

中华人民共和国国家标准

GB/T XXXXX—XXXX

消费品质量安全风险信息采集和处理指南

Guide for collecting and processing of consumer products safety risk
information

征求意见稿

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX - XX - XX 发布

XXXX - XX - XX 实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会 发布

目 次

前言	I
1 范围	II
2 规范性引用文件	1
3 术语和定义	1
4 风险信息采集	2
4.1 采集原则	2
4.2 风险信息源的识别	2
4.3 采集方法	3
4.4 采集程序	3
5 风险信息处理	3
5.1 预处理（筛选）	3
5.2 核心要素提取	3
5.3 风险信息挖掘	4
5.4 存储	4
5.5 共享与发布	4
附录 A（资料性附录） 常见的消费品质量安全风险信息	5
附录 B（资料性附录） 风险信息采集软件的功能示例	7
附录 C（资料性附录） 消费品质量安全风险信息采集案例	8
附录 D（资料性附录） 常见的风险信息挖掘算法	10
附录 E（资料性附录） 风险信息存储软件的功能构成	12
参考文献	13
图 D.1 关联规则应用步骤	10
图 D.2 决策树分类应用步骤	11
图 D.3 文本聚类应用步骤	11
表 A.1 常见的消费品质量安全风险信息	5
表 B.1 风险信息采集软件的功能构成	7
表 E.1 风险信息存储软件的功能构成	12

前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

本标准由全国消费品安全标准化技术委员会（SAC/TC 508）提出并归口。

本标准起草单位：

本标准主要起草人：

消费品质量安全风险信息采集和处理指南

1 范围

本标准给出了消费品质量安全风险信息采集和处理的原则、基本程序和方法。

本标准适用于有关组织开展消费品质量安全风险信息管理的活动，其他产品质量安全风险信息管理亦可参照适用。

本标准是通用指南，相关组织可结合自身情况和实际需要，设定更为具体的目标和信息采集与处理规范，对本标准提供的风险信息采集和处理过程进行必要细化或逐步完善。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅所注日期的版本适用于本文件。凡是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。

GB/T ××××（消费品质量安全风险信息数据项描述规范）

3 术语和定义

下列术语和定义适用于本文件。

3.1

消费品 consumer product

为满足社会成员生活需要而销售的产品。

[GB/T 22760—2008，定义2.1]

3.2

消费品质量安全 consumer product safety

消费品质量在预期使用和可预见的误用情况下符合安全的要求。

3.3

风险 risk

对伤害的一种综合衡量，包括伤害发生的可能性和伤害的严重程度。

[GB/T 28216—2011，定义3.6]

3.4

风险信息 risk information

描述潜在风险（3.3）或已发生的伤害事件的信息。

3.5

风险信息管理 risk information management

为及时掌握风险状况和识别可能存在的潜在风险而开展的风险信息采集、加工和处理的活动。

3.6

风险信息采集 risk information collection

为获得某个主题的风险信息（3.3）而开展的信息搜寻和获取过程。

3.7

风险信息处理 risk information processing

对风险信息的预处理、存储、分析挖掘、共享和发布的活动。

3.8

风险信息挖掘 **risk information mining**

利用数据挖掘算法对一定数量的风险信息进行深入计算和分析,并获取风险信息内在特征和潜在规律的过程。

4 风险信息采集

4.1 采集原则

风险信息的采集应遵循以下原则:

- a) 及时性原则。应及早采集风险信息,尽量在风险信息开始广泛传播之前第一时间采集,不应待风险聚集到一定程度之后才开始采集。
- b) 客观性原则。采集风险信息时应尽量保持信息原貌,不应对采集信息的核心要素进行随意更改。
- c) 合法性原则。风险信息的采集应符合国家法律法规的要求,不应以欺诈、窃取、贿赂、利诱、胁迫、侵入计算机等非法手段获取信息,并保护相关方的合法权益;对涉及到个人或者组织隐私的风险信息,应建立保密制度并进行严格保密,保证信息在可控范围内流转和使用。
- d) 可追溯性原则。应建立相应的编码规则和追溯制度,保留信息源记录。
- e) 全面性原则。应力求信息内容全面,尽量充分运用各种信息采集方法,从多个渠道去采集相应的风险信息,每一渠道也可以跨区域、跨时间去采集。

4.2 风险信息源的识别

4.2.1 总则

识别风险信息源是开展风险信息采集的基础。风险信息源的识别应从风险信息源分类入手,掌握各类风险信息源及风险信息在形成和流通过程中的特征和规律。

4.2.2 信息源角度的风险信息分类

4.2.2.1 按照风险信息的披露或掌握主体不同,消费品质量安全风险信息可分为:

- a) 政府部门发布或掌握的风险信息,主要是国内外与产品质量监管相关的政府部门发布的消费品质量安全风险信息或伤害和事故信息。该类信息具有权威度高、影响力大的特征。
 - b) 消费者反馈信息,主要是消费者通过电话、网络、信件等方式向相关组织提供的涉及消费品质量安全的风险信息。该类风险信息具有重复率高、主观性强、复杂多样的特征。
 - c) 企业掌握的风险信息,主要是企业在产品设计、生产、流通等过程中发现的消费品质量安全风险信息。
 - d) 相关组织发布或掌握的风险信息,包括国内外各种科研院所、检测技术机构、学术团体、医院、伤害事故鉴定机构、消费者维权机构以及专业媒体机构等组织掌握和发布的风险信息。
- 常见的消费品质量安全风险信息源见附录A。

4.2.2.2 根据风险信息源的开放状态即信息是否对外公开发布,消费品质量安全风险信息可分为:

- a) 公开的风险信息。主要是媒体信息和公告(通告)信息,即通过互联网、报纸、电视、广播、广告等各类媒体发布或以海报、公告、通告等形式对外公开发布的信息。
- b) 封闭的风险信息。未对外公开发布,仅在组织内部或由少数个人掌握的信息。

注：一般来说，公开风险信息更容易采集到，但封闭的风险信息可能因为信息尚未广泛传播更具有价值。

4.3 采集方法

风险信息采集方法包括但不限于：

- a) 调查法。对于掌握某类风险信息的个人或组织，有针对性地通过填写问卷、访谈等方式获取信息，包括原始信息和二次信息。
- b) 信息抓取工具法。对于在互联网上免费公开的风险信息，可以利用搜索引擎人工搜索和采集信息，或者利用自动采集工具软件，自动搜寻和采集相关风险信息。为提高采集信息的准确率，可在采集前建立关键词库。附录 B 给出了互联网风险信息采集软件的功能示例。
- c) 定向采集法。某一渠道的风险信息对本组织非常重要时，可通过与该渠道的相关机构建立起合作机制，以定向获取相关风险信息。定向采集信息时，宜事先设置好信息采集的格式和要求。
- d) 实验室检测验证法。对于某类不确定的风险信息需要利用技术手段进行验证和确认时，可由相关检测机构或实验室按照相关标准要求检测或模拟伤害发生环境进行试验验证。

4.4 采集程序

采集风险信息，首先要根据本组织实施风险信息管理的需要，整理出具体的风险信息源（清单），再针对不同来源的信息，确定不同的采集策略和方法，然后结合方法的便利程度和所需信息的紧急程度，以人工或借助相关计算机软件（平台），开展针对性的风险信息采集。

鉴于媒体信息本身已集成各类风险信息源，且具备信息量大、更新快等特点，各组织开展风险信息采集时，应重视对媒体信息尤其是互联网信息的采集。采集互联网信息，可以人工采集，有条件的宜使用自动采集工具来采集。

风险信息采集案例见附录 C。

5 风险信息处理

5.1 预处理（筛选）

风险信息的预处理（筛选）包括剔除重复信息、与本组织不相关的信息以及明显虚假的信息。

信息是否与本组织有关，由本组织开展风险信息管理的目的决定。一般情况下，不涉及产品质量安全的信息可以直接剔除。

采集到的风险信息可能包括虚假信息。在预处理阶段，需要根据风险信息所包含的内容，进行初步的筛选，剔除基于常识或一定的评估规则时可以认定为虚假的信息或其他干扰信息。风险信息的辨伪即识别虚假信息是一个复杂的过程，它贯穿于风险信息管理的始终而并不专属于预处理阶段，如采集过程中也会主动避免对部分干扰信息的采集，但预处理阶段的风险信息辨伪被认为是典型的和不可缺少的。

示例：利用采集软件采集互联网上消费者反馈的信息时，往往会设置一定的过滤或研判规则，若出现同一个IP地址在短时间内发布大量带有攻击性的消费品质量安全风险信息，可根据规则判定为非合理的恶意攻击信息并直接剔除。

5.2 核心要素提取

对于每一条风险信息，都要进行信息识别与分析处理，提取出核心要素，包括风险基础信息、危害因素信息以及伤害结果信息等。具体的提取要素参见GB/T ××××（消费品质量安全风险信息数据项描述规范）。

一般情况下，采集到的风险信息并不具备全部的核心要素，此时，要素提取工作是将已包含的要

素提取出来，同时明确还欠缺哪些要素。

5.3 分析和挖掘

经过一定时期的积累，组织一般会积累大量的风险信息。需要对这些风险信息进行深入分析和挖掘，才能把握信息的内在特征和规律。风险信息分析和挖掘应考虑但不限于以下方面：

——应建立风险信息分析的目标，明确分析的对象；

——对于分析已发生风险，可利用一些统计分析方法去实现，且应从多个维度分析，如时间维度、信息来源维度、区域维度等；

——风险信息分析可利用一些常用的统计分析方法，也可利用数据挖掘算法实现。附录D给出了几种常见的消费品质量安全风险信息挖掘的算法。

5.4 存储

采集到的风险信息需要进行储存，风险信息储存时应考虑但不限于以下方面：

——风险信息储存应建立持续性机制，保证风险信息能够不间断存取，并能随时获取历史某一时间段同类主题的信息；

——应实现两级储存模式，第一级储存是在刚采集到信息时开展的存储活动，第二级储存是在经过风险信息预处理、核心要素提取环节之后开展的存储活动；

——为方便检索和使用，可建立风险信息存储软件，其数据库的数据项设置可参考GB/T ××××（消费品质量安全风险信息描述规范），附录E给出了风险信息存储软件的功能构成。

5.5 共享与发布

一般的风险信息经过处理均可以在组织内共享，也可经过一定程序后对外发布，使利益相关方能及时了解风险信息，以便及时跟踪或采取其他进一步的措施。有些风险信息可能需要经过更为复杂的信息辨伪、加工处理和风险评估等环节才能实现共享和发布。

附 录 A
(资料性附录)
常见的消费品质量安全风险信息

表 A.1 给出了常见的消费品质量安全风险信息来源和存在形式。

表 A.1 常见的消费品质量安全风险信息

序号	信息类别	主要信息源或存在形式	备注
1	国内政府部门发布或掌握的风险信息	质检、工商部门发布的产品质量监督抽查和商品监测信息	公开信息
		农业、卫生、环保等部门发布或掌握的消费品质量安全风险信息	公开信息
		质检、工商等产品质量监管部门掌握的消费者举报投诉信息中的消费品质量安全风险信息	
		质检、工商等产品质量监管部门在开展产品质量监管活动中掌握的消费品质量安全风险信息	
		质检、工商、卫生、消防等部门开展打假活动发现的风险信息	公开信息
		质检总局发布的缺陷产品召回信息	
		海关和进出口检验检疫部门在进出口商品检验过程中发现的消费品质量安全风险信息	
		卫生和质检等部门掌握的产品伤害监测信息	
		公安、司法和交通管理等部门掌握的交通事故信息	
		公安、环保、安全生产等其他政府部门和法院在工作过程中掌握的消费品质量安全风险信息	
2	消费者反馈信息	消费者通过当地 12365、12315 和 315 等热线或平台向质检、工商等产品质量监管部门投诉的信息	与政府掌握的举报投诉信息同
		消费者向消费者维权机构或主流媒体反映的投诉信息	
		消费者在专门的投诉或维权网站上自行发布的质量申诉或产品投诉等信息	
		消费者向产品生产商和经销商等投诉的信息	
		消费者通过微博等方式发布的涉及消费品质量安全方面的反馈信息	公开信息
3	企业掌握的风险信息	生产企业内部如设计、生产部门发现的产品质量安全风险信息	
		企业在销售或售后服务过程中主动发现的消费品质量安全风险信息	
		产品的分销商、批发商、零售商等发现的消费品质量安全风险信息	
		企业通过行业交流或与其他企业的经济往来中发现风险信息	
4	国内相关组织发布或掌握的风险信息	行业协会或维权机构在日常工作中监测到的风险信息	
		产品检测技术机构在日常产品质量检测（不含受政府部门委托的抽查检测）过程中发现的消费品质量安全风险信息	
		产品质量监管部门、检测技术机构、企业和第三方机构等开展消费品质量安全主动风险监测获取的风险信息	
		科研院所的实验室在实验过程中发现的风险信息	
		医院在为消费者治疗过程中记录的风险信息	
		司法鉴定等机构掌握的交通等伤害事故信息	
		其他组织掌握的伤害信息或消费品质量安全风险信息	

序号	信息类别	主要信息源或存在形式	备注
5	境外信息	境外政府机构发布的消费品质量安全风险信息，例如美国消费品安全委员会（CPSC）和美国食品药品监督管理局（FDA）发布的召回等警示信息，欧盟委员会（EC）通过 RASFF 和 RAPEX 系统发布的食品和非食品消费品警示信息	公开信息
		境外非政府组织公开发布或向我国通报的有关产品预警通报的风险信息，例如国际绿色和平组织、美国消费者利益委员会（ACCI）、美国消费者联合会（ACU）等发布的消费品质量安全风险信息。	公开信息
6	媒体信息	互联网、电视、报纸、广播、期刊（含学术期刊）、广告等境内外媒体发布的风险信息	公开信息，且可能与其他各类信息重复
<p>注 1：为清晰列举常见的各类风险信息源，这里的信息类别列举与标准中 4.2.2.1 条的信息分类有所不同，主要是将境外信息和媒体信息独立列举出来。</p> <p>注 2：将媒体信息作为一类特殊的信息源列举，但并不意味着它是与前几种信息并列的一种信息分类方式，因为各种信息都可以通过媒体发布而成为媒体信息。在信息化时代，媒体信息具有信息量大、更新和传播快等特点，它也是消费品质量安全风险信息采集的重点之一。</p>			

附 录 B
(资料性附录)
风险信息采集软件的功能示例

表 B.1 给出了互联网风险信息采集软件的功能示例。

表 B.1 风险信息采集软件的功能构成

序号	名称	功能说明
1	网址管理	管理（增加、删除、修改）和配置拟采集网络信息的网址，使得软件能够自动获取网址所对应网页的正文内容。
2	采集规则设置	对采集的时间周期、采集频次、存储位置等关键参数进行设置。
3	关键词管理	配置符合行业主题的关键词以及关键词之间的逻辑表达式，使得只有满足关键词搭配规则的信息才被采集。
4	信息存储	可直接存入风险信息存储软件中，具体操作参考本标准附录 E。
5	风险信息挖掘	嵌入分类、聚类数据挖掘算法，将获得的大量信息初步进行分析。

附录 C
(资料性附录)
消费品质量安全风险信息采集案例

C.1 背景

A 组织受该市政府委托，负责消费品质量安全风险信息采集和管理的工作。

C.2 日常风险信息的采集

C.2.1 列举风险源清单

A 组织根据本标准附录 A 分析研究的常见的消费品质量安全风险信息源和存在形式，列举出需要采集的风险信息源（清单），包括 480 种主要媒体、本市 12365 和 12315 热线（消费者协会）积累的投诉举报信息、历年质监部门的监督检查数据信息、工商部门的商品监测数据信息、从事产品伤害监测的 1 家医院、公安交警大队的交通事故信息等。

C.2.2 确定风险信息采集的方法

根据以上清单，确定要借助采集软件来采集互联网信息，通过定向采集法采集质监、工商、消协、交警大队等政府部门的信息。另外，拟通过调查法征集本市产品质量安全专家有关特定领域的消费品质量安全风险信息，在一些专题领域，拟委托专业检测机构来进行风险信息的验证，获取具体的风险信息。

C.2.3 开发风险信息采集软件

组织开发互联网消费品质量安全风险信息采集软件，选择了 480 个主流媒体网站（网址）进行信息采集（监控），设定了采集规则和关键词。

C.2.4 建立定向采集机制

与质监、工商、消协、交警大队等部门沟通，在对方同意的基础上，建立定向采集机制，即通过定向采集方法获取监督检查、商品监测、消费者 12315 投诉及交通事故信息。

C.2.5 采集风险信息

在以上基础上，A 组织开始风险信息采集工作，全面采集各类消费品质量安全风险信息，并在信息采集基础上开展信息处理和报送等工作。

C.3 专题风险信息的采集

某日，该市突然发生一起幼儿因服装纽扣窒息住院事件，相关服装为本市一家企业生产。此时，A 组织要优先采集与婴幼儿和儿童服装安全（尤其是附件安全）相关的风险信息。A 组织列出了其专项风险信息采集的要解决的问题及信息源清单：

- a) 以往政府产品质量监管时是否抽检信息；
- b) 国内外有关标准要求；
- c) 相关实验室对婴童服装附件指标检测日常记录情况；
- d) 本市及全国消协中的类似的投诉记录；
- e) 本市风险监测定点医院中的类似伤害记录；
- f) 国内外有关类似案例的报道或文献记载（如国外通报）；

- g) 产品质量检测等质量领域专家对产品质量问题和该事件的看法;
- h) 事件中幼儿家庭等当事人反馈信息及医治医院医护人员的看法。

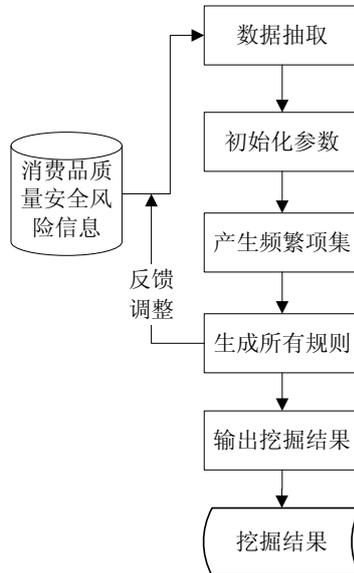
根据以上清单, A 组织选择相应的风险信息采集方法, 采取了以下措施:

- a) 从采集软件中的数据库及互联网上搜寻类似案例, 后者包括从新闻报道和专业技术文献中查找案例及专家观点等;
- b) 从以往积累的产品质量监督抽查和商品监测数据中找出婴童服装抽检数据信息;
- c) 查阅国内外相关产品标准和技术法规的对婴童服装附件要求的规定;
- d) 委托本市一家权威检测机构, 从市场上购买婴童服装并按照相关标准要求检测;
- e) 查阅本市消协的有关婴童服装伤害投诉记录, 并致电全国消协查询全国类似伤害记录;
- f) 查阅本市定点检测医院的产品伤害检测数据;
- g) 咨询或采访部分产品检测专家的意见;
- h) 访谈事件中幼儿家庭等当事人及医治医院医护人员, 了解经过及其看法;
- i) 获取其他信息作补充和印证。

附录 D
(资料性附录)
常见的风险信息挖掘算法

D.1 关联规则法

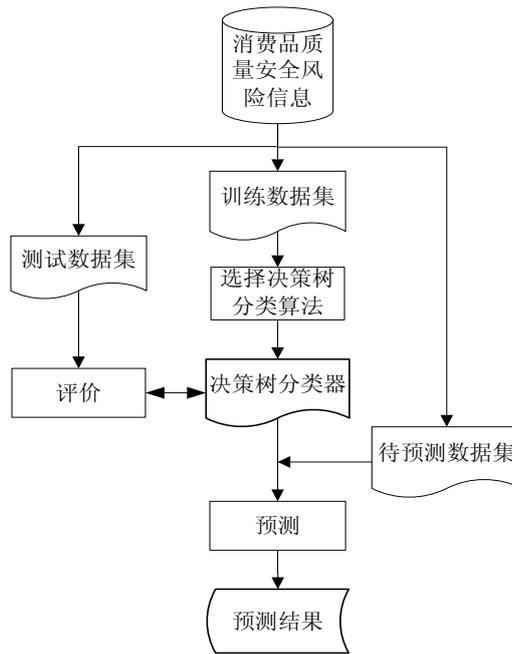
关联规则方法用于发现大量数据中项集之间的关联或相关关系，其挖掘效率取决于挖掘算法。利用关联规则方法对大量的消费品质量安全风险信息进行处理，获取关联规则，挖掘与消费品产生质量安全问题有关的因素之间的联系，通过诊断及时发现消费品质量安全隐患并预警，可以有效监管消费品生产、销售全过程，做到防患于未然，降低消费品质量安全风险。图D.1给出了关联规则的应用步骤。



图D.1 关联规则应用步骤

D.2 分类法

分类是把数据项映射到其中一个事先定义的类中的一个学习函数的过程。由一组输入的属性值向量和相应的类，用基于归纳学习算法得出分类。决策树和决策规则是解决实际应用中分类问题的一种数据挖掘方法。决策树算法的根本是构建决策树分类器。根据训练数据集训练得到决策树分类器，并通过测试数据集评价决策树分类器的分类效果。若分类的准确率在可接受的范围之内，则可利用该分类器进行预测。图D.2给出了决策树分类的应用步骤。

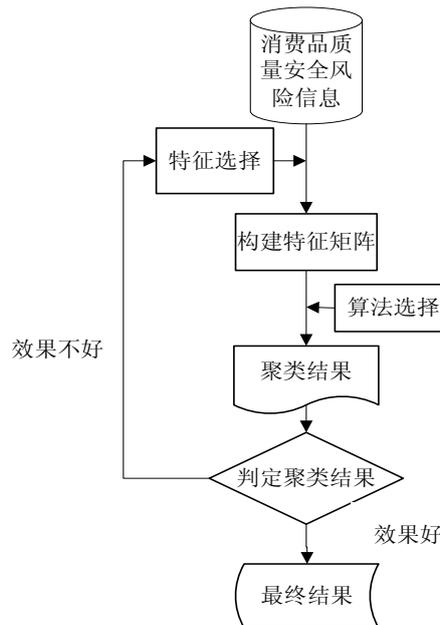


图D.2 决策树分类应用步骤

D.3 聚类方法

聚类是一个将数据集划分为若干组或类的过程，并使得同一个组内的数据对象具有较高的相似度，而不同组内的数据对象则是不相似的。相似或不相似的程度是基于数据对象描述属性的取值来确定的。聚类方法与前述分类方法不同，主要在于聚类所要求划分的类是未知的。

利用聚类方法，可以从大量的消费品质量安全风险信息中把相似的数据集归类成同一组，便于深度分析。图D.3给出了文本聚类的应用步骤。



图D.3 文本聚类应用步骤

附 录 E
(资料性附录)
风险信息存储软件的功能构成

表 E.1 给出了风险信息存储软件的功能构成示例。

表 E.1 风险信息存储软件的功能构成

序号	名称	功能说明
1	基本设置	对软件中用到的常用字典库进行设置,如消费品分类、伤害类型、伤害程度、伤害时的活动、物理类因子、化学类因子、生物类因子等。
2	风险信息添加/编辑/删除	对风险信息的管理和维护,包括添加、编辑和删除风险信息。
3	风险信息审核	需要由组织相关人员对填报的信息进一步审核,以确保信息的真实有效。
4	风险信息统计	从多个维度对 ([信息进行统计,包括消费品维度、地点维度、时间维度、因子维度等。
5	风险信息查询	可根据多个字段查询风险信息,包括消费品种类、地点、时间、安全因子等。

参考文献

- [1] GB/T 20134—2006 道路交通信息采集 事件信息集
- [2] GB/Z 20986—2007 信息安全技术 信息安全事件分类分级指南
- [3] GB/T 22118—2008 企业信用信息采集、处理和提供规范
- [4] GB/T 22760—2008 消费品安全风险评估通则
- [5] GB/T 23792—2009 信用标准化工作指南
- [6] GB/T 28216—2011 消费品质量安全因子评估和控制 通则