

量化私募基金系列之四 ——浅析量化基金所涉计算机技术及相关知识产权 作者：车小燕 | 汪健 | 杨维熙

一. 序言

量化投资交易作为一种区别于主观投资的投资交易方式，借助了许多前沿的计算机技术。量化基金管理人，借助数理金融理论、统计学分析、计算机程序化等技术，将基金管理人的投资策略和理念进行模型和算法化。区别于主观投资以“投资研究人员”为中心的投资交易方式，量化投资交易具有严格的投资纪律性，能够有效规避主观因素的干扰。

通力律师在服务量化基金管理人的过程中，对如何围绕相关技术开展知识产权风险排查和保护布局积累了一定的心得。本文以量化投资交易可能涉及哪些计算机技术及涉及到哪些知识产权而展开，并在下一篇中围绕如何进行知识产权风险排查和保护好自身的研发成果作进一步的探讨。

二. 量化基金概述

量化基金所涉及到的计算机技术至少包括以下几类：

1. 数据处理技术

量化基金需要大量的各类型数据进行建模和分析，因此在数据获取、存储、清洗和处理方面需要使用各种数据处理技术，包括数据挖掘、数据清洗、数据仓库、数据流处理等。同时，除了各类型金融数据，一些社交媒体、新闻等另类数据也是作为构建量化策略模型的重要数据来源，在这个过程中会使用到自然语言处理技术。例如 Palantir Metropolis 就利用其在大数据处理上的优势，在多源数据、另类数据、事实数据等多类信息应用的基础上进行进一步的探究，从而提供精准的数据集成、信息管理和定量分析服务。

.....
如您需要了解我们的出版物，
请联系：

Publication@llinkslaw.com

2. 程序设计和算法

量化基金通常使用编程语言开发交易程序。常见的编程语言包括 Python、R、C++等。此外，量化基金也需要设计各种算法来优化交易执行、风险管理等。

3. 人工智能

“AI+金融”逐渐成为量化投资领域的一种趋势，人工智能的加入在一定程度上能够解决量化投资策略同质化等问题。近年来，越来越多的量化基金开始使用人工智能技术，通过深度学习、强化学习等方法来改进交易策略模型、观察市场走势等。有的基金采用纯量化模型策略，基于深度神经网络模型构建策略框架，动态捕捉市场机会，不断自主优化投资策略。如某些量化基金管理人的选股模型不仅从盈利和营业收入等基本面角度去评估上市公司的价值和成长性，同时充分考虑市场情绪等因素为标的合理定价。此外，有的量化基金管理人则是通过量化加上主观干预的方式，在风格/行业层方面融入主观分析和判断，将组合进行主动偏离和调整，以获取相对基准的超额收益表现。

4. 云计算

云计算是将计算服务和资源远程交付给客户端，数据在通常称为“云”的基础结构上从客户端远程存储和处理。由于量化基金需要处理大量数据和复杂计算，因此需要利用云计算技术来加快模型训练和实时决策。

三. 与量化基金相关的知识产权

1. 商业秘密

量化策略模型算法存在“黑箱”特性，也是量化基金管理人的核心竞争力所在。因此，量化基金管理人对于其策略模型算法存在严格的保密期待和要求。结合我们长期深耕知识产权领域以及对量化基金管理人的服务经验，我们认为量化基金管理人的策略模型算法，可以选择以商业秘密的路径来进行保护。

在全国首例将算法作为商业秘密保护的案件中¹，法院认为虽然案涉算法的每个分部技术中所采用的模型均为公开模型，但技术模型选择及权重排序，权利人采取了相应的保密措施，并能为权利人带来商业收益和可保持的竞争优势，应当作为商业秘密予以保护。

值得关注的是，根据《关于规范金融机构资产管理业务的指导意见》第二十三条的规定，金融机构应当向金融监督管理部门报备人工智能模型的主要参数以及资产配置的主要逻辑。一方面，

¹ 深圳市智某信息技术有限公司与光某蜗牛(深圳)智能有限公司侵犯商业秘密纠纷案——深圳中院发布十件涉数字经济知识产权司法保护创新案例之一

量化策略模型算法属于量化基金管理人的核心资产，加之量化策略模型算法存在趋同或同质化的固有风险，因此量化策略模型算法的严格保密对于量化基金管理人至关重要。但量化基金本身具有受监管的金融业务属性，尤其量化投资交易在特定市场环境下可能存在加大市场波动等风险，金融监管机构要求量化机构进行信息报备确为必要。如何平衡监管信息报备需要及商业保密的合理诉求，相信将是长期的命题。但可以判断的是，金融监管机构要求量化基金管理人的人工智能模型进行报备，其目的不在于要求量化基金管理人突破保密的需要，而更多是从摸清量化交易底数、提升市场透明度和交易监管精准度等方面考虑，其目的在于维护市场交易秩序。而且，结合目前已经发布的程序化交易的信息报备要求，主要集中在量化的主辅策略方面，应不存在泄露商业秘密的风险。

2. 著作权

国家网信办等七部门联合发布的《生成式人工智能服务管理暂行办法》第四条第三项以及第七条第二项²规定了生成式人工智能在运行过程中不得侵犯他人依法享有的知识产权。由于人工智能模型训练需要多个来源数据库，如源自互联网、公共数据库、个人创作等，在算法和模型训练阶段，人工智能训练数据可能存在输入端的侵权的风险。虽然企业一般会在使用数据前进行数据清洗，删除或替换可能涉及侵犯著作权的内容，但仍存在使用受著作权保护的数据进行训练的侵权风险。例如程序员兼律师 Matthew Butterick 与 Joseph Saveri 律师事务所的法律团队合作，就专有 AI 工具 GitHub Copilot，对 GitHub 及其母公司微软、以及 OpenAI 发起集体诉讼，索赔 90 亿美元，Matthew Butterick 认为，GitHub Copilot 大模型在生成代码时使用了 GitHub 上开源代码的代码片段，但未经原作者的许可，构成侵犯著作权，GitHub 对此表示否认，称 Copilot 在使用 GitHub 上的开源代码进行训练时，只会使用公共领域的代码，不会使用任何受著作权保护的代码，该案件目前仍在审理。

3. 专利权

在专利方面，人工智能应用或算法是否能作为可授予专利权的计算机程序或软件，以及该专利客体审查规则如何细化一直备受关注。根据《专利审查指南》的相关规定，首先需要判断权利要求是否属于专利法第二十五条第一款第(二)项³规定的智力活动的规则和方法，如果仅仅是抽象的算法或商业规则和方法，则不会被授予专利权。其次，需要判断权利要求是否属于专利法第二条第二款⁴规定的技术方案，判断标准是该技术方案是否利用自然规律来解决技术问题并获得符合自然规律的技术效果，因此，在撰写该类专利时，不能够仅仅针对算法本身的计算过程进行描述，还需要与计算机实体技术特征相结合，例如可以增加对计算机执行算法指令的过程进行描述，比如包括计算机的处理器、存储器等组件执行数据读取、加载、运算、储存等指令的过程。

² 《生成式人工智能服务管理暂行办法》第四条提供和使用生成式人工智能服务，应当遵守法律、行政法规，尊重社会公德和伦理道德，遵守以下规定：(三)尊重知识产权、商业道德，保守商业秘密，不得利用算法、数据、平台等优势，实施垄断和不正当竞争行为。第七条生成式人工智能服务提供者(以下称提供者)应当依法开展预训练、优化训练等训练数据处理活动，遵守以下规定：(二)涉及知识产权的，不得侵害他人依法享有的知识产权。

³ 《专利法》第二十五条对下列各项，不授予专利权：(二)智力活动的规则和方法。

⁴ 《专利法》第二条第二款发明，是指对产品、方法或者其改进所提出的新的技术方案。

四. 结语

量化基金管理人作为计算机技术在金融领域融合应用的关键推动者，建议积极主动地开展知识产权的风险排查，建立知识产权合规框架和保护机制，确保数据和算法的合法使用，同时保证自身的知识产权得到有效保护。

如您希望就相关问题进一步交流, 请联系:



车小燕
+86 755 3391 7698
cherri.che@llinkslaw.com



汪健
+86 21 3135 8726
lawrence.wang@llinkslaw.com

如您希望就其他问题进一步交流或有其他业务咨询需求, 请随时与我们联系: master@llinkslaw.com

上海

上海市银城中路 68 号
时代金融中心 19 楼
T: +86 21 3135 8666
F: +86 21 3135 8600

北京

北京市朝阳区光华东里 8 号
中海广场中楼 30 层
T: +86 10 5081 3888
F: +86 10 5081 3866

深圳

深圳市南山区科苑南路 2666 号
中国华润大厦 18 楼
T: +86 755 3391 7666
F: +86 755 3391 7668

香港

香港中环遮打道 18 号
历山大厦 32 楼 3201 室
T: +852 2592 1978
F: +852 2868 0883

伦敦

1/F, 3 More London Riverside
London SE1 2RE
T: +44 (0)20 3283 4337
D: +44 (0)20 3283 4323



www.llinkslaw.com



Wechat: Llinkslaw

本土化资源 国际化视野

免责声明:

本出版物仅供一般性参考, 并无意提供任何法律或其他建议。我们明示不对任何依赖本出版物的任何内容而采取或不采取行动所导致的后果承担责任。我们保留所有对本出版物的权利。

© 通力律师事务所 2024