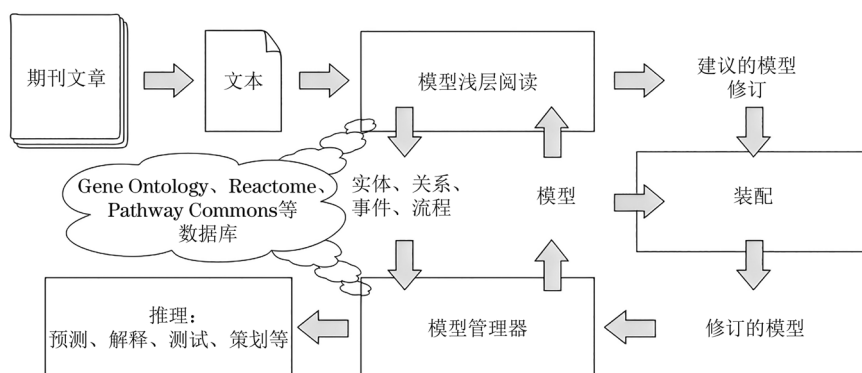


图1 大机制系统的粗略架构

3.2 复杂作战环境中的因果探索项目

随着现代军事科技的创新突破, 复杂的作战环境对军队和联盟伙伴及其行动提出了更高的要求, 亟须获得对作战环境的深度了解, 尤其是违反直觉的因果动态。同时, 为了摆脱数据库更新的滞后性以及数据库工具的严重依赖性, DARPA启动了因果探索(Causal Exploration, CausX)项目。该项目的核心目标是研发一个用于协助军事规划者深入剖析并应对导致复杂冲突局势的潜在原因的建模平台, 主要应用于以区域混

合冲突为例的场景挖掘, 着重解决由复杂因素交织引发的混合或非常态的冲突, 并能够在使用时根据作战环境迅速构建和维护因果模型。建立作战环境模型、分析敌我双方行动间的因果关系等是结构因果模型在该项目中的应用体现, 而评估不同作战方案潜在结果、提供最有效的打击方案则展示出了鲁宾因果模型的指导作用。借助建模平台可以探究影响作战环境的各类因素, 帮助军事规划者梳理作战脉络, 科学制定军事决策, 全面把握作战环境的复杂性。该项目实施流程如图2^[23]所示。

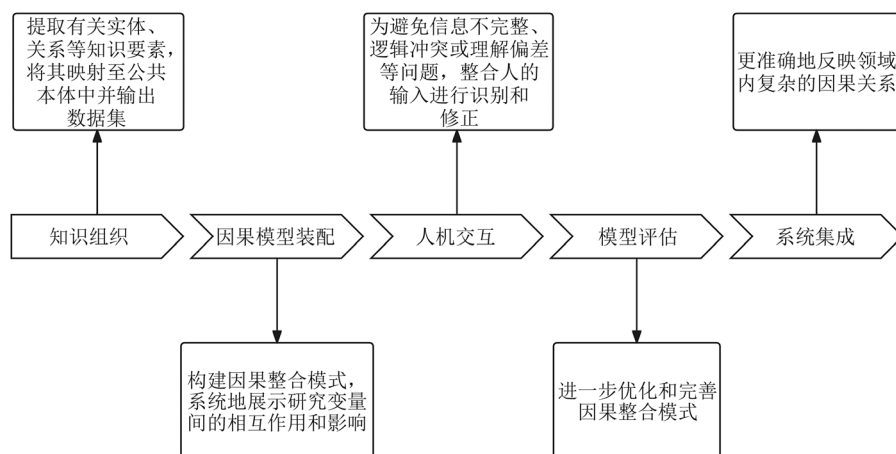


图2 CausX项目实施阶段

3.3 可解释人工智能项目

尽管人工智能研究已经取得举世瞩目的突破,但因许多技术应用包含的算法、模型以及系统存在着决策逻辑透明度低、结果的可解释性弱的问题,人工智能技术及其系统在国防、金融、医疗等关键决策领域的应用受到了极大的限制。可解释人工智能(Explainable Artificial Intelligence, XAI)项目的目标在于创建一套新的或改进的机器学习技术,生成可解释的模型,当

模型与有效的解释技术相结合时,用户能够理解、适当信任和有效管理新一代人工智能系统,其计划架构如图3^[24]所示。该项目利用图模型等推理出决策过程中的因果逻辑,评估模型框架的有效性,充分体现了因果推断框架的结合应用。XAI项目聚焦两个技术领域:①可解释的学习者,主要是指开发原型可解释学习系统,包括可解释模型和解释界面;②解释的心理学模型,以总结当前心理学解释理论、开发用于解释的计算模型为主要工作内容^[25]。

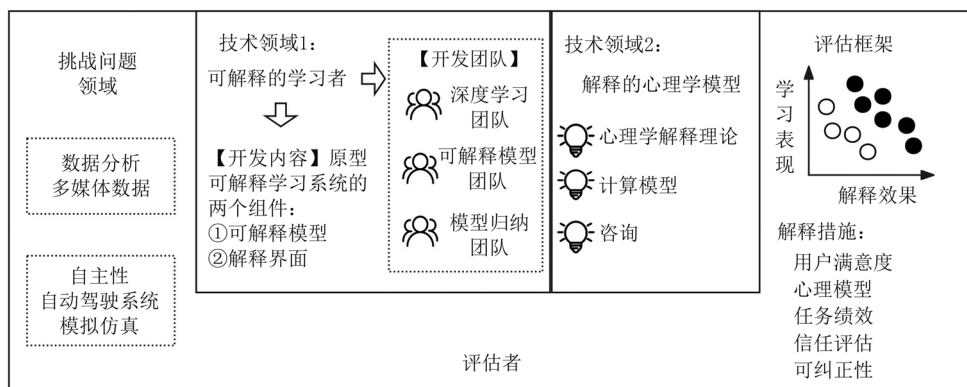


图3 XAI项目架构

3.4 知识导向人工智能推理图谱项目

随着非结构化多媒体信息量的爆炸式增长,迅速理解和分析全球事件成为维护国家安全、把握局势动态、精准制定应对策略的信息基石。知识导向人工智能推理图谱(Knowledge-Directed Artificial Intelligence Reasoning over Schemas, KAIROS)项目旨在实现一

个知识导向的(基于模式的)人工智能系统,可以通过学习大量的时间信息和事件模式识别并展现不同事件间的相似性,进而识别复杂事件并引起用户注意。如图4^[26]所示,该项目第一阶段主要利用丰富的语言和事件数据,进行深入的语言学推断和常识推理检测,分类和集中子事件并以此为基础构建强大的模式库;第二阶段将创建的模式库应用于多媒体、多语言等新信息的处

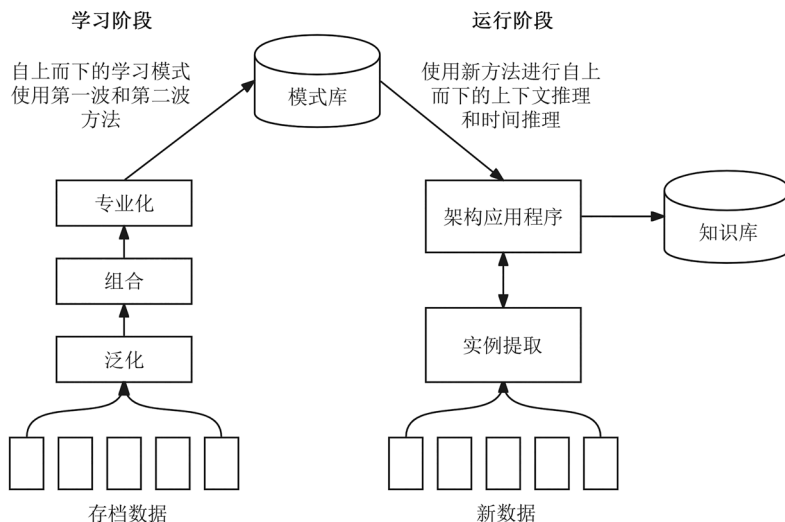


图4 KAIROS项目流程

理, 识别事件与实体, 致力于发现和提炼其中的复杂事件及关联, 进而构建或扩充知识库^[27]。KAIROS项目对因果推断框架的应用主要体现在人工智能知识图谱的构建与分析上, 结合时间变化对图谱进行分析, 可以高效

推理出事件的因果关系和动态变化, 支持正确决策。

CausX、XAI、KAIROS项目及BMP是DARPA在因果推断研究中的重要尝试, 对其项目目标、项目特点、目标对象和项目周期进行归纳, 结果如表2所示。

表2 因果推断项目内容归纳

要素	BMP	CausX项目	XAI项目	KAIROS项目
项目目标	开发方法, 能够自动识别和建构复杂系统中的大机制	创建平台, 理解导致复杂冲突局势的潜在因果因素	创建生成高度可解释的模型, 提升用户对人工智能系统的信任度并进行有效管理	构建知识导向的人工智能系统, 识别、追踪复杂事件并引起用户注意
项目特点	装配并不断完善大机制, 解释模型行为	允许用户根据当前作战环境快速构建和维护因果模型, 并将其与人类专家经验相结合	强调人工智能系统的可解释性、透明度和用户信任	以时序信息和事件模式为核心脉络, 赋予机器深入解析事件模式的能力
目标对象	期刊文章	地缘政治事件	新一代人工智能系统	社会复杂事件
项目周期	2014—2017年	2017—2021年	2017—2021年	2019—2024年

4 DARPA因果推断智能项目趋势分析

4.1 坚持问题导向原则, 强调前瞻性视角引领

在国防领域, 因果推断智能项目的实施紧密围绕国家安全和战略需求, 具有明确的研究目标和关注重点, 不仅需要解决当前的问题, 还要具有前瞻性视角, 引领未来发展方向。表3所示为DARPA 4个项目的研究问题、关注重点和前瞻性的研究角度。BMP专注于从知识发现角度探索海量文献数据中的大机制模式, 进而提升模型行为的可解释性。CausX项目主要从系统动力学的角度, 将作战环境看作一个动力驱动系统, 探索导致冲突的潜在因素, 如激进分子的攻击、政府的腐败、安全部队的滥用等, 以帮助军事规划者捕捉、应对和预

测复杂冲突局势。XAI项目从人机协作的角度, 研究如何使人工智能决策过程能够被人类理解, 重点关注可解释性和透明度的提升, 进而增强人类对人工智能系统的控制和信任。KAIROS项目前瞻性地从事件理解的角度, 自动抽取时间信息和事件模式, 构建社会复杂事件的知识图谱, 致力于深入解析复杂事件并反馈至用户, 引起关注。

4.2 数据驱动向知识导向转变, 聚焦技术集成与创新应用

随着大数据、人工智能、云计算等技术的迅猛发展, 情报工作愈发强调知识导向以及技术集成和创新应用的重要性。在分析上述4个因果推断智能项目的基础上, 从数据源和技术方法两个方面绘制因果推断智能项目数据源拓展与技术革新趋势图, 如图5所示。数据是训练精确且高效的因果模型的关键要素, 数据源的选择将会直接影响因果推断的结果。BMP收集癌症期刊文本数据进行浅层阅读和深度阅读以训练模型。CausX项目数据主要来自军事和情报文件、开源文本、政府或非政府组织数据等。XAI项目将异构的多媒体信息作为数据源, 从中发现感兴趣的事件。KAIROS项目以知识为导向, 依赖文本、演讲、图片、视频等多媒体、多语言数据构建知识图谱, 挖掘潜在规律。因果推断智能项目重视数据源的多样性和融合, 更强调将数据转化为可解释、可应用的知识体系, 呈现出由数据驱动转向知识导向的趋势。

表3 因果推断智能项目研究问题、关注重点和研究角度

类别	研究问题	关注重点	研究角度
BMP	如何理解和分析复杂生物系统中的机制	自动提取和构建生物机制模型	知识发现
CausX项目	导致冲突的潜在因素是什么	识别和量化复杂系统中的因果链	系统动力学
XAI项目	如何解释人工智能系统的决策过程	提高人工智能系统决策和推理的可解释性、可信度、透明度	人机协作
KAIROS项目	如何以自动化方式理解和构建复杂事件和行为序列	事件理解和事件链的自动构建	事件理解

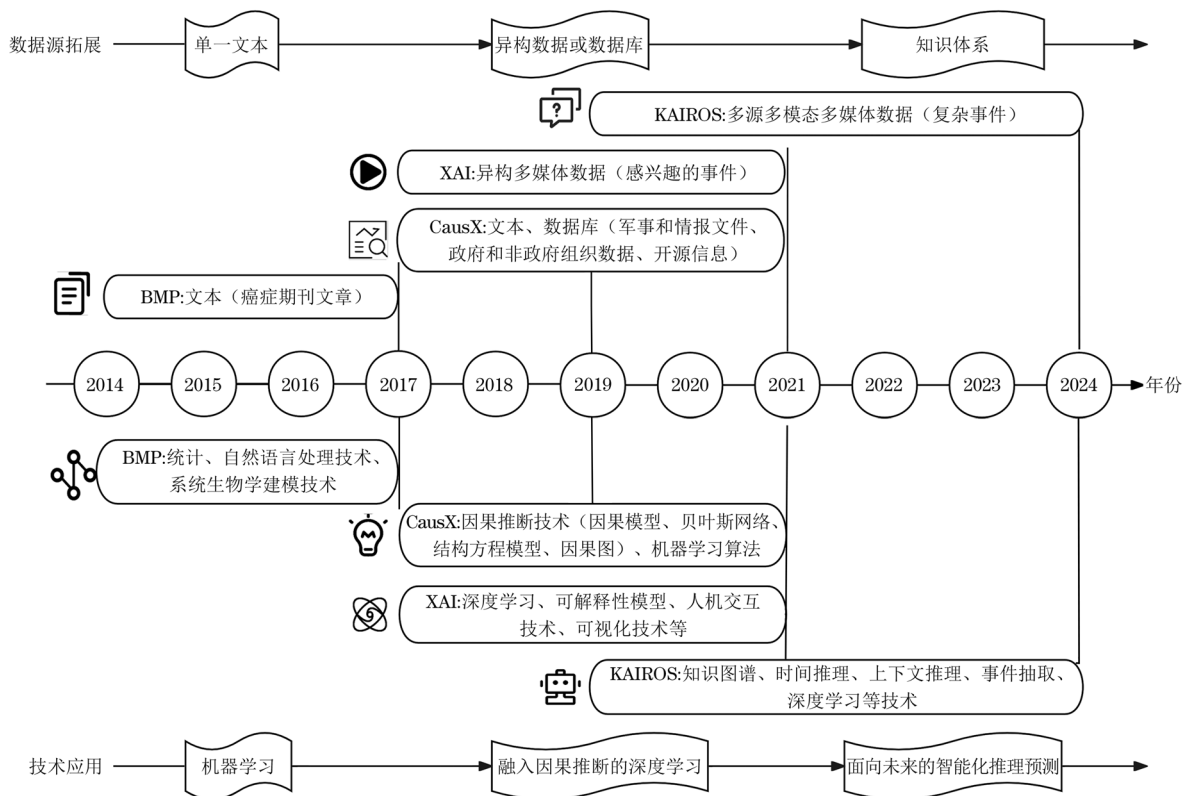


图5 数据源拓展与技术革新趋势

大数据和人工智能等前沿技术迅猛发展,为因果关系分析提供了极为丰富的数据源和强大的计算工具。BMP主要利用机器阅读、自然语言处理、生物学建模等技术构筑大机制,但对于因果推断技术的研究尚处在相对初级的阶段;CausX项目综合运用贝叶斯网络、结构方程模型等因果推断技术和机器学习算法,实现潜在因果因素的发现;XAI项目利用深度学习、人机交互等技术辅助决策,提升决策的可解释性与透明度;KAIROS项目将叙事结构引入复杂事件序列,并以结构化表示呈现给用户,与知识图谱、时间推理、上下文推理等技术相结合,可以感知并剖析多媒体信息中的复杂事件。依据目标任务,4个项目在技术上从单纯的机器学习技术研究逐渐向融合、应用和创新的方向发展,并且呈现出面向未来的智能化、自动化推理预测的趋势。

4.3 明确角色划分与责任担当,持续优化团队协作模式

角色划分和责任担当在团队协作中起到至关重要的作用,也是推动团队项目实施和提升成果质量的重要

保证。鉴于因果推断智能项目的复杂性与广泛的应用潜力,DARPA往往会将项目细分成多个聚焦的领域,每个细分领域被赋予独特的战略重要性。同时,有针对性地组建跨学科研究团队,汇聚各领域精英力量,攻坚克难,推动领域内的前沿探索与突破。表4所示为4个项目的研究机构和科技公司参与情况。

XAI项目在可解释的学习者领域分配了俄勒冈州立大学、帕洛阿尔托研究中心、雷神BBN公司、罗格斯大学等11个研究团队,在解释的心理学模型领域安排了佛罗里达人机认知研究所的团队攻关。每个团队各自承担具体的工作任务,注重团队间工作的配合,合力促成项目实施。在研究成果方面,查尔斯河实验室团队开发了深度强化学习策略的因果模型,通过回应反事实查询来实现解释增强训练,开发出了行人检测算法的精炼版本^[28];加州大学洛杉矶分校团队提出概念和反事实解释(Conceptual and Counterfactual Explanations, CoCoX)模型,利用故障线路这一新颖反事实解释取代了标准的基于注意力的解释,实现了图像的分类^[29];得克萨斯大学达拉斯分校团队利用可处理的概率模型,使用由逻辑、概率、有向图生成的全局解释,生成有关视

频数据中是否存在活动的解释,实现视频活动的识别等,推动因果推断在智能决策领域发挥强大功能^[30]。

4.4 精准对接项目需求,实现落地成果价值最大化

DARPA的因果推断项目紧密围绕实际问题 and 应用需求,根据项目目标和应用场景对项目功能及技术方法进行设计,并加快推进研究成果的实际应用转化。如表5所示,这些项目的研究成果大部分还处于试验和开发阶段,部分技术在特定领域进行应用和测试,进入早期应用阶段,但是已经在一些行业中展现出了巨大的潜力和实际效益,应用前景十分广阔。

上述项目中,XAI项目现已推出专为分析和自主应用程序而构建的开源的、可解释的人工智能工具包 `xaitk.org`,该工具包结合了一个可搜索的独立贡献存储库和一个更加集成的通用软件框架。虽然CausX项目

表4 研究团队参与情况

类别	研究阶段或技术领域	研究团队来源
BMP	阅读	卡内基梅隆大学、MITRE公司、佛罗里达人机认知研究所、南加州大学信息科学研究所等
	组装机制	
	生成解释	
CausX项目	知识组织	雷神BBN公司、美国科学情报研究所、Lum AI公司等
	因果模型装配	尚不明确
	人机交互	
	模型评估	
	系统集成	
XAI项目	可解释的学习者	俄勒冈州立大学、帕洛阿尔托研究中心、雷神BBN公司、罗格斯大学等
	解释的心理学模型	佛罗里达人机认知研究所
KAIROS项目	模式库构建	雷神BBN公司、伊利诺伊大学厄巴纳-香槟分校、Next Century科技公司等
	运行推理与预测	
	用户交互	
	训练数据生成评估	

表5 研究成果应用情况

类别	研究成果	应用场景	应用状态
BMP	大机制系统(自适应软件系统框架)	在癌症生物学等领域改变知识的组织及沟通方式,判断模型的生物学合理性,解释驱动复杂系统的要素,实现新科学模式等	已存档
CausX项目	Causal Exploration软件	军事国防、经济金融、公共卫生等	暂未完全投入使用
XAI项目	可解释的人工智能工具包库,现已推出xaitk.org工具包	情报分析、人工智能等	已投入使用
KAIROS项目		社会科学研究、情报分析等	正在开发

周期已经结束,开发的Causal Exploration软件暂时还未完全投入使用,但该软件已经能够挖掘并提取丰富且相互关联的行为,未来有望在军事国防、经济金融、公共卫生等领域发挥重要作用。BMP的自适应软件系统框架也将应用于新科学模式,改变知识的组织及沟通方式,解释驱动复杂系统的要素。KAIROS项目目前正处于积极的研发阶段,其进展令人瞩目,有望在世界事件动态理解领域树立新的里程碑。

5 启示与建议

美国DARPA在因果推断道路上不懈探索,为世界各国国防决策分析工作提供了参考借鉴。然而,不同国家之间在顶层设计上可能存在差异,因此在借鉴相关

经验时,需要充分考虑自身国情和建设需求。从成功经验中获取有效适用的策略和方法,从失败教训中洞察疏漏与不足并及时修正,对加快因果推断在情报分析中的应用具有重要意义。

5.1 正视因果推断作用,强化因果推断基础研究

国防建设智能化步伐加快,对国防智能决策系统提出了更高的可解释性要求。在国防情报分析系统中引入因果推断研究,可以识别影响国防决策的关键因素并揭示这些因素之间的相互作用和反馈机制,更清晰地描绘出因素间的因果关系和作用路径,从而增强决策的可解释性。从“是什么”跨越至“为什么”,是一个

由浅及深、从现象到本质的认知过程，在促进批判思维和创新思维发展的同时，提供了更广阔的视野和更深刻的洞察力。因此，必须正视因果推断的作用，明确因果事件中存在的因果逻辑而非仅限于简单的相关性分析。特别是情报机构，应认识到仅凭相关性分析并不足以支持复杂的决策过程，必须引入因果推断来提升分析的深度和准确性。

夯实因果推断基础研究根基，激发理论创新活力，提升技术方法先进性与实用性。首先，要加强因果推断基础理论研究，深入理解其概念、原理和方法论，支持科研机构 and 高校开展前沿性、探索性因果推断研究，巩固基础研究根基。其次，鼓励科研人员自由探索，敢于提出新理论、新方法、新技术，勇于挑战传统观念，同时建立激励机制，加强在因果推断研究领域的交流，激发理论创新活力。最后，利用政策支持、团队协作、资源整合等形成合力，共同推动因果研究在理论和实践方面迈向新高度。国防智能决策系统是一个高度动态化的复杂系统，因果推断基础研究是促进该系统正确决策的“证据底座”，只有深度把握系统内各变量间的因果逻辑关系，才能提升系统可解释性，进而提升国防智能化建设水平。

5.2 重视数据资源建设，促进知识传播与高效利用

数智时代，深度学习、人工智能等技术井喷式发展，工作模式正在朝向以数据为基础、以人工智能技术为驱动的全新形态演进^[31]。数据是辅助国防军事领域智能系统实现精准分析、决策、指挥、打击的关键要素之一，高质量训练数据背景下的智能决策系统能够提升部队的快速响应和精确打击能力^[32]。要想谋胜未来，就必须重视数据赋能的重要作用。在支撑决策方面，数据驱动摆脱了依赖经验和直觉而产生的结果主观性过高的限制，提供了更为科学合理的分析方法与结果。然而，数据更新的滞后性、数据迷雾、数据不完备以及数据缺失等数据质量问题可能会导致因果结论的失真或误导，使研究人员深陷“伪数据”泥潭，寸步难行。鉴于此，获取数据时应当加大资源和精力的投入，以确保数据的准确性和完整性。

加强数据及数据资源平台建设，促进知识传播与高效利用。一方面，高质量的数据是训练精准的因果推断模型的必要条件，通过整合多领域、多场景的数据，

建立高质量因果推断数据集，进一步实现数据的组织，构建数据资源体系，为用户提供安全可靠的数据支持。另一方面，公共数据资源的有序开发、开放和流通共享可以促进知识传播、提升数据利用率。借助该思想，构建国家层面的数据共享平台，完善数据共享机制，打破数据机构之间的壁垒与信息孤岛，实现跨机构、跨领域的的数据互联互通，形成全国多源统一的公共数据资源体系。但值得注意的是用于国防智能决策的数据往往涉及国家机密信息，在进行数据资源建设时必须考虑到数据的安全性、隐私保护以及特定业务需求的独特性，根据实际情况作出判定。

5.3 提升技术智能化水平，推动跨领域合作与研究成果转化

在复杂的情报环境中，因果关系往往被噪声数据掩盖，而目前缺乏有效的因果推断工具和方法，推断结果的可靠性被质疑，因此必须加强因果推断技术的集成与创新。首先，不同的模型和算法有着各自的特点与显著优势。例如，一些模型在捕捉数据间的线性关系方面尤为出色，而其他模型在洞察非线性关系方面能力卓越，通过技术集成，可以将不同优势的模型与算法进行有机融合与叠加，汇聚所长，从而提升因果推断的准确率与效率。其次，通过技术集成和创新可实现实时或动态的推理，缓解时间滞后和长期效应问题^[33-34]，满足高度动态化智能战场的时效性要求。最后，技术集成为跨领域的知识融合提供了新方法，通过技术集成，将不同领域知识有机结合，利用因果推断更专业、更高效地解决国防情报分析中的因果关系问题。

推动跨领域合作与交流，搭建开放共享的因果推断模型库，拓展因果推断应用领域及场景。因果推断研究涉及计算机科学、统计学、医学、军事等多个学科领域，因果推断任务的完成往往需要多领域学者的鼎力合作，因此，有必要加强跨学科、跨领域的合作与交流，创新不同学科领域交叉融合的途径。其中，搭建开放共享的因果推断模型库不失为一种能够有效促进因果推断交流与合作的方式，它可以为用户提供数据处理、模型训练、结果可视化等功能，能够使更多专家学者参与因果推断研究。同时，加强与产业界的合作，共同研发具有市场前景的因果推断产品和技术解决方案，加速因果推断技术成果向实际应用转化。

5.4 加强监督管理, 制定并严格执行道德伦理审查制度

为提高决策准确性与有效性, 因果推断研究密切结合智能技术工具。但无论是人工智能、大语言模型还是机器学习, 都主要依赖算法和数据来处理海量信息, 这就容易导致其只能按照预设的逻辑或规则进行处理, 缺乏人类所具备的价值观和道德伦理观念, 道德伦理问题明显。例如, 在推断过程中, 智能化技术或工具可能会无意放大或引入数据中的偏见, 导致对特定群体产生不公平的影响, 因此十分有必要加强道德伦理层面的监管。

严格执行伦理审查与监管措施, 确保合规性。必须重视道德伦理审查过程并施以监管措施, 为因果推断研究守牢健康、可持续发展底线。2023年, 中国科学院自动化研究所曾毅团队研发出人工智能伦理道德与社会规范的群智数据与知识平台“智善·如流”, 该平台允许公众参与刻画人类在不同情景下的道德认知和行为, 把脉、把关和引导人工智能的健康稳健发展^[35]。同样, 在利用因果推断推动情报分析和决策工作时, 有关部门也需要加强监督管理工作, 严格执行道德伦理审查制度, 防止因滥用、误用导致操纵舆论、传播虚假信息等危害。

参考文献

- [1] 新华社. 国家国防科工局: 务实推动国防科技工业高质量发展[EB/OL]. [2024-03-22]. <https://baijiahao.baidu.com/s?id=1774900541796301255&wfr=spider&for=pc>.
- [2] 蔡瑞初, 郝志峰. 大数据中的因果关系发现[M]. 北京: 科学出版社, 2018: 2.
- [3] SAVAGE N. Why artificial intelligence needs to understand consequences[EB/OL]. [2024-02-24]. <https://doi.org/10.1038/D41586-023-00577-1>.
- [4] 刘智锋, 王继民. 可解释机器学习在信息资源管理领域的应用研究综述[J]. 数据分析与知识发现, 2024, 8(1): 16-29.
- [5] YAO L Y, CHU Z X, LI S, et al. A survey on causal inference[J]. ACM Transactions on Knowledge Discovery from Data, 2021, 15(5): 74.
- [6] 珀尔, 麦肯齐. 为什么?[M]. 江生, 于华, 译. 北京: 中信出版社, 2019.
- [7] PEARL J. Causal inference in statistics: an overview[J]. Statistics Surveys, 2009, 3: 96-146.
- [8] 珀尔. 因果论: 模型、推理和推断[M]. 2版. 刘礼, 译. 北京: 机械工业出版社, 2022.
- [9] PEARL J. Causal inference: history, perspectives, adventures, and unification (an interview with Judea Pearl) [J]. Observational Studies, 2022, 8(2): 22-36.
- [10] ZHOU B, LI X Y, LIU T Y, et al. CausalKGPT: industrial structure causal knowledge-enhanced large language model for cause analysis of quality problems in aerospace product manufacturing[J]. Advanced Engineering Informatics, 2024, 59: 102333.
- [11] PU H Y, CHEN Z, LIU J, et al. Research on decision-level fusion method based on structural causal model in system-level fault detection and diagnosis[J]. Engineering Applications of Artificial Intelligence, 2023, 126: 107095.
- [12] LI H R, LIU X M, HUAI X F, et al. A causal inference method for canal safety anomaly detection based on structural causal model and GBDT[J]. LHB, 2023, 109(1): 2218844.
- [13] CHE B W, YUAN S L, ZHANG H Y, et al. Causal inference between pernicious anemia and cancers: a bidirectional two-sample Mendelian randomization analysis[J]. BMC Cancer, 2024, 24(1): 586.
- [14] LEE Y K, KIM J, SEO S W. Discovery of genes positively modulating treatment effect using potential outcome framework and Bayesian update[J]. BMC Medical Informatics and Decision Making, 2022, 22(1): 113.
- [15] ZHANG Y, HUANG F H, DENG X Y, et al. Air target intention recognition and causal effect analysis combining uncertainty information reasoning and potential outcome framework[J]. Chinese Journal of Aeronautics, 2024, 37(1): 287-299.
- [16] MCJAMES N, PARNELL A, O'SHEA A. Little and often: causal inference machine learning demonstrates the benefits of homework for improving achievement in mathematics and science[J]. Learning and Instruction, 2024, 93: 101968.
- [17] ZHANG Y P, WU M, WANG H H, et al. Genetic predicted causal inferences between antioxidants and birth weight[J]. Journal of Clinical Biochemistry and Nutrition, 2024, 75(1): 54-59.
- [18] LI Y L, LENG Y M, TANG H B, et al. Assessment of the causal effects of obstructive sleep apnea on atrial fibrillation: a Mendelian randomization study[J]. Frontiers in Cardiovascular Medicine, 2022, 9: 843681.
- [19] DU L, SUN Z H, DING X, et al. Causal-guided active learning

- for debiasing large language models[EB/OL]. [2024-09-08]. <https://doi.org/10.48550/arXiv.2408.12942>.
- [20] LIU C, XIANG W, WANG B. Identifying while learning for document event causality identification[EB/OL]. [2024-09-08]. <https://doi.org/10.48550/arXiv.2405.20608>.
- [21] MA P C, DING R, DAI H Y, et al. ML4S: learning causal skeleton from vicinal graphs[C]//Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington DC USA. New York: ACM Press, 2022: 1213-1223.
- [22] COHEN P R. DARPA's Big Mechanism program[J]. Physical Biology, 2015, 12 (4): 045008.
- [23] DARPA. Causal Exploration[EB/OL]. [2024-08-26]. <https://www.darpa.mil/attachments/CausalExplorationProposersDayPMBriefing20161219.pdf>.
- [24] GUNNING D, AHA D W. DARPA's Explainable Artificial Intelligence program[J]. AI Magazine, 2019, 40 (2): 44-58.
- [25] DARPA. Explainable Artificial Intelligence (XAI) [EB/OL]. [2024-09-10]. <https://www.darpa.mil/attachments/DARPA-BAA-16-53.pdf>.
- [26] DARPA. Knowledge-Directed Artificial Intelligence Reasoning over Schemas (KAIIROS) [EB/OL]. [2024-09-10]. <https://www.darpa.mil/program/knowledge-directed-artificial-intelligence-reasoning-over-schemas>.
- [27] 张慧, 张骁雄, 丁鲲, 等. 美军智能数据情报KAIIROS项目分析[J]. 指挥信息系统与技术, 2021, 12 (1): 45-49.
- [28] WITTY S, LEE J K, TOSCH E, et al. Measuring and characterizing generalization in deep reinforcement learning[J]. Applied AI Letters, 2021, 2 (4): e45.
- [29] AKULA A, WANG S, ZHU S C. CoCoX: generating conceptual and counterfactual explanations via fault-lines[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34 (3): 2594-2601.
- [30] ROY C, NOURANI M, HONEYCUTT D R, et al. Explainable activity recognition in videos: lessons learned[J]. Applied AI Letters, 2021, 2 (4): e59.
- [31] 张钦, 薛海丽, 梁昊光. 国家战略情报体系发展的国际态势、经验与启示[J]. 图书情报工作, 2024, 68 (5): 13-21.
- [32] 中工网. 为推动军事大数据发展注入源头活水[EB/OL]. [2024-04-17]. <https://www.worker.cn/c/2022-12-30/7687020.shtml>.
- [33] DU S Z, SONG G J, HONG H K. Collective causal inference with lag estimation[J]. Neurocomputing, 2019, 323: 299-310.
- [34] DU S Z, SONG G J, HAN L, et al. Temporal causal inference with time lag[J]. Neural Computation, 2018, 30 (1): 271-291.
- [35] 新京报. 中科院等发布人工智能伦理道德平台 用户可对大模型道德判断[EB/OL]. [2024-04-19]. <https://baijiahao.baidu.com/s?id=1770915881748162976&wfr=spider&for=pc>.

作者简介

任前前, 女, 硕士研究生, 研究方向: 科技情报分析与文本挖掘。

白如江, 男, 博士, 教授, 通信作者, 研究方向: 科技情报分析与文本挖掘, E-mail: brj@sdut.edu.cn。

陈鑫, 女, 硕士研究生, 研究方向: 科技情报分析与文本挖掘。

刘睿琳, 女, 硕士研究生, 研究方向: 科技情报分析与文本挖掘。

Causal Inference and Its Application in the Field of National Defense: Taking DARPA Causal Inference Intelligence Project in the United States as an Example

REN QianQian¹ BAI RuJiang¹ CHEN Xin¹ LIU RuiLin²

(1. School of Information Management, Shandong University of Technology, Zibo 255049, P. R. China;

2. Faculty of Informatics, University of Toronto, Toronto M5S3G6, Canada)

Abstract: Affected by the rapid development of technology and changes in the combat environment, the process of national defense construction is facing unprecedented complexity. How to assist decision-makers in more accurately understanding the causal logic between complex system elements, and further consolidate and improve the level of national defense construction is one of the important issues in the field of national defense. Based on the elaboration of causal knowledge such as Rubin causal model and structural causal model, this study takes a series of causal inference intelligence projects carried out by Defense Advanced Research Projects Agency (DARPA) in the United States as an example to analyze the important value and specific application scenarios of causal inference in national defense construction. Combining the trends of causal inference intelligence projects, this article explores the inspirations and suggestions for China's national defense construction. Research has shown that causal inference has unique advantages in improving the accuracy and interpretability of national defense decisions. In the future, it is necessary to face up to and actively introduce causal inference, attach importance to data resource construction, enhance the level of technological intelligence, strengthen supervision and management, and accelerate the application of causal inference in national defense construction.

Keywords: Causal Inference; National Defense Construction; America; Defense Advanced Research Projects Agency; Rubin Causal Model; Structural Causal Model

(责任编辑: 王玮)