ICS
CCS

T/

团　　体　　标　　准

T/XXXX XXX—XXXX

面向银行的智能化服务规范

（征求意见稿）

XXXX-XX-XX 发布　　XXXX-XX-XX 实施

XXXX　　发 布

目　　次

前　　言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由XXXX提出。

本文件由XXXX归口。

本文件起草单位：。

本文件主要起草人：。

面向银行的智能化服务规范

1 范围

本文件规定了面向银行的智能化服务指标，包括应用层、技术层、基础层。

本文件适用于银行自助设备提供的智能化服务。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

DB14/T 2164-2020 生物特征识别技术应用指南 公共安全领域

GB 40560-2021 人民币现金机具鉴别能力技术规范

GB/T 20979-2019 信息安全技术 虹膜识别系统技术要求

GB/T 35783-2017 信息技术 虹膜识别设备通用规范

GB/T 37742-2019 信息技术 生物特征识别 指纹识别设备通用规范

GB/T 37076-2018 信息安全技术 指纹识别系统技术要求

GB/T 31488-2015 安全防范 视频监控人脸识别系统技术要求

JR/T 0164-2018 移动金融基于声纹识别的安全应用技术规范

GB/T 33135-2016 信息技术 指静脉识别系统指静脉采集设备通用规范

T/CBA 220-2021 远程银行人工智能客服评价指标规范

GB/T 36464.3-2018 信息技术 智能语音交互系统 第3部分：智能客服

JR/T 0263—2022 机器学习金融应用技术指南

GB/T 18789.1-2013 信息技术 自动柜员机通用规范 第1部分：设备

GA 745-2017 银行自助设备、自助银行安全防范要求

GB/T 35678-2017 公共安全 人脸识别应用 图像技术要求

JR/T 0171-2020个人金融信息保护技术规范

3 术语和定义

GB 40560-2021、GB/T 31488-2015、GB/T 20979-2019、GB/T 37076-2018、GB/T 36464.3-2018和GB/T 18789.1-2013界定的术语和定义适用于本文件。

4 应用层

4.1 智能识别

4.1.1 身份识别

4.1.1.1 个体识别

应支持以下至少一种识别比对方式：

a） 能验证用户的真实身份是否与其声称的身份一致；

b） 能辨识已注册用户的身份，并能拒绝未注册用户。

4.1.1.2 识别时间

识别时间应不大于5秒。

4.1.1.3 呈现攻击检测

应具有侦测或防止静态图像、动态视频、假体、异物等呈现攻击的能力。

4.1.1.4 警告与报警

进行生物特征验证时，如果用户的样本特征与所比对的已登记特征不符，或在进行生物特征辨识时，如果在特征数据库中检索不到与样本特征相符的候选者，应给出警告信息；

检测出伪造识别图像、识别数据，或复制图像、识别数据，或非授权数据库操作，或其他异常攻击时，应给出报警信息。

4.1.2 钞票识别

对于GB 40560-2021《人民币现金机具鉴别能力技术规范》中规定的自助批量纸币鉴别机具，漏识率应为0%，误识率不高于0.1%，冠字号码字符误读率不高于0.03%，拒钞率不高于1%。

4.1.3 证件识别

宜采用光学识别、射频识别、芯片识别等技术，能够准确识别证件的信息。

应支持多种证件类型的识别，包括但不限于身份证、护照、驾驶证、港澳通行证等。

应具备多语言支持能力，能够识别不同语言的证件信息。

应能够自动识别证件的信息，包括但不限于姓名、性别、出生日期、证件号码、地址等信息。

应能够根据识别结果进行智能验证，验证证件的真伪和有效性。

4.1.4 单据识别

应支持多种票据类型的识别，包括但不限于支票、汇票、本票等；

应支持多种纸张形式票据的识别，包括但不限于A4、A5、B5、B6等；

具备多语言支持能力，能够识别不同语言的银行票据信息；

应能够自动识别银行票据的信息，包括但不限于出票人名称、票据号码、出票日期、金额等信息。

应能够根据识别结果进行智能验证，包括但不限于验证码、防伪码、防伪标签、统一票号等识别手段，验证票据的真伪和有效性。

4.2 智能客服

4.2.1.1 技术要求

具备自然语言处理、语音识别、语音合成等技术能力，能够理解和回答客户的问题，支持语音、触屏、文字或动作等多模态交互方式。

应支持多种渠道的接入，包括网页、移动应用、微信、电话等。

应具备多语言支持能力，能够满足不同语言环境下的客户需求。

4.2.1.2 功能要求

应能够自动回答客户的问题，包括业务咨询、交易查询、账户管理等方面的问答；

应能够根据客户的对话内容进行智能推荐，提供更加精准的服务。

能够进行语音对话记录，保存相关的对话记录和文本分析结果，宜从用户的聊天记录中抽取相关的信息，自动存储为用户画像，以备后续查询和分析。

4.2.1.3 性能要求

响应时间应不大于3s，确保客户的咨询和服务请求能够及时得到响应。

问题识别率不低于90%,问题解决率不低于70%。

4.2.1.4 安全性要求

智能客服系统应遵守相关的数据保护和隐私保护法律法规，确保客户信息的合法使用和保护。应建立完善的数据保护和隐私保护机制,对数据保护满足JR/T 0171-2020《个人金融信息保护技术规范》相关要求。

4.3 智能风控

应能对投资风险进行检测、识别、预警。

应能对交易进行欺诈风险评估，对可能的欺诈交易进行拦截。

利用机器学习、知识图谱等技术，构建客户风险画像，可结合投资者特定的需求及风险承受程度，提供投资组合建议。

5 技术层

5.1 生物特征识别

5.1.1 人脸识别

应支持多种图像来源，包括但不限于摄像头、照片、视频等；

应支持多种图像格式的识别，包括但不限于jpeg、bmp、tif等；

应具备实时处理能力，能够实时识别人脸特征并进行身份验证。

应在人脸比对步骤之前，利用红外图像、可见光图像、三维图像等技术进行采集源的活体判断，进一步提升人脸识别的可靠性。

5.1.2 虹膜识别

在总比对次数不小于500万次、样本来源不少于3000只眼睛，当错误接受率不大于0.0001%时，错误拒绝率应不大于3%。

5.1.3 指纹识别

对于指纹验证，在总比对次数不小于1000万次、样本来源不少于5000枚指纹，当错误接受率为0.01%时，错误拒绝率应不大于3%。

对于指纹辨识，系统指纹库规模应不低于10000枚指纹，错误接受辨识率为0.2%时，辨识率宜不小于98%。

1：N（N小于等于10000）对比识别速度应不大于1秒。

指纹传感器类型宜采用电容面式指纹传感器。

成像技术宜采用电容反射式感应成像。

5.1.4 指静脉识别

图像采集质量要求分辨率不低于300PPI，畸变率不高于5%，灰度等级256，算法应能够进行图像增强、去噪和分割等操作，提高图像质量，应能够准确提取手指静脉的特征信息，应能够处理各种手指状态和尺寸的图像。

应能够进行1:1身份验证，即根据输入的指静脉特征与预先存储的指静脉特征进行比对，判断是否匹配。

应能够进行1:N身份验证，即根据输入的指静脉特征在数据库中查找最相似的指静脉特征，并返回最相似的结果。

1：N（N小于等于10000）对比识别速度应不大于1秒。

指静脉数据输出应可输出指静脉加密图像。

接口通讯应对人体输入学设备免驱。

应具有工作指示灯，显示识别通过、失败状态。

5.1.5 声纹识别

错误接受率(FAR)应不大于0.5%。

错误拒绝率(FRR)应不大于3.0%。

5.2 智能语音

语音交互平均响应时间宜在2秒以内。

在低噪环境（噪声强度在50 dB以下）中，关键词语音识别的字正确率宜在90%以上；在高噪环境（噪声强度在60 dB~65 dB）中，关键词语音识别的字正确率宜在85%以上。

在低噪环境（噪声强度在50 dB以下）中，连续语音识别的字正确率宜在85%以上;在高噪环境(噪声强度在60 dB~65 dB)中,连续语音识别的字正确率宜在80%以上。

语义理解能力客观测试准确率应达到95%以上，主观测试准确率应达到85%以上。主观测试方法、客观测试方法见GB/T 36464.3—2018第6章。

语音交互成功率应在80%以上。

宜支持方言识别及服务功能。

宜支持语种识别功能、混合语言识别功能。

5.3 机器学习

5.3.1 训练模型

机器学习的训练模型需考虑的内容如下：

a)宜根据金融应用场景需求建立模型选择策略,策略选择依据包括但不限于模型复杂度和可解释性、原始数据及中间数据规模和维度、存储与计算资源成本、金融业务诉求精度和准确率、模型结果的时效性以及潜在的外围干扰项。

b)宜根据金融应用场景需求定义约束条件和评价指标,并将其转换为对应监督学习或无监督学习的性能指标和评价函数，具体内容如下。

——监督学习评价指标，例如混淆矩阵、查准率、查全率、准确率、置信度、误差平方和、决定系数、对数似然损失函数、受试者工作特征曲线(ROC) 的曲线下面积(AUC) 值等。

——无监督学习评价指标，例如方差、还原误差、置信度、困惑度、类间距离、类内距离等。

c)当数据量相对较少时，宜采用交叉验证，例如简单交叉验证、留一交叉验证等。

d)金融应用系统中机器学习引擎的模型训练过程宜根据金融应用场景需求选择合适的训练方法，宜使用特征工程方法提升模型效能，宜根据场景需要提升训练模型的稳定性，宜根据应用场景数据规模优化训练效果。

5.3.2 特征管理

机器学习的特征管理需考虑的内容如下：

a)宜保证特征的保密性，使特征在产生、传输、处理和存储的各个环节中不被泄露给未授权的个人和实体。

b)宜保证特征的完整性，即特征在传输过程中不被篡改、破坏和插入，不发生延迟、乱序和丢失的情况，特别是采用分布式存储方式时，确保内容的一致性。

c)宜保证特征的可用性，使特征可被已授权的其他机器学习子系统或实体使用。

d)根据算法的需要，宜支持自动及手动选择与构建特征。

e)宜支持离散特征、连续特征、时序特征、组合特征的抽取。

f)宜支持显式特征构造方法，以增强特征的可解释性。

g)宜保证离线特征与在线特征的一致性。

5.3.3 算法管理

在基于机器学习的金融系统核心业务中，例如核心交易系统、清结算系统、量化交易系统等，算法设计者宜提高算法的鲁棒性，增强安全性，需考虑的内容如下：

a)当训练数据中有恶意数据破坏原有训练数据分布时，宜能区分和识别恶意数据，防止模型精度降低。具体对策可包括增加算法参数、丰富特征库、优化权重比例等。

b)在算法设计过程中，宜综合考虑安全性、泛化性能和算法开销，合理权衡算法的安全性、实用性和性能，以更好地满足实际应用需求。

c)宜采用集成学习、模型融合等提升手段集成不同的决策方法，以增强模型泛化能力，更好地适应未知数据分布情况。

5.3.4 推理决策

推理决策宜满足金融应用场景需求，需考虑的内容如下：

a)宜保证模型的可用性,对样本数据有较好的容忍度,在极端情况下，保证模型可以正常返回结果，供系统进行决策处理。

b)宜保证输入样本数据及输出返回结果的保密性，确保不被未授权用户非法获取。

c)宜保证输入样本数据及输出返回结果的完整性，确保不被非法篡改。

d)宜保证关键性场景中模型的可解释性,从业务建模、参数设置和样本选择等多方面提升模型可解释性。

e)宜使用模型压缩或者剪裁提升推理速度，并对使用的压缩、蒸馏、裁剪方法记录备案。

f)从多方面提升推理决策服务的可用性和效率，主要内容如下。

——具备任务按需调度能力，根据模型、业务特点确定计算节点和存储节点。

——统一管理中心集群与边缘节点，使边缘业务就近调度到边缘设备执行。

——具备任务跨集群调度与多级资源拉通共享能力，可将本地任务调度到另一个集群中计算。

6 基础层

6.1 数据库

6.1.1 数据采集

数据采集过程需考虑的内容如下：

a)对数据采集的来源和方式、数据范围、数据形式、数据状态等方面进行说明。

b)对数据采集的环境、工具、技术以及校验方法等采取必要的管控措施，保证所采集数据及其标记的完整性、真实性和一致性。.

c)数据采集后，宜对采集数据进行初步清洗，包括去除重复及错误数据、补充残缺数据等。

d)宜记录数据采集过程中的活动及其责任人，保证采集活动的可追溯性和可审计性。

e)个人金融信息数据和金融业务敏感数据的采集过程宜符合JR/T 0171-2020《个人金融 信息保护技术规范》相关要求。

6.1.2 数据存储

数据存储过程需考虑的内容如下：

a)在必要时，宜对采集或应用生成的数据进行持久化存储，以便后续用于模型的训练及校验。

b)数据的存储宜设置存储时限,并满足时间最小化原则,即为满足机器学习目的所必需的最短时间。

c)在超出数据存储时限后，宜立即对所存储的数据进行处理(例如删除、销毁、脱敏等)。

d)宜支持结构化存储方式(例如关系型数据库)和非结构化存储方式(例如键值数据库、图数据库等)。

e)个人金融信息数据和金融业务敏感数据的数据存储过程宜符合JR/T 0171-2020 相关要求。

6.1.3 数据处理

数据处理过程需考虑的内容如下：

a)宜采用规范化的数据处理过程和数据校验机制，保证数据处理结果的一致性和准确性。

b)宜支持数据聚合功能，即对分散的数据进行融合或拼接。

c)宜对数据处理过程进行管理和记录，保证数据处理活动的可追溯性和可审计性。

d)个人金融信息数据和金融业务敏感数据的委托处理和加工处理过程宜符合JR/T 0171-2020相关要求。

6.1.4 数据传输

采用机器学习的金融应用系统的数据传输过程中,对适用于JR/T 0171-2020相关要求的个人隐私及业务敏感数据,宜建立相应的数据传输安全策略和规程并采用有效的技术手段和控制措施，确保数据在传输过程中的保密性、完整性和可用性。

6.1.5 数据调度管理

数据调度管理过程需考虑的内容如下：

a)宜支持数据宿主选择，即选择数据服务过程中承接数据或驱动调度的节点。

b)对于智能金融应用场景中高频次的实时推理决策需求(例如事中不当操作检查),宜使用适当拥塞控制策略以确保整个系统处理链条的健康状态。

c)宜提供弹性扩展功能以增强系统整体吞吐量并分摊数据服务压力。

d)宜支持执行算子抽象技术,将数据流控制算法、数据获取算法等功能抽象为统一数据服务 接口,并遵循标准通信协议,避免跨机构金融协作活动场景中出现数据格式、协议不--致等情况。

e)宜支持数据虚拟化和基于规则的报警机制(例如日志审计)等,确保数据服务的整体健壮性。

6.1.6 数据安全

金融机构从数据的采集、传输、存储、使用、删除、销毁等方面建立全生命周期防护措施,需考虑的内容如下:

a)宜符合JR/T 0171-2020和GB/T 27910-2011的相关要求,确保机器学习应用过程中的个人隐私数据安全和业务敏感数据安全。

b)宜制定机器学习数据采集过程中所涉及的硬件投入计划、源数据量、源数据格式、数据来源的采集结果评定要求。

c)宜对机器学习金融应用模型的测试数据进行脱敏、权限控制等安全保护处理,测试数据要隔离存储,并基于线上数据定期更新、扩充。

6.2 硬件系统

6.2.1 智能柜员机

除GB/T 18789.1-2013《信息技术 自动柜员机通用规范 第1部分:设备》规定的存取款、查询、修改密码等功能外,还可受理圈存、转账汇款、缴费、开卡、开通电子渠道、存折补登、投资理财、公司业务、生活缴费、个人外汇、个人贷款等交易。

除GB/T 18789.1-2013《信息技术 自动柜员机通用规范 第1部分:设备》规定的显示模块、输入模块、卡处理模块、存取款模块、凭条打印模块等配置外,还应具备身份识别、视觉交互、语音交互、安全监控及报警等设备。

6.2.2 安全防范

应符合GA 745-2017《银行自助设备、自助银行安全防范要求》对实体防范、入侵和紧急报警系统、视频监控系统等的要求。