

# 资本市场金融科技创新试点（上海） 项目公示表

填报时间：2024年02月20日

试点公示（对于通过试点申请的项目，《公示表》将在项目公示阶段对社会公开）

辅导公示（对于通过辅导申请的项目，《公示表》将在项目公示阶段对社会公开，  
标\*项目可酌情填写，或填“暂无”、“不适用”）

一、 项目 基本 信息	1.1 申报单位	1.1.1 牵头申报单位： 海通证券股份有限公司
		1.1.2 联合申报单位： 交通银行股份有限公司 蓝象智联（杭州）科技有限公司
	1.2 项目名称	基于隐私计算的银行证券风险信息共享平台
	1.3 项目类型	<input type="checkbox"/> 金融服务 <input type="checkbox"/> 科技产品 <input type="checkbox"/> 业务辅助 <input type="checkbox"/> 合规科技 <input type="checkbox"/> 监管科技 <input checked="" type="checkbox"/> 行业平台 <input type="checkbox"/> 行业基础设施 <input type="checkbox"/> 其他(需补充说明): _____
	1.4 应用场景	<p>本次申报的基于隐私计算的银行证券风险信息共享平台项目中包含两个场景方向，分别是“券商银行风险共享方向”和“集团风险共享方向”。</p> <p>一、券商银行风险共享方向</p> <p>场景 1.1 银行券商基于多方安全计算提升风控能力</p> <p>基于多方安全计算技术的安全求交能力和匿踪查询能力，交通银行/海通证券能够查询本地的风险机构客户是否同时也存在于海通证券/交通银行的风险机构客户当中。利用安全求交和匿踪查询可以保护双方的数据安全，查询方只能查询单一信息，同时被查询方无法知道查询方查询了什么内容。最终能够实现双方风控能力的补充。</p> <p>场景 1.2 券商基于多方安全计算和联邦学习提升高风险机构客户识别能力</p>

		<p>海通证券依据投资者交易行为评估高风险机构客户，基于多方安全计算和联邦学习技术，结合交通银行的高风险机构借贷客户信息可以提升海通证券高风险机构客户评估模型的识别能力。基于多方安全计算的安全求交能力，找到海通证机构投资者和交通银行高风险机构借贷客户的交集。使用分类算法进行模型训练，同时将这部分交通银行高风险机构借贷客户信息作为额外的特征向量补充在海通证券的高风险机构客户评估模型当中，在保证双方数据不出库的同时最终提升评估模型的识别能力。</p> <p>场景 1.3 贷后监控</p> <p>基于多方安全计算技术的安全求交能力和匿踪查询能力，交通银行能够发现近期借贷的机构客户是否在海通证券进行大额的股票买入交易。同时交通银行可以把既存在借贷又存在大额股票买入交易的机构客户作为新的目标变量优化其原有的风控模型。</p> <p>二、集团风险共享方向</p> <p>场景 2.1 集团内部基于多方安全计算提升风控能力</p> <p>基于多方安全计算技术的安全求交能力和匿踪查询能力，集团子公司能够查询本地的风险客户是否同时也存在于海通证券总部的风险客户名单中。利用安全求交和匿踪查询可以保护双方的数据安全，查询方只能查询单一信息，同时被查询方无法知道查询方查询了什么内容。最终能够实现集团子公司风控能力的补充。</p>															
	<p><b>*1.5 数据应用</b></p>	<p>海通证券数据：</p> <p>1.海通证券依据投资者交易行为评估的高风险机构客户；</p> <p>2.海通证券大额交易记录的机构客户；</p> <p>交通银行数据：</p> <p>1.交通银行风险机构客户信息；</p> <p>本项目通过多方安全计算及联邦学习技术探索建立跨主体数据安全共享隐私计算平台，在保障原始数据不出域前提下规范开展数据共享应用，确保数据交互安全、使用合规、范围可控，实现数据可用不可见、数据不动价值动。</p>															
	<p><b>*1.6 实施计划</b></p>	<table border="1"> <thead> <tr> <th>项目阶段</th> <th>起始</th> <th>结束</th> <th>工作内容</th> <th>输出</th> </tr> </thead> <tbody> <tr> <td>方案详细规划设计阶段</td> <td>T</td> <td>T+1月</td> <td>方案设计与评审</td> <td>详细需求调研、功能详细设计、详细数据需求分析、场景模型详细设计；</td> </tr> <tr> <td>开发及定制化阶段</td> <td>T+1月</td> <td>T+2月</td> <td>定制开发</td> <td>实现数据流通相关的计算功能组件，达到运行稳定、功能完整、操作方便、建模正确。</td> </tr> </tbody> </table>	项目阶段	起始	结束	工作内容	输出	方案详细规划设计阶段	T	T+1月	方案设计与评审	详细需求调研、功能详细设计、详细数据需求分析、场景模型详细设计；	开发及定制化阶段	T+1月	T+2月	定制开发	实现数据流通相关的计算功能组件，达到运行稳定、功能完整、操作方便、建模正确。
项目阶段	起始	结束	工作内容	输出													
方案详细规划设计阶段	T	T+1月	方案设计与评审	详细需求调研、功能详细设计、详细数据需求分析、场景模型详细设计；													
开发及定制化阶段	T+1月	T+2月	定制开发	实现数据流通相关的计算功能组件，达到运行稳定、功能完整、操作方便、建模正确。													

					实现算法组件化和流程化研发,支持更直观设计建模过程
	实施部署 联调测试 阶段	T+2 月	T+3 月	部署上线	包括各方平台部署上线和测试、数据接入、测试和准备。
	场景模型 搭建与测 试阶段	T+3 月	T+10 月	业务实现	基于平台实现包括各类数据流通业务场景部署、调参和模型发布。
	试运营阶 段	T+10 月	T+12 月	试运行	平台试运行、模型试运行
	项目验收 阶段	T+11 月	T+12 月	项目验收	验收资料整理、验收报告
	<b>1.7 面临的困难及解决思路</b>	<p>难点 1: 风险信息共享平台的计算效率能否满足业务的需求 当前隐私计算的速度较之明文计算依然存在 1 至 2 个数量级的差距。随着场景应用的不断成熟,将会面临大批量、高并发的计算场景,同时相关应用对计算结果的实时性也存在较高的要求。因此会对平台的计算效率构成挑战。总体来看,平台需要兼顾计算体系内的隐私保护程度和计算信任度的同时提升计算效率。短期内可以通过拓展硬件配置的方式缓解潜在的计算效率压力,长远来看需要不断优化隐私计算相关算法来实现效率的突破。</p> <p>难点 2: 隐私计算技术的安全可解释性 隐私计算应用了安全增强、信息保护的手段进行人工智能运算,对算法可解释性提出了更高的要求,也是亟待解决的实际问题。部分技术在隐私保护的同时对隐私信息做了类似于差分、混淆和散列等熵增动作,数据的无序化使可解释性规则更加难以挖掘,需要更多的技术理论突破。</p> <p>中国人民银行印发《金融科技发展规划(2022-2025年)》中的重点任务中明确提出,要全面加强数据能力建设,在保障安全和隐私前提下推动数据有序共享与综合应用,充分激活数据要素潜能,有力提升金融服务质效。在技术方面,积极应用多方安全计算、联邦学习、差分隐私、联盟链等技术,探索建立跨主体数据安全共享隐私计算平台,在保障原始数据不出域前提下规范开展数据共享应用,确保数据交互安全、使用合规、范围可控,实现数据可用不可见、数据不动价值动。在管理方面,探索建立多元化数据共享和权属判定机制,明确数据的权属关系、使用条</p>			

		<p>件、共享范围等，通过模型计算、模糊查询、智能核验等方式实现跨机构、跨地域、跨行业数据资源有序共享，在确保最小必要、专事专用前提下增强金融数据规模效应和正外部性，提升数据要素资源配置效率。</p> <p>隐私计算在信息保护的前提下进行联合运算，对于可解释性的要求需要在该框架下进行传统算法的再认识，这也需要在实践中积累工程实施的可解释性经验；</p> <p>难点 3：跨公司资源协调</p> <p>本项目需要多家公司共同协作完成建设和运营，涉及数据使用方、数据提供方以及平台提供方，管理难度比一般项目要高，可能会存在高沟通成本问题，产生项目延期。</p> <p>需要在项目建设前期，明确和同步各家公司的数据权责、同时将需要涉及的数据提前接入到各方平台，避免在项目中期出现权限审批相关的流程工作，耽误整体项目进度。除此之外，还需要通过建立高效的组织管理机制，建立领导层和实施层双层协调机制；加强项目过程管理，首先指定详细的项目实施计划，为项目实施各阶段编制相应的规范、标准和目标。在项目各阶段的活动，定期召开项目例会和专项问题讨论会，针对具体工作和问题，探讨应对策略，落实解决方案和实际解决效果，保障项目进度。</p>															
	1.8 专利、认证或奖项	<table border="1"> <thead> <tr> <th data-bbox="596 1160 727 1249">序号</th> <th data-bbox="727 1160 957 1249">奖项名称</th> <th data-bbox="957 1160 1066 1249">类别</th> <th data-bbox="1066 1160 1241 1249">颁发单位</th> <th data-bbox="1241 1160 1437 1249">颁发日期</th> </tr> </thead> <tbody> <tr> <td data-bbox="596 1249 727 1615">1</td> <td data-bbox="727 1249 957 1615">《一种基于匿名化数据的纵向逻辑回归建模方法》</td> <td data-bbox="957 1249 1066 1615">专利</td> <td data-bbox="1066 1249 1241 1615">国家知识产权局</td> <td data-bbox="1241 1249 1437 1615">20220906</td> </tr> <tr> <td data-bbox="596 1615 727 1968">2</td> <td data-bbox="727 1615 957 1968">《一种不暴露中间结果的私有数据隐匿求交方法》</td> <td data-bbox="957 1615 1066 1968">专利</td> <td data-bbox="1066 1615 1241 1968">国家知识产权局</td> <td data-bbox="1241 1615 1437 1968">20220225</td> </tr> </tbody> </table>	序号	奖项名称	类别	颁发单位	颁发日期	1	《一种基于匿名化数据的纵向逻辑回归建模方法》	专利	国家知识产权局	20220906	2	《一种不暴露中间结果的私有数据隐匿求交方法》	专利	国家知识产权局	20220225
序号	奖项名称	类别	颁发单位	颁发日期													
1	《一种基于匿名化数据的纵向逻辑回归建模方法》	专利	国家知识产权局	20220906													
2	《一种不暴露中间结果的私有数据隐匿求交方法》	专利	国家知识产权局	20220225													

		3	《一种用于安全多方计算的数据隐匿查询方法》	专利	国家知识产权局	20221209
		4	《一种基于秘密分享的共享数据等宽分箱方法》	专利	国家知识产权局	20220908
		5	《一种一方加密多方联合解密的数据加解密方法》	专利	国家知识产权局	20230509
		6	《一种联邦衍生特征逻辑回归建模方法》	专利	国家知识产权局	20220712
		7	《一种用于联邦学习的数据点乘运算方法—发明专利证书》	专利	国家知识产权局	20230227

		8	《一种用于联邦学习的数据点乘处理方法》	专利	国家知识产权局	20230510
		9	《一种安全隐患三要素查询方法》	专利	国家知识产权局	20210625
		10	《一种用于联邦学习的特征过滤方法》	专利	国家知识产权局	20221125
		11	《国家金融科技测评中心（BCTC）颁发的多方安全计算金融应用评测》	认证	银行卡检测中心	20220801
		12	《信通院颁发的联邦学习安全专项测评证书》	认证	信通院	20211220

		13	《金融科技产品认证证书(多方安全计算金融应用)》	认证	北京国家金融科技认证中心有限公司	20230330
		14	《国家金融科技测评中心(BCTC)颁发的联邦学习金融应用评测证书》	认证	信通院	20220228
		15	《信通院颁发的多方安全计算基础能力测评证书》	认证	信通院	20201218
二、依法合规原则评估	*2.1 涉及的业务场景是否由持牌机构提供	<p>2.1.1 申报机构已取得的证券期货相关法定业务资格名称：证券经纪；证券自营；证券承销与保荐；证券投资咨询；与证券交易、证券投资活动有关的财务顾问；直接投资业务；证券投资基金代销；为期货公司提供中间介绍业务；融资融券业务；代销金融产品；股票期权做市业务；中国证监会批准的其他业务，公司可以对外投资设立子公司从事金融产品等投资业务。</p> <p>2.1.2 本次申报项目业务场景涉及的业务资格：证券经纪；融资融券证券投资基金代销；代销金融产品。</p>				

2.2 现行法律法规和监管规定符合情况	<p>2.2.1 证券监管部门的相关法规及符合情况： 本项目履行证券监管部门的相关法规如下：</p> <ul style="list-style-type: none"> <li>✓ 证监会《证券投资基金经营机构信息技术管理办法》</li> <li>✓ 证监会《证券公司和证券投资基金管理公司合规管理办法（2020年修订）》</li> </ul> <p>证监会《证券公司风险控制指标管理办法（2020年修订）》</p>	<p>2.2.2 行业协会、交易所等自律组织的相关规范及符合情况： 本项目履行行业协会、交易所的相关法规如下：</p> <ul style="list-style-type: none"> <li>✓ 中证协《证券公司全面风险管理规范》</li> <li>✓ 中证协《证券公司合规管理实施指引》</li> <li>✓ 中证协《证券公司反洗钱工作指引》</li> <li>✓ 中证协《证券公司投资者权益保护工作规范》</li> </ul> <p>中证协《证券公司外部接入信息系统评估认证规范》</p>
		<p>2.2.3 国家或其他管理部门的相关法规及符合情况：</p> <p>1.本项目履行如下法律法规条款：</p> <ul style="list-style-type: none"> <li>✓ 《中华人民共和国网络安全法》</li> <li>✓ 《中华人民共和国数据安全法》</li> <li>✓ 《中华人民共和国个人信息保护法》</li> <li>✓ 《民法典》</li> </ul> <p>目前根据《个人信息保护法》第二十三条，“个人信息处理者向其他个人信息处理者提供其处理的个人信息的，应当向个人告知接收方的名称或者姓名、联系方式、处理目的、处理方式和个人信息的种类，并取得个人的单独同意”，目前本项目不涉对外提供任何个人信息，不存在个人用户告知问题。</p> <p>2.本项目的数据开发利用建设依据遵循：</p> <ul style="list-style-type: none"> <li>✓ 《中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见》</li> </ul> <p>3.本项目的技术利用遵循如下标准：</p> <p>1) 多方安全计算国际标准</p> <ul style="list-style-type: none"> <li>✓ ISO/IEC 19592-1:2016 《Information Technology-Security Techniques-Secret Sharing-Part 1: General》 国际标准化组织/国际电工委员会联合技术委员会 19592-1:2016 《信息技术-安全技术-秘密共享-第1部分：概述》</li> <li>✓ ISO/IEC 19592-2:2017 《Information Technology-Security Techniques-Secret Sharing-Part 2: Fundamental Mechanisms》 国际标准化组织/国际电工委员会联合技术委员会 19592-2:2017 《信息技术-安全技术-秘密共享-第2部分：基本机制》</li> </ul>



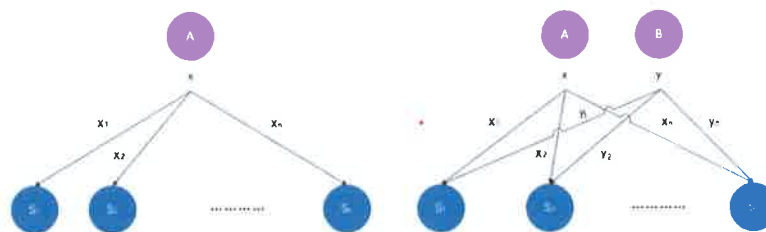
		<ul style="list-style-type: none"> <li>✓ ISO/IEC CD 4922-1 《Information Security-Secure Multiparty Computation-Part 1: Genera》国际标准化组织/国际电工委员会联合技术委员会 CD 4922-1 《信息安全-多方安全计算-第1部分：概述》</li> <li>✓ ISO/IEC WD 4922-2.2 《Information Security-Secure Multiparty Computation-Part 2: Mechanisms Based on Secret Sharing》国际标准化组织/国际电工委员会联合技术委员会 4922-2.2 《信息安全-多方安全计算-第2部分：基于秘密共享的机制》</li> <li>✓ IEEE P2842 《Recommended Practice for Secure Multi-party Computation》电气与电子工程师协会 P2842 《多方安全计算的推荐实践》</li> <li>✓ 密标委 《多方安全计算密码技术框架》</li> <li>✓ 金标委 《JR/T 0196-2020 多方安全计算金融应用技术规范》</li> </ul> <p>2) 同态加密国际标准：</p> <ul style="list-style-type: none"> <li>✓ ISO/IEC 18033-8 《Information technology - Security techniques - Part 8: Fully homomorphic encryption》国际标准化组织/国际电工委员会联合技术委员会 18033-8 《信息技术-安全技术-第8部分：完全同态加密》</li> <li>✓ ISO/IEC 18033-6 《Information technology - Security techniques - Part 6: Homomorphic encryption》国际标准化组织/国际电工委员会联合技术委员会 18033-6 《信息技术-安全技术-第6部分：同态加密》</li> <li>✓ 密标委 《全同态加密技术研究》</li> </ul> <p>3) 联邦学习国际标准：</p> <ul style="list-style-type: none"> <li>✓ IEEE P3652.1 《Guide for Architectural Framework and Application of Federated Machine Learning》电气与电子工程师协会 P3652.1 《联邦机器学习体系结构框架和应用指南》</li> <li>✓ IEEE P2830 《Standard for Technical Framework and Requirements of Trusted Execution Environment based Shared Machine Learning》电气与电子工程师协会 P2830 《基于可信执行环境的共享机器学习的技术框架标准及要求》</li> <li>✓ IEEE P2986 《Draft Recommended Practice for Privacy and Security for Federated Machine Learning》电气与电子工程师协会 P2986 《联邦机器学习的隐私与安全的推荐实践（草案）》</li> </ul>
--	--	--

		<p>行业实践参考：          ✓ 交通银行 《隐私计算金融应用蓝皮书》</p> <p>四、本项目在执行过程中遵循如下原则：          1.本项目在执行过程中确保数据用途可控，并在数据使用的整个生命周期过程中，确保数据授权、最小够用、全程防护。          2.数据使用方对数据源的使用需由数据提供方进行事前审核、授权和使用监督，保证数据使用的合法性和证据固化，实现数据应用的全流程可监控、可审计。          3.双方的业务系统间采用了逻辑隔离的方式相互隔离，仅通过一对一专线进行数据传输，最大化保护用户隐私及数据安全。</p>
	<p>*2.3 出具合规评估意见的机构、评估时间及评估结论</p>	<p>2.3.1 评估机构名称          海通证券股份有限公司 合规管理部</p> <p>2.3.2 出具时间：          2024年2月20日</p> <p>2.3.3 评估结论（最终结论）：          基于隐私计算的银行证券风险信息共享平台属于行业平台类项目，目前根据《个人信息保护法》第二十三条，“个人信息处理者向其他个人信息处理者提供其处理的个人信息的，应当向个人告知接收方的名称或者姓名、联系方式、处理目的、处理方式和个人信息的种类，并取得个人的单独同意”，共享平台中预计使用的信息是证券公司和银行独立采集的客户敏感信息，如需进行跨公司分享，则应以取得客户同意为前提，方能具备合法性与可行性。项目组将针对上述前提条件继续做好各项准备工作，建立并完善风险控制措施，取得试点资格后稳步推进项目实施。</p>
<p>三、有序创新原则评估</p>	<p>3.1 技术创新情况</p>	<p>本项目基于多方安全计算技术及联邦学习技术，实现对证券及银行的风控类业务的科技赋能，并具备行业领先优势。</p> <p>一、本项目所使用的新兴技术介绍          1.多方安全计算          多方安全计算是一种基于多方数据协同完成计算目标，实现除计算结果及其可推导出的信息之外不泄漏各方隐私数据的密码技术。国外对应这一概念的是 Secure Multi-party Computing，简称 SMPC 或 MPC，学术界一般称为“安全多方计算”，产业内则普遍称“多方安全计算”，如中国人民银行发布的金融行业标准</p>

《多方安全计算金融应用技术规范》(JR/T 0196-2020)、工业和信息化部发布的团体标准《基于多方安全计算的数据流通产品技术要求与测试方法》等。多方安全计算是由一系列密码学安全计算协议组成的协议栈，常采用的技术有秘密分享 (Secret Sharing)、不经意传输 (Oblivious Transfer)、混淆电路 (Garbled Circuit)、同态加密 (Homomorphic Encryption) 等。

### 1.1 秘密分享

秘密分享 (Secret Sharing, SS) 最早由 Shamir 和 Blakley 在 1979 年提出。其技术思路是，将每方数据分割成随机秘密分片，并将每个分片分发给不同的参与方分别管理，各参与方基于获得的随机秘密分片进行计算，并基于各参与方计算的结果进一步运算得到最终结果。由于秘密分片的分割是随机的，单个参与方无法用分片恢复出秘密，需要由所有参与者或满足门限数量的参与者一同协作才能恢复整体秘密消息，以此实现了原始数据的保护。



### 1.2 不经意传输

不经意传输 (Oblivious Transfer, OT) 由 Rabin 在 1981 年首次提出，在不经意传输中，通常包括发送方和接收方两类角色，发送方拥有一个“消息-索引”对  $(M_1, 1), \dots, (M_N, N)$ 。在每次传输时，接收方选择一个满足  $1 \leq i \leq N$  的索引  $i$ ，并接收  $M_i$ 。接收方不能得知关于数据库的任何其他信息  $(M_j, j \neq i)$ ，发送方也不能了解接收方  $i$  选择的任何信息，即不知  $i$  值。

### 1.3 混淆电路

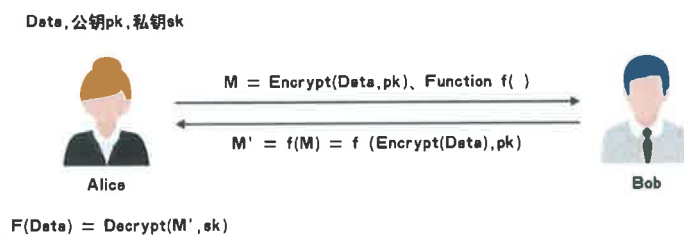
混淆电路 (Garbled Circuit, GC) 由姚期智在 20 世纪 80 年代首次提出，是指将计算逻辑编译成计算电路 (与、或、非电路)，并将计算电路分解为产生阶段和求值阶段。每一方都负责一个阶段，由于每一阶段中电路都被加密处理，所以任何一方都不能从其他方获取信息，但仍然可以根据电路获取整体运算结果。混淆电路由一个不经意传输协议和一个分组密码组成。电路的复杂度至少是随着输入内容的增大而线性增长的。

### 1.4 同态加密

同态加密 (Homomorphic Encryption, HE) 是一种允许用户直接在

密文上进行运算的加密形式，其运算得到的结果仍是密文，并且该密文结果的解密结果与对明文运算的结果一致。目前的同态加密实现多为非对称加密算法，即所有拥有公钥的参与方都可以加密、执行密文计算，但只有私钥所有者可以解密。数学上可以将同态性表述为，加密算法满足

$Dec(sk, Enc(pk, m1) \odot Enc(pk, m2)) = m1 \otimes m2$ ，其中  $m1$ 、 $m2$  为明文， $pk$ 、 $sk$  分别为公钥、私钥， $Enc()$ 、 $Dec()$  分别为同态加密、同态解密， $\odot$ 、 $\otimes$  分别为明文域和密文域上的运算。当明文域上的运算为加法时，称满足加法同态；当明文域上的运算为乘法时，称满足乘法同态。

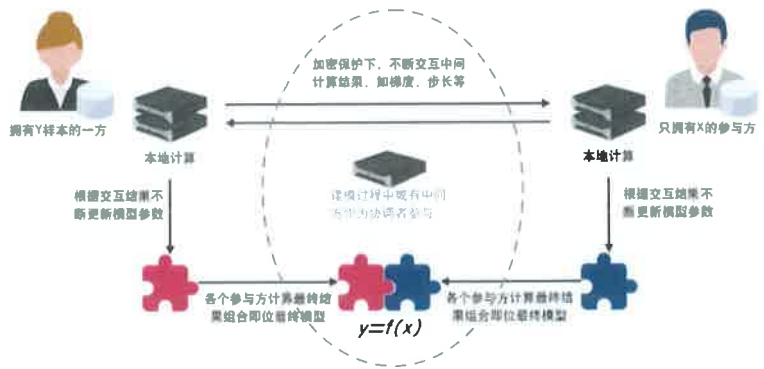


同态加密最早在 1978 年由 Ron Rivest、Leonard Adleman 和 Michael L. Dertouzos 以银行为应用背景提出概念。随后发现了部分同态算法（半同态），支持做有限次的同种运算。2009 年，由 Craig Gentry 提出了首个全同态加密方案，支持做无限次的多种运算及运算组合，并逐渐发展出多种不同的全同态加密方案。同态加密按密文计算能力可以分为三类：

- (1) 部分同态加密(Partial Homomorphic Encryption, PHE)：仅支持单一类型的密文运算(加法或乘法同态)。
- (2) 类同态加密(SomeWhat Homomorphic Encryption, SWHE)：支持有限次的加法和乘法密文同态。
- (3) 全同态加密(Fully Homomorphic Encryption, FHE)：支持任意次数的多类型密文运算(加法和乘法同态)及运算组合。

## 2. 联邦学习

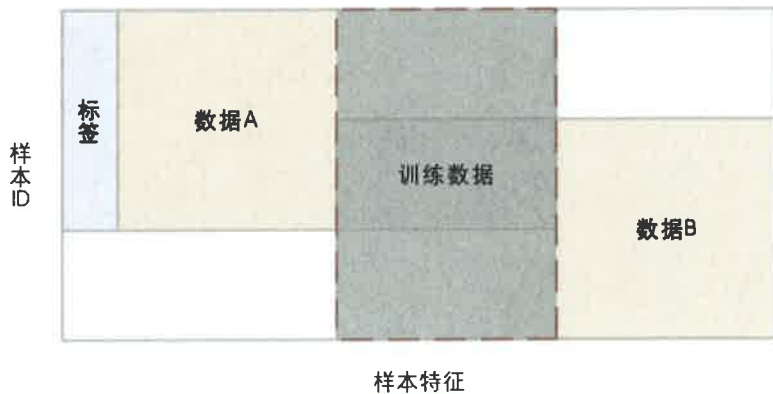
联邦学习 (Federated Learning, FL) 是一种多个参与方在保证各自原始私有数据不出数据方定义的私有边界的前提下，协作完成某项机器学习任务的机器学习模式。



Google AI 团队在 2016 年率先提出了“联邦学习”算法框架，主要针对设备数据集进行协同机器学习模型训练，让数据在不离开设备的情况下，可以在多种设备上训练共享机器学习模型。2019 年 Google 实现了第一个产品级的移动端联邦学习系统，并把该系统从联邦学习推广到联邦计算和联邦分析。其后，英伟达也在 NVIDIA NGC-Ready 服务器上开发了用于分布式协作联邦学习训练的 Clara FL。联邦学习技术分为横向联邦学习、纵向联邦学习、联邦迁移学习三类，在联合营销、联合风控、反洗钱等业务领域均有较多应用。

### 2.1 横向联邦学习

横向联邦学习又称作基于样本的联邦学习，用于特征集( $X_1$ 、 $X_2$ 、 $X_3 \dots$ )相同但样本集( $I_1$ 、 $I_2$ 、 $I_3 \dots$ )不同的场景。可形式化表示为： $X_i = X_j, Y_i = Y_j, I_i \neq I_j \forall D_i, D_j, i \neq j$ 。其中， $X$ 表示特征空间， $Y$ 表示标签空间， $I$ 是样本 ID 空间， $D$ 表示共同数据集，下同。



例如，不同城市的两家城商行，双方经营业务相似、特征集交集较多，但由于地区差异，客户集交集较少，这时可采用双方数据集中特征重叠的、不同客户数据实现横向联邦学习。

### 2.2 纵向联邦学习

纵向联邦学习又称作基于特征的联邦学习，用于样本集( $I_1$ 、 $I_2$ 、 $I_3 \dots$ )相同但特征集( $X_1$ 、 $X_2$ 、 $X_3 \dots$ )不同的场景，可形式化表示为： $X_i \neq X_j, Y_i = Y_j, I_i = I_j \forall D_i, D_j, i \neq j$ 。



样本特征

例如，同一城市的金融机构和运营商，双方的客户集交集较多，但业务类型相差较大、特征集交集较少，这时可采用双方数据集中客户重叠的、不同特征数据实现纵向联邦学习。

### 2.3 联邦迁移学习

联邦迁移学习用于样本集( $I_1$ 、 $I_2$ 、 $I_3 \dots$ )不同且特征集( $X_1$ 、 $X_2$ 、 $X_3 \dots$ )也不同的场景。可形式化表示为：

$X_i \neq X_j, Y_i \neq Y_j, I_i \neq I_j \forall D_i, D_j, i \neq j$ 。联邦迁移学习是将统的迁移学习引入到联邦学习的范式中。



样本特征

例如，不同城市的银行和运营商，双方客户集交集较少，特征集交集也较少，这时不对数据集进行切分，而是引入数据迁移的方式，使用有限的公共样本集学习两个特征空间之间的公共表示，然后应用于仅具有单侧特征的样本预测。

目前，谷歌、Facebook、百度等机构均在其开源的机器学习框架中提供联邦学习支持，如 TensorFlow-Federated、PaddleFL 等；同时，微众银行、字节跳动、矩阵元等推出了联邦学习开源框架，如 FATE、FedLearner 等。

### 二、新兴技术应用在本项目的场景中带来的价值

在“银行券商基于多方安全计算提升风控能力”场景、“集团内部基于多方安全计算提升风控能力”场景和“贷后监控”场景中（详细说明位于1.4应用场景中），以“银行券商基于多方安全计

		<p>算提升风控能力”场景为例，传统的实现方式有两种，第一种方式：海通证券将风险机构客户信息同步给交通银行，由银行自行匹配。这种方式会导致海通证券将不必要的信息透露给交通银行；第二种方式：海通证券开放数据接口，交通银行通过接口查询信息。这种方式会导致交通银行将自有的机构客户信息暴露给海通证券。利用基于多方安全计算技术的安全求交和匿踪查询可以保护双方的数据安全，查询方只能查询单一信息，同时被查询方无法知道查询方查询了什么内容。最终能够实现双方风控能力的补充。</p> <p>在“券商基于多方安全计算和联邦学习提升高风险机构客户识别能力”场景中（详细说明位于1.4应用场景中），传统的实现方式：海通证券获取银行的高风险机构客户信息，并将相关信息加入海通证券高风险机构客户评估模型。该种方式存在敏感数据传输，将违反国家《数据安全法》和《个人信息保护法》。同时也无法保证数据的时效性。通过应用联邦学习技术，海通证券和交通银行可以各自训练模型，而后将模型相关信息（模型的权重更新和梯度信息）采取加密的方式反复交互优化，最终通过模型聚合得到一个全局模型。已训练好的联邦学习模型不共享，分别置于各参与方。上述过程，每一个参与方拥有的数据都不会离开该参与方，其特点可以总结归纳为“数据不动模型动，数据可用不可见”。</p>
	<p><b>3.2 技术领先优势</b></p>	<p>本项目采用的隐私计算技术包含如下技术路线领先优势：</p> <p>一、提升数据安全及隐私保护水平</p> <p>金融数据要素的价值释放，既应在数据资源的基础上引入先进技术，又要在安全底线的基础上依法依规，在数据融合应用和隐私保护之间寻求动态平衡。本项目应用到隐私计算技术，实现数据“可用不可见”、“可用并可控”、“可控可计量”的综合化解决方案提供了关键技术路径。</p> <p>二、降低金融信息壁垒，激发业务创新</p> <p>围绕隐私计算技术特点形成的综合化解决方案，能有效缓解数据价值共享当中面临的“不愿、不敢、不能”难题，降低了金融数据与其他数据的跨机构共享壁垒，既降低了数据资源保护、数据安全等方面的成本，又实现数据价值融合、释放数据红利，为金融业务模式创新提供技术支撑。</p> <p>三、提升金融风险防范能力，促进行业健康发展</p> <p>隐私计算技术可促进跨机构、跨行业之间数据的有效协同，实现数据的交叉验证，降低风险。通过数据价值融合，可实现风控模</p>

		<p>型的精细化、用户画像精准化，提升金融服务的风险评估能力。与此同时，隐私计算可通过联合统计、匿踪查询、安全求交、联邦学习等功能，实现金融机构间的黑灰名单“可用不可见”、多头借贷行为统计共享、建模特征共享、银行贷后监控，提高整个行业风险联防联控能力，促进金融行业健康、稳定、高效发展。</p>
	<p><b>3.3 服务对象与渠道</b></p>	<p>项目上线后，服务对象为海通证券风控部门及交通银行风控部门。服务对象不涉及个人投资者。</p>
<p><b>四、风险可控原则评估</b></p>	<p><b>4.1 业务风险防控</b></p>	<p><b>4.1.1 业务风险点：</b> 业务风险点主要有如下几点： 1.操作风险 项目参与方在隐私计算平台进行操作时，人员、系统或外部事件因素影响而产生系统运行故障、业务阻塞、数据泄漏或数据丢失问题导致损失。</p> <p>2.数据有效性风险 因项目参与方对项目目标，数据定义或数据采集方法等方面理解不一致的原因，导致数据准确性不足，数据时效性不足或数据完整性不足导致损失。</p> <p>3.模型有效性风险 因业务理解偏差，建模方法不匹配或业务规则调整的原因，导致模型准确度不及预期，或模型失效导致损失。</p> <hr/> <p><b>4.1.2 风险监测机制</b></p> <p>1.操作风险监测 加强业务和流程跟踪，包括工作更新、协调、监督、汇总及反馈等工作的落实；建立风险事件报告机制，明确报告责任、风险事件范围，报告形式及期限，监督落实情况。</p> <p>2.数据有效性风险监测 完善数据质量监测方法，通过技术方法监测数据采集、传输及落地质量，制定每日数据监测报告；建立风险事件报告机制，对数据不能及时提供等影响业务开展的重要风险事件及时反馈，及时处理。</p> <p>3.模型有效性风险监测 完善模型有效性评估方法，采用定期业务评估的方法，监测模型</p>



		<p>上线效果；加强业务规则变化监测，通过专人负责的方式，及时汇报相关法律、法规变化；加强项目团队内部的协调，通过即时联络及定期会议等方式，把控建模方向。</p> <p><b>4.1.3 风险控制措施：</b></p> <p>1.操作风险控制措施</p> <p>建立操作管理制度，制定详细的操作准则，为项目各阶段操作编制相应的规范、标准及目标。在项目开展的各个阶段，对重点环节增加评审检查工作，通过上级评审等方式，确保操作准则得到有效落实；对重点操作增加复核机制，通过双人复核等方式检查验证，防止错误的结果流转至业务服务环节。定期召开项目讨论会，针对工作中出现的具体问题，探讨应对措施，落实解决方案，评估解决效果。</p> <p>2.数据有效性风险控制措施</p> <p>构建数据管理体系。数据采集过程，在多方沟通的基础上，通过数据采集协议等方式明确数据采集方法、数据范围、使用目的及目的；数据存储过程，通过加密技术对原始信息进行处理，数据及时备份并严格控制访问权限；传输过程，采用加密信道进行传输，实时监控传输日志，出现传输失败及时通知技术团队解决。</p> <p>3.模型有效性风险控制措施</p> <p>构建全流程模型管理体系。调研阶段，业务部门协同技术部门进行充分的沟通，通过需求文件等方式明确业务逻辑、建模目标等；研发阶段，全面评估模型适用场景，准确评估模型效果；上线阶段，定期监测模型效果，优化迭代；建立合规风控结果审核、客户反馈处置等全流程人工干预机制，并及时通知技术团队优化模型。</p> <p><b>4.1.4 应急预案：</b></p> <p>1.成立应急保障团队，制定主动监控和评估机制，对重点环节进行评审检查，对系统运行状况进行定期验收及问题排查，发现问题及时介入解决。</p> <p>2.针对相关单位反馈风险认定不准确事件，项目团队针对反馈的情况，对相关问题成因进行梳理，对模型进行迭代优化。</p> <p>3.针对运维部门反馈风险事件，项目团队应及时介入，通过系统重置、主备切换等方式尽快恢复系统功能，尽量减少对业务环节的影响。事后总结问题原因，完善相应操作流程，落实系统升级措</p>
--	--	---



		<p>施，预防类似情况的再次发生。</p>
	<p>4.2 技术风险防控</p>	<p><b>4.3.1 技术风险点：</b>  根据对“基于隐私计算的银行证券风险信息共享平台”项目实施方案的评估可能存在如下几类技术风险：</p> <p>1.平台功能性风险  平台功能性风险包括平台计算结果时效性风险和平台网络通信风险。</p> <p>1.1 平台计算结果时效性风险  由于当前隐私计算的计算速度较之明文计算依然存在1至2个数量级的差距，本次项目建设当中的场景会涉及大量数据的并发计算，且同时对计算结果的实时性也存在较高的要求，因此可能会因性能不足或算法优化程度不足而导致平台无法在预期的时间内产生结果。</p> <p>1.2 平台网络通信风险  本项目涉及多方计算，计算过程中会涉及大量中间参数通过网络进行传输，当网络产生波动，或出现一方通信中断可能会导致平台无法产生最终的计算结果。</p> <p>2.网络安全风险  作为需要进行互联网通信的数据平台势必需要分析和梳理网络安全风险，基于项目实施方案整理如下：</p> <p>2.1 平台漏洞风险  平台漏洞风险是指平台自身存在高危漏洞，例如SQL注入、跨站脚本、反序列化漏洞、访问控制失效等。攻击者可以通过WEB页面、端口通信等方式利用漏洞获取平台的管理权限，从而最终获取平台中的数据权限。</p> <p>2.2 不安全的平台环境风险  不安全的平台环境风险是指因作为平台支撑环境的主机或是网络存在安全风险，从而最终导致平台自身产生安全风险。例如攻击者通过其它信息系统入侵海通内网，并获取海通内网服务器管理权限，再利用海通内部的访问控制规则获取平台所在服务器的管理权限，最后利用漏洞或是挂马、病毒利用操作人员的疏忽获取平台关键的访问权限，从而最终拿到平台的管理权限或者数据权限。</p> <p>3 数据安全风险  平台于本地存储敏感数据，因此需要考虑数据安全风险，如下：</p>

		<p><b>3.1 数据传输风险</b></p> <p>由于多方安全计算和联邦学习算法在计算过程中涉及加密的中间数据传输，这些传输的中间数据的包括经过加密算法和聚合算法处理过的梯度值，理论上存在推回原始数据的可能性。</p> <p><b>3.2 数据违规访问风险</b></p> <p>数据违规访问是指非授权人员绕过身份授权违规访问或传播敏感数据。</p> <p><b>4.3.2 风险监测机制：</b></p> <p>针对上述技术风险，项目风险监测机制如下：</p> <p>加强项目过程管理，首先制定项目详细计划，为项目实施各阶段编制相应的规范、标准和目标。在项目各阶段的活动，特别注意重点环节的评审检查工作，定期召开项目例会和专项讨论会，针对具体工作和问题，探讨应对策略，落实解决方案和实际解决效果。</p> <p><b>4.3.3 风险控制措施：</b></p> <p>针对平台计算结果时效性风险，本次海通证券将通过使用高配置的硬件资源来降低潜在的时效性风险。经过详细周密的调研，业务部门协同技术部门进行联合评估，最终明确本次部署的硬件规格为 64c128g1T。</p> <p>针对平台网络通信风险，排除各种不可抗力，海通证券将从访问控制和负载分配两个维度保障通信的可靠程度。在访问控制维度，各参与计算的成员将遵循最小化部署原则，仅保留业务通信端口的同时关闭多余网络端口，并开通防火墙白名单策略，确保通信不会被意外拦截。在负载分配维度，各参与计算的成员将在其负载均衡设备中下发固定的带宽策略，确保不会被其它通信挤占。</p> <p>针对平台漏洞风险，需在平台正式上线前进行相应的漏洞扫描，确保不存在 OWASP TOP10 漏洞，同时平台的前端访问页面需要采用 HTTPS 协议。此外平台自身的安全基线配置也要满足等级保护三级的基本要求。</p> <p>针对不安全的平台环境风险，要确保平台部署在满足等级保护三级的安全域内，充分利用安全域中的网络安全设备（如防火墙、入侵防御、流量探针等）和主机安全设备（如杀毒软件、HIDS 等）对平台进行防护加固，同时平台自身要开启日志留存功能，确保任何违规的操作能够被记录。</p>
--	--	--

		<p>针对数据传输风险，要确保数据的加密方式支持 SM2 非对称加密等国密算法保障核心技术信息安全。此外，虽然本次项目中涉及的多方计算通信路线会途径互联网，通过截取数据包的方式存在获取加密后的中间结果的可能性，但若要基于传输的中间结果反推原始数据，需要知道计算采用了哪种算法，同时还需要大量的计算资源和大量的计算时间，针对第三方人员反推出原始数据的可能性，目前仅存在于理论中。而针对参与项目的人员，将遵从海通证券的数据管理体系，通过签署保密协议和制定详细的项目操作管理制度方式来规避潜在的数据传输风险。</p> <p>针对数据违规访问风险，要利用技术手段和管理手段同步进行规避，技术手段为，确保所有参与项目的人员仅能通过堡垒机统一对平台进行访问，在保留操作过程的同时也对人员身份进行了二次验证。管理手段为，所有参与项目的人员，需要签署保密协议，明确职责。</p>
		<p><b>4.3.4 应急预案：</b></p> <p>建立开发、测试、生产不同的环境，在平台上线前进行充分测试。建立完备的技术风险应急预案，并定期进行应急演练，在职责范围内完善管理流程、优化监控系统，对信息技术风险进行 24 小时监测和检查，实时感知系统运行、对外服务状态，及时发现技术系统的异常与故障，防范数据安全风险的发生。</p> <p>建立技术应急处置小组和工作制度，在发生上述风险时，应急处置小组及时召开应急处置会议，确定处置负责人，按照应急处置流程执行相关应急预案，持续跟踪处置情况。事后组织应急处置通报会议，针对性修改风险控制方案及完善应急预案。</p>
*4.3 投资者保护机制		<p><b>4.3.1 客户投诉渠道：</b></p> <p>本项目服务于内部合规风控人员，不涉及客户投诉。</p>
		<p><b>4.3.2 投诉处理机制：</b></p> <p>不涉及。</p>
		<p><b>4.3.3 风险补偿机制：</b></p> <p>不涉及。</p>
		<p><b>4.3.4 项目退出机制：</b></p> <p>本项目根据试点情况及监管意见执行项目下线，在确保投资者数据安全的前提下，根据各种可能的业务风险、技术风险等做好相应预案，实现平稳退出。</p> <p>1.业务方面,按照退出方案终止有关服务,及时告知相关业务部门,进行妥善的内部服务切换。</p> <p>2.技术方面,充分开展相关评估,制定完善的系统停用和数据迁移保管方案,确保退出过程中数据清理、隐私保护等工作符合国家</p>



		及证券行业相关规范要求，切实保障平台用户、合作机构合法权益。
--	--	--------------------------------

附页：


<p>牵头申报单位 承诺</p>	<p>本单位郑重承诺：</p> <ol style="list-style-type: none"><li>1.本单位在申报资本市场金融科技创新试点（上海）项目过程中，所提供的一切申报材料信息真实、准确和完整。</li><li>2.申报项目符合依法合规、有序创新、风险可控的申报原则。</li><li>3.申报项目不存在违反法律和行政法规情况，不包含国家秘密信息。</li><li>4.本单位将配合监管部门完成后续评审公示、监督检查或风险处置等工作。</li><li>5.本单位已全面开展合规性评估和内控审计，能够有效保障业务连续性和用户信息安全，保证资金安全。</li></ol> <p>以上承诺如有违反，愿承担相应责任与后果。</p> <p>单位（公章） </p> <p>法定代表人（签字）： </p> <p>年 月 日</p>
----------------------	--

（注：联合申报单位如多于1家，承诺签章栏请相应增加）

附页：

<p>联合申报单位1 承诺</p>	<p>本单位郑重承诺：</p> <ol style="list-style-type: none"><li>1. 本单位在申报资本市场金融科技创新试点（上海）项目过程中，所提供的一切申报材料信息真实、准确和完整。</li><li>2. 申报项目符合依法合规、有序创新、风险可控的申报原则。</li><li>3. 申报项目不存在违反法律和行政法规情况，不包含国家秘密信息。</li><li>4. 本单位将配合监管部门完成后续评审公示、监督检查或风险处置等工作。</li><li>5. 本单位已全面开展合规性评估和内控审计，能够有效保障业务连续性和用户信息安全，保证资金安全。</li></ol> <p>以上承诺如有违反，愿承担相应法律责任。</p> <p> </p> <p>法定代表人（签字）：  年 月 日</p>
-----------------------	--

附页：

<p>联合申报单位 2 承诺</p>	<p>本单位郑重承诺：</p> <ol style="list-style-type: none"><li>1.本单位在申报资本市场金融科技创新试点（上海）项目过程中，所提供的一切申报材料信息真实、准确和完整。</li><li>2.申报项目符合依法合规、有序创新、风险可控的申报原则。</li><li>3.申报项目不存在违反法律和行政法规情况，不包含国家秘密信息。</li><li>4.本单位将配合监管部门完成后续评审公示、监督检查或风险处置等工作。</li><li>5.本单位已全面开展合规性评估和内控审计，能够有效保障业务连续性和用户信息安全，保证资金安全。</li></ol> <p>以上承诺如有违反，愿承担相应责任与后果。</p> <p>单位（公章） 法定代表人（签字）</p>  <p>2024年2月20日</p>
------------------------	---