

ZX

人 民 银 行 征 信 系 统 标 准

Q/PBCCRC 1.1-2016

数据采集规范 通用要求

（二代试行）

2017-02 首次发布

2019-11 最新修订

中国人民银行征信中心 发 布

目 次

前 言	III
1 范围	1
2 规范性引用文件	2
3 术语和定义	3
3.1 数据提供机构	3
3.2 数据提供机构区段码	3
3.3 客户资料维护机构	3
3.4 业务管理机构	3
3.5 还款逾期	3
3.6 信息报告日期	3
3.7 数据项	4
3.8 信息段	4
3.9 信息记录	4
3.10 数据报送文件	4
4 基本约定	5
4.1 数据类型	5
4.2 填报约束	6
4.3 空值填报约定	7
4.4 数据项重复出现约定	7
4.5 企业身份标识约定	7
4.6 个人身份标识约定	7
4.7 基本信息记录模板约定	7
4.8 特殊字符转义约定	8
4.9 信息交换编码约定	8
5 数据文件设计	9
5.1 概述	9
5.2 命名规则	10
5.3 文件头	11
5.4 文件体	13
6 常用代码表	16
6.1 身份类别代码表	16
6.2 个人证件类型代码表	16
6.3 企业身份标识类型代码表	16
6.4 分次放款标志代码表	17
6.5 五级分类代码表	17
6.6 资产转让标志代码表	17
6.7 债权转移时的还款状态代码表	17

6.8 授信协议额度循环标志代码表.....	18
6.9 授信协议额度状态代码表.....	18
6.10 代偿（垫款）标志代码表.....	18
6.11 国民经济行业分类代码表.....	18
6.12 世界各国和地区名称代码表.....	19
6.13 行政区划代码表.....	19
6.14 币种代码表.....	20
7 通用校验规则.....	21
7.1 规则说明.....	21
7.2 通用校验规则.....	21
附 录 A （规范性附录） 中征码校验位计算规则.....	25
附 录 B （资料性附录） 修订说明	27
B.1 本次修订记录.....	27
B.2 历史修订记录.....	27

前 言

本规范依据 GB/T 1.1-2009《标准化工作导则 第1部分：标准的结构和编写规则》给出的规则起草。

《人民银行征信系统数据采集规范》系列标准规定了人民银行征信系统从数据提供机构采集企业、事业单位等组织（以下统称企业）和个人信用数据的范围、内容、时点、格式、形态、校验规则以及错误更正机制等方面的具体要求。

本部分为该系列标准的第一部分，重在对本术语进行定义，对各部分共同遵守的约定、对通用的数据文件设计进行说明，并列出了常用代码表及通用校验规则等。

本规范由中国人民银行征信中心提出。

本规范起草单位：中国人民银行征信中心。

本部分主要起草人：杜鲲、邓蕊、冯庄、赵燕。

1 范围

本部分规定了《人民银行征信系统数据采集规范》系列标准其他部分共同遵守的通用要求。

数据提供机构在阅读《人民银行征信系统数据采集规范》其他部分时，应结合本部分内容进行阅读。

2 规范性引用文件

本规范制定主要引用了以下标准或文件。如无特殊说明，凡是注日期的引用文件，其随后所有的更正单（不包括勘误的内容）或修订版均不适用于本标准；凡是不注日期的引用文件，其最新版本适用于本标准。

GB/T 1.1-2009 标准化工作导则 第1部分：标准的结构和编写规则

GB/T 2260-2007 中华人民共和国行政区划代码

GB/T 2659-2000 世界各国和地区名称代码

GB/T 4754-2017 国民经济行业分类

GB 11643-1999 公民身份号码

GB 11714-1997 全国组织机构代码编制规则

GB/T 12406-2008 表示货币和资金的代码

GB 32100-2015 法人和其他组织统一社会信用代码编码规则

其中，行政区划代码最新版本以征信系统使用的为准，具体的可登陆征信系统网站查询、下载。

3 术语和定义

3.1 数据提供机构

依据有关规定或按照信息共享协议向人民银行征信系统提供数据的机构。

本规范主要指从事信贷业务的机构，包括：银行业金融机构，提供融资服务的证券公司、融资租赁公司、保理公司、小额贷款公司、P2P 网络借贷机构，以及融资性担保公司、提供信用保证保险服务的保险公司等。

3.2 数据提供机构区段码

征信系统要求，数据提供机构在报送各类实体（如客户、账户、授信协议、抵（质）押合同）的信息时，应给每个实体分配一个全局（征信系统范围内）唯一的标识码。为防范各机构自行编码带来的冲突（即两家或多家机构编制的标识码有重合），征信系统给每个数据提供机构分配了各自的区段码用作标识码的前缀，这个区段码在本规范中被称为“数据提供机构区段码”。

通常情况下，征信系统给每家数据提供机构仅分配一个唯一且终生不变的区段码。发生数据提供机构合并时，合并后新成立的数据提供机构可以继承被合并机构的区段码，也可以申请新的区段码。但发生数据提供机构拆分时，从原数据提供机构独立出来的新机构应申请新的区段码。

3.3 客户资料维护机构

数据提供机构内部负责客户资料信息管理、核实和相关异议处理的网点/分支机构。

3.4 业务管理机构

数据提供机构内部具体负责业务管理、信息核实和相关异议处理的网点/分支机构。

3.5 还款逾期

借贷交易中，借款人如未能按照约定及时归还本金、利息以及其他费用，均视为还款逾期。

3.6 信息报告日期

为了保证信息更新的及时性，征信系统针对各类信息规定了一系列的状态观测点，要求数据提供机构在这些状态观测点的日终提取需要报送的信息内容，并在规定的时间内上报至征信系统。这些状态观测点，在本规范中被称为“信息报告日期”。

正常情况下，每一条信息记录中都有一个“信息报告日期”，用于说明：截至该“信息报告日期”，信息的有效内容如记录所示。

3.7 数据项

数据项是用一组属性描述定义、标识、表示和允许值的基础数据单元。

3.8 信息段

信息段是一个已标识、命名和结构化的、在功能上相互关联的数据项的集合。每个信息段都由各自定义的数据项组成。

3.9 信息记录

信息记录是征信系统采集信息的基本单位，数据提供机构向征信系统报送数据时，需要组织成信息记录。从组成结构上看，信息记录分为两类：

- 一类由一个基础段和不少于一个其他信息段构成。每个信息段最多出现一次，其中，基础段必须出现，其他信息段视情况出现。
- 另一类不划分信息段，由多个数据项直接构成。

3.10 数据报送文件

数据报送文件是报送数据的载体，由文件头、文件体构成。其中文件体由一条或多条信息记录组成。

4 基本约定

4.1 数据类型

在对数据项进行描述时，本规范定义了 11 种基本数据类型。在采用 XML 描述信息记录时，各数据项的取值本质上都是以字符串的形式表示。

表 4-1 数据类型说明

数据类型	定义	XML 文本格式
N	数字字符串	数字字符串
AN	数字字符和英文字符组成的字符串	数字字符和英文字符组成的字符串
ANC	可以包含所有字符（包括汉字）的字符串	可以包含所有字符（包括汉字）的字符串
Enum	枚举（代码）型数据。代码可由数字和英文字符构成	数字和英文字符构成的字符串
Year	年份	由数字字符组成的用以表示年份的字符串（格式为 yyyy）
Month	年月（包含年、月）	由数字字符和“-”组成的用以表示年月的字符串（格式为 yyyy-mm）
Date	日期（包含年、月、日）	由数字字符和“-”组成的用以表示日期的字符串（格式为 yyyy-mm-dd）
Time	时间（包含年、月、日、时、分、秒）	由数字字符、“-”、“:”和大写字母“T”组成的用以表示时间的字符串（格式为 yyyy-mm-ddTHH:MM:SS, 其中 T 为时间和日期的分隔符）
uInt	自然数（含 0）	由数字字符组成的表示十进制自然数的字符串
Int	整数	由数字字符和“-”组成的表示十进制整数的字符串，其中“-”加在数字前表示该整数为负数
Float	实数	由数字字符、“-”、“.”组成的表示十进制实数的字符串，其中“.”用以分隔整数和小数部分、“-”加在数字前表示该实数是负数

对于前 3 种数据类型而言，本规范通过在其名称后附上数字的方式来表达该数据类型的长度。具体而言，本规范使用了两种长度说明方式：

1. 固定长度：将一个长度数字直接放在基本数据类型之后。例如，“N4”表示一个由数字字符形成的长度为 4 的字符串。
2. 最大长度：将两个点号（“..”）和一个长度数字放在基本数据类型之后。例如，“N..4”表示一个由数字字符形成的最大长度为 4 的字符串。

需要强调的是，与一代接口规范不同，本规范中提到的字符串长度均指字符串中的字符个数，与字符编码方式及单个字符编码所占字节数无关¹。

对于自然数、整数这两种数据类型，本规范通过在其名称后面附上数字来表示该整数的取值范围。具体的表示方法如下：

1. 对于自然数 `uInt`，将两个点号（“..”）和一个整数放在自然数类型之后，表示其取值范围。例如，“`uInt..5`”表示这个自然数的取值范围是大于等于 0 小于 10^5 （不含 10^5 ）。
2. 对于整数 `Int`，将两个点号（“..”）和一个整数放在整数类型之后，表示其取值范围。例如，“`Int..5`”表示这个整数的取值范围是大于 -10^5 小于 10^5 （不含 -10^5 和 10^5 ）。

对于实数类型的数据，采用将两个用逗号分隔开的数字放在实数类型之后的括号中，表示其取值范围以及需要精确的小数点位数。例如，“`Float（5，2）`”逗号后面的数字 2 代表其需要精确到小数点后 2 位，逗号前面的数字 5 代表整数部分取值范围是大于 -10^5 小于 10^5 （不含 -10^5 和 10^5 ）。

4.2 填报约束

本规范从两个维度对数据项的填报进行约束：一是出现约束，规定数据项是否出现在相应的信息段中；二是空值约束，规定当数据项出现时，数据项的取值是否允许为空。

■ 针对数据项是否需要在信息段中出现，本规范定义了如下两种类型的填报约束：

1. **A（必须报送出现）**，表示必须出现在相应的信息段中。
2. **S（条件报送）**，表示当满足特定条件时，此类数据项应出现，否则不出现。

■ 针对数据项是否可以报空值，本规范使用了 3 种类型的填报约束：

1. **M（非空型）**：表示此数据项必须填报有意义的值，不能填写空值。
2. **C（条件非空）**：在特定条件为真的情况下，必须填报有意义的值，不能填写空值；反之，可以为空值。
3. **O（可空型）**：当数据提供机构无法获取相关信息时，该数据项填报空值。当数据提供机构内部有对应数据项时，该数据项需要填报有意义的值。

对于这两种填报约定，前者是从业务层面对信息段完整性提出的要求，后者是从技术层面对数据项取值范围的限定。对于一个数据项，可能对于所有业务都是必须出现的（A），但是技术允许其取值为空（O）；也可能仅仅对于部分业务需要出现（S），但是在技术上要求其取值必定不能为空（M），即：此数据项可能不出现，一旦出现必须填写有意义的值。

¹ 例如，二代系统中 `ANC80` 表示英文字母和汉字一共 80 个，而对于一代系统来说 `ANC80` 表示字符串的长度按字节计算不超过 80 个，通常情况下，只能容纳 40 个汉字。

4.3 空值填报约定

本规范中，空值指空文本（即不含任何字符）或空白字符串（即所含字符均为空格）。

对于 C 型和 O 型数据项，若报空值，推荐用空文本。即：在应填写数据项取值的位置，填写空文本。

4.4 数据项重复出现约定

信息段中带有“→”符号的数据项可以重复出现，出现次数应与表示出现重复次数的数据项取值保持一致。

4.5 企业身份标识约定

通常情况下，征信系统使用“企业身份标识类型+企业身份标识号码”来唯一识别一个企业；如果出现这两项无法唯一确定一个企业的情况，征信系统会使用“企业名称+企业身份标识类型+企业身份标识号码”作为识别企业身份的标识项。

当企业身份标识类型出现在信息记录的基础段时，如无特别说明，身份标识类型只能在选择“中征码、统一社会信用代码、组织机构代码”之一。

填写组织机构代码时，请省略其中的“-”符号。

4.6 个人身份标识约定

通常情况下，征信系统使用“证件类型+证件号码”来唯一识别一个自然人；如果出现这两项无法唯一确定一个自然人的情况，征信系统会使用“姓名+证件类型+证件号码”作为识别自然人身份的标识项。

填报以下两类个人证件号码需要注意：

- 报送公民身份号码时，请将老的 15 位公民身份号码转换为新的 18 位号码；
- 报送护照号码时，请在护照号前补充国籍代码（采用国标《世界各国和地区名称代码》（GB/T 2659-2000）中规定的 3 位拉丁字母编码）。

4.7 基本信息记录模板约定

考虑到不同类型数据提供机构间的差异，二代征信系统不再用一套统一的标准要求所有机构上报相同的基本信息，而是定义了多套模板采集基本信息，同一数据项在不同模板中的填报约束可能不同。数据提供机构可根据自身实际选择合适的模板。

4.8 特殊字符转义约定

当信息记录采用 XML 组织时，需要对“<”、“>”、“&”、“'”、“””五类特殊字符进行转义处理。转义规则如下：

特殊字符	转义后的字符
<	<
>	>
&	&
'	'
”	"

4.9 信息交换编码约定

在信息交换时采用“Unicode”字符集以及无 BOM 的 UTF-8 编码。

5 数据文件设计

5.1 概述

数据提供机构与征信系统之间以文件作为数据交互的载体。

数据提供机构向征信系统上传数据必须先按照规定格式生成一条条的信息记录、再组织成一到多个数据报送文件。决定数据报送文件数量的因素有两个：一是信息记录类型，征信系统规定一个数据报送文件只能包含一类信息记录；二是信息记录数量，征信系统规定数据报送文件包含的信息记录数量不得超过 100 万。

针对接收到的每一个数据报送文件，征信系统在处理完毕后，会唯一生成一个对应的数据反馈文件返给数据提供机构，以通知数据处理结果。即：数据报送文件与数据反馈文件之间一一对应，每一个数据反馈文件唯一对应一个数据报送文件。

数据报送文件与反馈文件遵循同样的命名规则、同样的文件结构。两类文件均由文件头、文件体组成。

文件名、文件头是一串由数字、英文字母以及符号构成定长的字符串，每一位占一个字节。

文件头固定为文件的第一行，文件头用于说明文件体的概要信息以及文件体如何解析。文件头与文件体之间以一个回车换行符（“\r\n”）分隔。

文件体由一条条信息记录构成，每条信息记录为一行，信息记录之间以一个回车换行符分隔。一个文件只能包含同一类型的信息记录。原则上，文件体不能为空，即至少应包含一条信息记录。唯一的例外是，若数据报送文件中的所有信息记录均成功入库，则对应的数据反馈文件中只有文件头，没有文件体。文件体结构如图 5-1 所示。

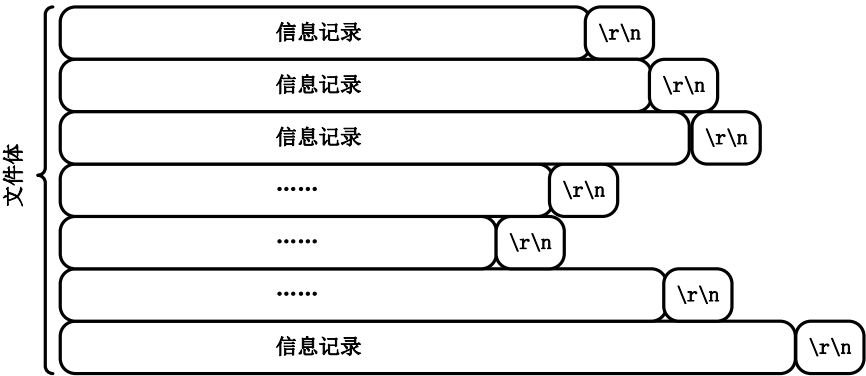


图 5-1 文件体结构示意图

5.2 命名规则

文件名为 34 位的字符串，其中文件名前缀 30 位、文件名后缀 3 位、前缀与后缀的分隔符“.”占 1 位。

为了保证数据传输的安全及减少网络传输的压力，征信系统要求数据报送文件与数据反馈文件在交换前进行加密加压。加密加压后的文件名后缀统一为“enc”，加密加压前的文件后缀统一为“txt”。

文件名前缀仅由英文、数字构成，两类数据文件的前缀具体构成见表 5-1 和表 5-2。

表 5-1 数据报送文件文件名数据元素

位置	数据项名称	数据类型	数据项说明
1-14	数据提供机构 区段码	AN14	数据提供机构区段码是征信系统分配给数据提供机构的，用于作为标识前缀，防止各数据提供机构自行编码导致冲突的编码。
15-22	文件生成日期	N8	数据报送文件的生成日期。 格式为“YYYYMMDD”。
23-25	所含信息 记录类型	N3	数据报送文件中所含记录的类型。 具体参见系列标准其他部分中各信息记录类型代码。
26-26	预留位	AN1	目前统一填“0”。
27-29	流水号	AN3	当文件名前 26 位相同时，用于区分不同文件的编号，以保证数据报送文件命名的唯一性。
30-30	反馈标识	N1	对于数据报送文件，统一填“0”。

表 5-2 数据反馈文件文件名数据元素

位置	数据项名称	数据类型	数据项说明
1-29	数据反馈文件的文件名前 29 位与其对应的原数据报送文件保持一致。		
30-30	反馈标识	N1	数据反馈文件的类型。 代码表如下： 1-代表原数据报送文件中所有信息记录已成功入库 2-代表原数据报送文件的文件名或者文件头出现错误，文件中所有信息记录未能入库 3-代表原数据报送文件中部分信息记录已成功入库，部分记录存在错误未能入库

5.3 文件头

数据报送文件头 50 位，具体构成见表 5-3。

表 5-3 数据报送文件文件头数据元素

位置	数据项名称	数据类型	数据项说明
1-1	文件头起始标识	AN1	填“A”表示文件头开始。
2-4	文件头长度	N3	说明文件头所含字节数（含“文件头起始标识”和“文件头长度”）的十进制数字串。不足 3 位的，左侧用“0”补齐。
5-7	所含记录类型	N3	数据报送文件中所含记录的类型。
8-12	信息记录版本号	AN5	用于标识数据报送机构生成文件体中信息记录依据的是哪个版本的规格定义。 格式为 N.N.N。
13-14	信息记录模板类型	AN2	信息记录采用的模板类型代码。 ➤ 对于基本信息，填写其模板类型代码； ➤ 对于其他信息，填写空值（即，两个空格）。
15-28	数据提供机构区段码	AN14	数据提供机构区段码是征信系统分配给数据提供机构的、用作标识前缀、防止各数据提供机构自行编码导致冲突的编码。
29-42	文件生成时间	N14	生成数据报送文件的具体时间。 格式为“YYYYMMDDHHMMSS”。
43-43	预留位	AN1	对应文件名中的第 26 位。 目前统一填“0”。
44-50	信息记录数	N7	说明文件中所含信息记录数的十进制数字串。不足 7 位的，左侧用“0”补齐。

数据反馈文件头 72 位，具体构成见表 5-4。

表 5-4 数据反馈文件文件头数据元素

位置	数据项名称	数据类型	数据项说明
1-1	文件头起始标识	AN1	填“A”表示本行为文件头。
2-4	文件头长度	N3	说明文件头所含字节数（含“文件头起始标识”和“文件头长度”）的十进制数字串。不足 3 位的，左侧用“0”补齐。
5-9	反馈文件版本号	AN5	用于标识征信系统生成反馈文件依据的是哪个版本的规格定义。 格式为 N.N.N。 所有反馈文件规格定义的初始版本号为“2.0.0”。
10-38	原数据报送文件 文件名	AN29	反馈文件对应的原数据报送文件文件名前 29 位。
39-39	反馈标识	N1	代码表如下： 1-代表原数据报送文件中所有信息记录已成功入库 2-代表原数据报送文件的文件名或者文件头出现错误，文件中所有信息记录未能入库 3-代表原数据报送文件中部分信息记录已成功入库，部分记录存在错误未能入库
40-44	原数据报送文件 信息记录版本号	AN5	反馈文件对应的原数据报送文件的信息记录版本号。
45-58	反馈文件 生成时间	N14	数据反馈文件的生成具体时间。 格式为“YYYYMMDDHHMMSS”。
59-65	正确入库记录数	N7	说明原数据报送文件中正确记录条数的十进制数字串。数量不足 7 位的，左侧用“0”补齐。
66-72	错误反馈记录数	N7	说明原数据报送文件中错误记录条数的十进制数字串。数量不足 7 位的，左侧用“0”补齐。

5.4 文件体

5.4.1 数据报送文件文件体

数据报送文件的文件体由信息记录构成。

信息记录用 XML (EXtensible Markup Language) 描述。每条完整的信息记录为一行，即每条信息记录的 XML 文本中不得含有回车换行符 (“\r\n”)，信息记录之间以回车换行符分隔。

数据报送文件文件体中所含信息记录的具体组成及 XML 样例详见系列标准其他部分。

5.4.2 数据反馈文件文件体

数据反馈文件的文件体存在 3 种不同的情况：

1. 当原数据报送文件中所有信息记录已成功入库时，反馈文件的文件体不包含任何信息记录。
2. 当原数据报送文件的文件名或者文件头发生错误时，反馈文件的文件体只包含一条信息记录。信息记录不分段，组成数据项描述参见表 5-5。

表 5-5 文件名或文件头错误时反馈信息记录数据项说明

数据项名称	数据项说明	数据类型	出现约束	空值约束
反馈错误数	文件名、文件头中的错误数量。	uInt..3	A	M
→ 错误反馈代码	具体参见各部分附录。	AN6	A	M
→ 出错位置		ANC..100	A	M

数据反馈文件中的反馈信息记录由 XML 描述。

反馈信息记录根元素说明见表 5-6。

表 5-6 反馈信息记录根元素说明

元素名称	<XML Tag>	出现次数 ²	下级元素
根元素	Document	-	参见表 5-7

表 5-7 反馈信息记录根元素构成说明

元素名称	<XML Tag>	出现次数	下级元素
反馈信息记录	FbInf	1:1	参见表 5-8

表 5-8 反馈信息记录元素构成说明

元素名称	<XML Tag>	出现次数	下级元素
反馈错误数	FbErrNum	1:1	无
反馈错误信息	ErrInf	1:n	参见表 5-9

² 1:1 代表出现且仅出现一次，1:n 代表至少出现 1 次。

表 5-9 反馈错误信息元素构成说明

元素名称	<XML Tag>	出现次数	下级元素
错误反馈代码	FbCode	1:1	无
出错位置	FbMsg	1:1	无

反馈信息记录样例见图 5- 2。

<pre><Document> <FbInf> < FbErrNum >2</ FbErrNum ><!--反馈信息数--> <ErrInf> <FbCode > ABN003</FbCode><!--反馈信息代码 1--> <FbMsg>0000</FbMsg><!--反馈信息 1--> </ErrInf> <ErrInf> <FbCode > ABN004</FbCode><!--反馈信息代码 2--> <FbMsg> 0000</FbMsg><!--反馈信息 2--> </ErrInf> </FbInf> </Document></pre>
--

图 5- 2 原数据报送文件的文件名或文件头错误时，反馈信息记录 XML 示意

3. 当原数据报送文件中部分信息记录成功入库、部分记录存在错误时，针对每一条错误的信息记录，在反馈文件的文件体中会存在一个对应的反馈信息记录，反馈记录组成见表 5- 10。

表 5- 10 文件中部分信息记录错误时反馈信息记录数据项说明

数据项名称	数据项说明	数据类型	出现约束	空值约束
出错记录标识	填写出错记录在原数据报送文件中的行号。	uInt..8	A	M
反馈错误数	出错记录中的错误数量。	uInt..3	A	M
→错误反馈代码	具体参照各部分附录。	AN6	A	M
→出错位置	出错位置一般填写错误元素的 XML 标签。	ANC..100	A	M
出错信息记录	填写原出错信息记录的 BASE64 编码。 在解析本数据项时，应先解码获取原 XML 格式记录。	ANC..100000	A	M

数据反馈文件中的反馈信息记录由 XML 描述。

反馈信息记录根元素说明见表 5- 11。

表 5- 11 反馈信息记录根元素说明

元素名称	<XML Tag>	出现次数	下级元素
根元素	Document	-	参见表 5-12

表 5- 12 反馈信息记录根元素构成说明

元素名称	<XML Tag>	出现次数	下级元素
反馈信息记录	FbInf	1:1	参见表 5-13

表 5- 13 反馈信息记录元素构成说明

元素名称	<XML Tag>	出现次数	下级元素
出错记录标识	ErrRecId	1:1	无
反馈错误数	FbErrNum	1:1	无
反馈错误信息	ErrInf	1:n	参见表 5-9
出错信息记录	ErrRec	1:1	无

反馈信息记录样例见图 5- 3。

```
<Document>
  <FbInf>
    < ErrRecId >20</ ErrRecId ><!--出错记录标识-->
    < FbErrNum >2</ FbErrNum ><!--反馈信息数-->
    <ErrInf>
      <FbCode >BBR001</FbCode><!--反馈信息代码 1-->
      <FbMsg>0000</FbMsg><!--反馈信息 1-->
    </ErrInf>
    <ErrInf>
      <FbCode > ABE004</FbCode><!--反馈信息代码 2-->
      <FbMsg> IDNum</FbMsg><!--反馈信息 2-->
    </ErrInf>
    < ErrRec >原上报信息记录 BASE64 转码后的字符内容</ ErrRec><!--原出错记录-->
  </FbInf>
</Document>
```

图 5- 3 原数据报送文件中部分记录存在错误时，反馈信息记录 XML 示意

6 常用代码表

6.1 身份类别代码表

代码	中文名称	说明
1	自然人	
2	组织机构	

6.2 个人证件类型代码表

代码	中文名称	说明
1	户口簿	
2	护照	
5	港澳居民来往内地通行证	
6	台湾同胞来往内地通行证	
8	外国人居留证	
9	警官证	
A	香港身份证	
B	澳门身份证	
C	台湾身份证	
X	其他证件	
10	居民身份证及其他以公民身份证号为标识的证件	
20	军人身份证件	

6.3 企业身份标识类型代码表

代码	中文名称	说明
10	中征码（原贷款卡编码）	当企业身份标识类型出现在信息记录的基础段时，如无特别说明，身份标识类型只能在“10、20、30”中选择一个。
20	统一社会信用代码	
30	组织机构代码	
01	工商注册号	
02	事业单位证书号	
03	社会团体登记号	
04	民办非企业登记号	
05	基金会登记号	
06	宗教证书登记号	
41	纳税人识别号（国税）	
42	纳税人识别号（地税）	

6.4 分次放款标志代码表

代码	中文名称	说明
0	贷款额度一次性发放	
1	贷款额度分次发放，且每笔放款对应一个独立的 D1 账户报送	
2	贷款额度分次发放，且所有放款使用统一的 D1 账户合并报送	

6.5 五级分类代码表

代码	中文名称	说明
1	正常	
2	关注	
3	次级	
4	可疑	
5	损失	
9	未分类	

6.6 资产转让标志代码表

代码	中文名称	说明
0	未转让	
1	转让	

6.7 债权转移时的还款状态代码表

代码	中文名称	说明
1	逾期 1-30 天	
2	逾期 31-60 天	
3	逾期 61-90 天	
4	逾期 91-120 天	
5	逾期 121-150 天	
6	逾期 151-180 天	
7	逾期 180 天以上	
0	债务人违约时自动垫款	

6.8 授信协议额度循环标志代码表

代码	中文名称	说明
0	否	
1	是	

6.9 授信协议额度状态代码表

代码	中文名称	说明
1	有效	
2	到期/失效	

6.10 代偿（垫款）标志代码表

代码	中文名称	说明
0	否	
1	是	

以上未注明引用来源的，均属于征信系统自定义代码表。如下代码参照相关国标，部分在国标基础上补充部分代码。

6.11 国民经济行业分类代码表

完整国民经济行业分类代码表详见《国民经济行业分类》（GB/T 4754-2017），具体门类代码以及本规范新增的部分代码见下表：

门类代码	中文名称	说明
A	农、林、牧、渔业	
B	采矿业	
C	制造业	
D	电力、热力、燃气及水生产和供应业	
E	建筑业	
F	批发和零售业	
G	交通运输、仓储和邮政业	
H	住宿和餐饮业	
I	信息传输、软件和信息技术服务业	
J	金融业	
K	房地产业	
L	租赁和商务服务业	
M	科学研究和技术服务业	
N	水利、环境和公共设施管理业	
O	居民服务、修理和其他服务业	

门类代码	中文名称	说明
P	教育	
Q	卫生和社会工作	
R	文化、体育和娱乐业	
S	公共管理、社会保障和社会组织	
T	国际组织	
9	未知	征信系统补充代码。 无法获知单位所属行业门类信息时，填写“9”。

6.12 世界各国和地区名称代码表

参照国标《世界各国和地区名称代码》（GB/T 2659-2000）执行，并补充部分代码。

代码	中文名称	说明
ATG	安提瓜和巴布达	完整代码表见国标。
ARG	阿根廷共和国	
ARM	亚美尼亚共和国	
...	
991	中国（含港澳台）	征信系统补充代码。 ➤ 对于明确是中国人（含港澳台），无法区分大陆与港、澳、台的自然人，填写此代码； ➤ 对于无法获知登记注册地是大陆还是港、澳、台的企业，填写此代码。
992	外国（不含港澳台）	征信系统补充代码。 ➤ 对于明确是外国人（不含港澳台），无法区分具体国籍的自然人，填写此代码； ➤ 对于港、澳、台以外的外国企业，无法区分具体国家的企业，填写此代码。
999	未知	征信系统补充代码。

6.13 行政区划代码表

参照国标《中华人民共和国行政区划代码》（GB/T 2260-2007）执行，目前正在使用的最新版本可登陆征信系统网站查询、下载，部分代码示意如下：

代码	中文名称	说明
110000	北京市	完整代码表见国标。
120000	天津市	
130000	河北省	
...	
710000	台湾省	
810000	香港特别行政区	
820000	澳门特别行政区	

6.14 币种代码表

采用国标《表示货币和资金的代码》(GB/T 12406-2008)，部分代码示意代码如下：

代码	中文名称	说明
ADP	安道尔比塞塔	完整代码表见国标。
AED	UAE 迪拉姆	
AFA	阿富汗尼	
ALL	列克	
AMD	亚美尼亚达姆	
...	...	
ZWD	津巴布韦元	

以上引用国标的完整代码表可登录征信系统查询及下载。

7 通用校验规则

7.1 规则说明

为了保证数据质量，征信系统在加载数据报送文件的过程中需要进行数据校验。数据校验分为 4 个层次，分别是文件校验、信息记录校验、信息段校验、数据项校验。各层次主要校验内容如下。

- 文件校验：主要针对报送文件格式、文件名要素、文件头要素进行验证。
- 信息记录校验：主要针对信息记录构成、信息记录之间关系进行验证。
- 信息段校验：主要针对信息段构成、信息段之间关系进行验证。
- 数据项校验：主要针对信息段内部及信息段之间的数据项的取值进行验证。

为了提前发现问题，数据提供机构在报送文件前，需要依据一部分校验规则对文件格式、记录内容进行预校验。

为了方便解读和引用，定义了 8 位校验规则编码。编码规则如下：

- 第 1 位代表校验规则的 4 个层次。“F”代表文件校验，“R”代表信息记录校验，“S”代表信息段校验，“I”代表数据项校验。
- 第 2-4 位代表信息记录类型，引用信息记录类型代码。其中，通用校验规则用“000”表示。
- 第 5 位为数字 0 或 1。“0”代表此规则需要数据提供机构报送文件前进行校验，“1”代表此规则仅在征信中心端进行校验。
- 第 6-8 位为顺序码，保证编码的唯一性。

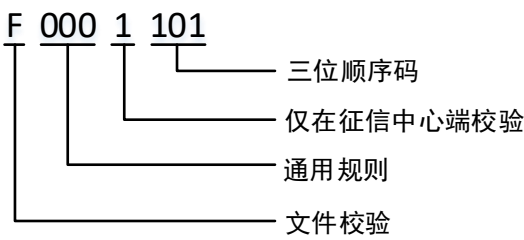


图 7-1 校验规则编码示意

需要注意的是，这里的编码仅仅是为便于引用，并不是反馈文件中的“出错信息代码”。

7.2 通用校验规则

本节校验规则适用于《人民银行征信系统数据采集规范》中所有部分。
对一个文件进行校验时，需要结合通用校验规则与其专用校验规则共同校验。

表 7-1 通用文件检验规则

规则编码	校验规则
	文件名组成校验
F0000201	文件名前缀必须符合文件名命名规则。 <ul style="list-style-type: none"> ➤ 长度 30 位； ➤ 各位置填写的内容符合文件名的要求。
F0001202	文件名中“数据提供机构区段码”应在库中存在。
	文件头组成校验
F0000301	文件头应符合文件头要求： <ul style="list-style-type: none"> ➤ 长度应符合要求，不含回车换行符（“\r\n”）； ➤ 各位置填写的内容应符合文件头的要求。
F0001302	信息记录版本号、信息记录模板号必须是该数据提供机构可以使用的。
	文件头与文件名内容一致性校验
F0000401	文件头中的数据项要与文件名已出现的数据项取值保持一致。 <ul style="list-style-type: none"> ➤ 数据提供机构区段码相同； ➤ 反馈标识为“0”时，文件名的“文件生成日期”与文件头“文件生成时间”中的日期相同； ➤ 反馈标识为“0”时，文件名与文件头“所含信息记录类型”相同。
	文件体格式校验
F0000501	文件体中各行包含的字符个数不能超过上限（1000 万）。
F0000502	<ul style="list-style-type: none"> ➤ 文件头与文件体之间应用一个回车换行符（“\r\n”）分隔； ➤ 信息记录之间应用一个回车换行符（“\r\n”）分隔。
	文件间校验
F0000601	文件名必须和系统已经接收的文件不重复。

表 7-2 通用信息记录校验规则

规则编码	校验规则
	记录自校验
R0000101	信息记录的格式必须符合要求： <ul style="list-style-type: none"> ➤ 对于一个信息记录，同一个信息段最多出现一次； ➤ 信息记录中只能包含本记录中应有的信息段、数据项 XML 标签； ➤ 信息记录的类型需要与文件头中的信息记录类型相匹配。
R0000102	信息记录中的信息报告日期必须小于等于文件生成时间。
	记录间校验
R0000201	对于正常情况下的数据报送，报送文件中不应存在标识项相同的信息记录。
R0001203	信息记录入库时，库中不能存在与其标识项完全相同的信息记录。

表 7-3 通用信息段校验规则

规则编码	校验规则
	信息段自校验
S0000101	A 型数据项必须出现在相应的信息段中。
S0000102	信息段不能为空段（信息段为空指两个信息段的 XML 标签中没有任何数据项）。

表 7-4 通用数据项校验规则

规则编码	校验规则
	空值约束
I0000101	M 型数据项不能为空值且不能为空格（含全角和半角空格）。
	数据类型校验
I0000201	记录中数据项取值（指去掉前后的空值）应符合类型规定。 ➤ 对于枚举型数据项，取值需要在代码表中； ➤ 对于非枚举型数据项需要长度合规，类型相符。
	数据提供机构校验
I0001301	各记录中包含的“数据提供机构区段码”应在库中存在，且与文件头中的“数据提供区段码”对应同一数据提供机构。
I0001302	客户资料维护机构/业务管理机构应在库中存在且未注销，且应是数据提供机构总部或其分支机构。
	信息主体身份标识校验
I0000401	<p>各类身份标识号码应符合相应的编码规则。</p> <p>➤ 公民身份证号码则应符合如下规则：</p> <ul style="list-style-type: none"> ● 长度必须为 18 位； ● 前 17 位为数字，第 18 位可为数字或字母“X”、“x”； ● 第 7-14 位为出生日期，日期必须在 19010101-20991231 之间； ● 校验位必须满足校验规则（参见国标《公民身份号码》（GB 11643-1999））。 <p>➤ 护照的证件号码应符合如下规则：</p> <ul style="list-style-type: none"> ● 最多 12 位； ● 前 3 位为国籍代码。 <p>➤ “港澳居民往来大陆通行证”的证件号码应符合如下规则：</p> <ul style="list-style-type: none"> ● 长度必须为 9 位或 11 位； ● 第一位为字母“H”或“M”，后 8 位或 10 位必须为数字。 <p>➤ “台湾居民往来大陆通行证”的证件号码应符合如下规则：</p> <ul style="list-style-type: none"> ● 长度必须为 8 位或 10-14 位； ● 8 位证件必须为 8 位数字；10-14 位证件，前 10 位必须为数字，其后若只有 3 位则应为“（”+字母+“）”，若有 4 位则应为“（”+两位数字+“）”。 <p>➤ 组织机构代码应符合如下规则：</p> <ul style="list-style-type: none"> ● 应由 8 位数字（或大写拉丁字母）以及 1 位数字（或大写拉丁字母）组成； ● 校验位应满足校验规则（参见国标《全国组织机构代码编制规则》（GB 11714-1997））。 <p>➤ 统一社会信用代码应符合如下规则：</p> <ul style="list-style-type: none"> ● 应由 18 位数字或大写英文字母（不使用 I、O、Z、S、V）组成； ● 第 1 位和第 2 位应为有效的登记管理部分代码和机构分类代码； ● 第 9-17 位应符合组织机构代码规则； ● 校验位应符合校验规则（参见国标《法人和其他组织统一社会信用代码编码规则》（GB 32100-2015））。 <p>➤ 中征码应符合如下规则：</p> <ul style="list-style-type: none"> ● 编码长度应为 16 位； ● 第 1 到 14 位应为字母或者数字；

规则编码	校验规则
	<ul style="list-style-type: none"> ● 校验位（最后两位）符合校验规则（参见附录 A）。
I0001402	中征码应在库中存在。
	其他数据项与“信息报告日期”校验
I0000501	信息记录中的“信息报告日期”应不早于同一条记录中业务日期（“业务日期”具体见各分册专用校验规则）。
	日期有效性校验
I0000601	<p>所有时间必须是有效时间：1901-01-01T00:00:00—2099-12-31T23:59:59。</p> <p>所有日期必须是有效日期：1901-01-01—2099-12-31。</p> <p>所有月份必须是有效月份：1901-01—2099-12。</p> <p>所有年份必须为有效年份：1901—2099。</p>
	可出现多次数据项组校验
I0000701	对于一组可出现多次的数据项，其出现次数必须与其个数相匹配。
I0000702	对于一组可出现多次的数据项，当出现多次时，内容不能重复。
	标识有效性校验
I0001801	对于已发生标识变更的记录，新上报的记录中不能再包含原业务标识码。

附录 A
(规范性附录)
中征码校验位计算规则

中征码的校验位按照下列公式计算：

$$A = 1 + MOD(\sum_{i=1}^{14} C_i \times w_i, 97)$$

式中：

- MOD ——表示求余函数；
- i ——表示代码字符从左至右位置序号；
- C_i ——表示第 i 位置上的代码字符的值，见表 A-2；
- A ——表示校验位，是由 C₁₅*10+C₁₆得到的数字（十进制）。
- W_i ——表示第 i 位置上的加权因子，其数值如下表：

表 A-1 加权因子

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
W _i	1	3	5	7	11	2	13	1	1	17	19	97	23	29

表 A-2 代码字符的值

代码字符	机器处理用代码字符数值
0	0
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
A	10
B	11
C	12
D	13
E	14
F	15
G	16
H	17
I	18
J	19
K	20
L	21

代码字符	机器处理用代码字符数值
M	22
N	23
O	24
P	25
Q	26
R	27
S	28
T	29
U	30
V	31
W	32
X	33
Y	34
Z	35

附录 B
(资料性附录)
修订说明

B.1 本次修订记录

本次系第 3 次修订，较 2018 年 09 月修订的版本，除文字上的编辑性调整外，主要调整内容如下：

1. 在“4.9 信息交换编码约定”中明确编码方式采用“无 BOM 的 UTF-8 编码”。
2. 调整校验规则 I0000401：
 - a) 取消统一社会信用代码“第 3-8 位为行政区划，代码应为已实施或现行的行政区划国标代码之一”的校验要求；
 - b) 将中征码“第 1 到 3 位应为字母或者数字；第 4-14 位应为数字；”的检验要求调整为“第 1 到 14 位应为字母或者数字”。
3. 调整校验规则 F0000501，将文件体中各行包含的字符个数上限调整为 1000 万。

B.2 历史修订记录

2018 年 6 月**第 1 次修订**，较 2017 年 02 月首次发布的版本，除文字上的编辑性调整外，主要调整内容如下：

1. 增加及明确了如下术语和定义，详见术语和定义章节：
 - a) 增加“还款逾期”；
 - b) 明确组织信息记录时，各信息段最多只应出现一次。
2. 基本约定章节中明确了特殊字符的转义约定；
3. 调整及明确了如下通用的代码表的填报要求，详见常用代码表章节：
 - a) 企业身份标识代码表中删除“07-律师事务所执业许可证号”、“08-司法鉴定许可证号”。
 - b) 报送行业分类时，调整为采用《国民经济行业分类》（GB/T 4754-2017）规定的代码表，包括门类、小类等；
 - c) 明确报送行政区划时，应使用征信系统规定行政区划代码表。
4. 调整了 I0000201、I0001302、I0000401 对应的通用校验要求，详见通用校验规则章节。
5. 更正错误：删除了原数据报送文件中部分记录存在错误时，反馈信息记录 XML 示意（图 5-2）中错误反馈信息中的取值。
6. 明确了中征码校验位的计算规则，详见附录 A。

2018 年 9 月**第 2 次修订**，较 2018 年 06 月修订的版本，除文字上的编辑性调整外，主要调整内容如下：

1. 补充校验规则 I0000701，详见通用校验规则。