

团 体 标 准

T/ISC 0033—2023

互联网信息推荐技术影响性评估指南

Impact assessment guidelines for Internet information recommendation
technology

2023 - 11 - 15 发布

2023 - 12 - 15 实施

中 国 互 联 网 协 会 发 布

目 次

前言	II
1 范围	1
2 规范性引用文件	1
3 术语、定义和缩略语	1
3.1 术语和定义	1
3.2 缩略语	2
4 推荐技术影响评估框架及流程	2
4.1 评估原则	2
4.2 推荐技术影响要素	3
4.3 推荐技术影响分析原理	3
4.4 推荐技术影响评估流程	4
5 推荐技术影响评估方法	4
5.1 评估模型概述	5
5.2 算法导向性	5
5.3 算法客观性	7
5.4 算法风险性	8
5.5 茧房系数	9
5.6 收敛系数	10
6 推荐技术影响评估实施	11
6.1 影响评估准备	11
6.2 影响要素识别和计算	12
6.3 影响结果判定	12
6.4 影响评估文档记录	13
6.5 影响风险处理	14
7 影响评估工作形式	14
7.1 自评估	14
7.2 检查评估	14

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这些专利的责任。

本文件由中国互联网协会提出并归口。

本文件起草单位：国家计算机网络应急技术处理协调中心、长安通信科技有限责任公司、浙江工业大学、北京师范大学、北京工业大学、深圳市腾讯计算机系统有限公司、北京百度网讯科技有限公司。

本文件主要起草人：韩晗、刘美辰、张夏、王鲁华、赵芸伟、崔牧凡、魏海潇、闵勇、刘茜、李童、武杨、李克鹏、许锐。

互联网信息推荐技术影响性评估指南

1 范围

本文件确立了互联网信息推荐技术影响评估框架及流程，描述了评估方法、评估实施步骤及评估工作形式。

本文件适用于互联网内容提供商、第三方评估机构以及监管部门等开展互联网信息推荐技术影响评估工作。

2 规范性引用文件

本文件没有规范性引用文件。

3 术语、定义和缩略语

3.1 术语和定义

下列术语和定义适用于本文件。

3.1.1

互联网信息推荐技术 Internet information recommendation technology

互联网内容提供商向用户提供互联网信息服务时应用个性化推送类、排序精选类、检索过滤类等算法进行信息推荐的技术。

3.1.2

互联网信息推荐技术影响评估 Internet information recommendation technology impact assessment

运用科学的方法和手段，对使用互联网信息推荐技术可能引发的社会公众影响进行系统地识别、分析和评价。

3.1.3

评估对象 assessment object

被评估的主体，本标准指采用互联网信息推荐技术进行信息推荐的互联网内容提供商。

3.1.4

互联网内容提供商 Internet content provider

向广大用户综合提供互联网信息业务和增值业务的运营商。

3.1.5

目标数据 target data

经评估对象授权许可用于影响评估的限定业务范围和限定时间内实际运营产生的数据集合。

3.1.6

评估模型 evaluation model

通过构建评估指标并明确指标量化方式，形成适用于评估算法推荐对社会公众的影响的模型。

3.1.7

信息茧房 information cocoon

基于用户兴趣和习惯，将更多相同类型内容分发给用户，造成用户获取信息内容狭隘、单一的现象。

3.1.8

时间茧房 time cocoons

通过量化用户在连续时间段内的行为特征评价信息茧房风险的一种方式。

3.1.9

访问茧房 access cocoons

通过量化用户在连续访问次数内的行为特征评价信息茧房风险的一种方式。

3.2 缩略语

下列缩略语适用于本文件。

ICP 互联网内容提供商 (Internet content provider)

4 互联网信息推荐技术影响评估框架及流程

4.1 评估原则

4.1.1 标准化原则

互联网信息推荐技术影响评估应参照本文件确立的评估流程和方法进行实施。

4.1.2 可控性原则

在影响评估实施过程中，应严格按照标准的项目管理方法对服务过程、人员和工具等进行控制，保证影响实施过程中的安全可控。

- a) 服务可控性。评估方应在评估工作前介绍评估服务流程，明确需要得到评估对象协调配合的工作内容，确保评估工作顺利开展；
- b) 人员与信息可控性。所有参与评估的人员应签署保密协议，应对评估过程数据和结果数据严格管理，未经授权不得泄露给任何单位和个人；
- c) 过程可控性。成立评估实施团队，采取项目组长负责制，应对评估工作进行记录并形成相关记录文档，达到项目过程的可控；
- d) 工具可控性。评估使用的评估工具应事先通告用户，并在评估实施前获得用户的许可，包括工具软件、测试策略等。

4.1.3 最小影响原则

对于在线业务系统的影响评估，应采用最小影响原则，对于影响评估所需要的数据，需与用户沟通并进行应急备份。

4.2 影响要素

互联网信息推荐技术影响评估围绕算法导向性、算法客观性、算法风险性、茧房系数、收敛系数等影响要素展开。

- a) 算法导向性。算法导向性评价互联网信息推荐算法的舆论导向影响，以政治导向和推荐主题情感导向综合评价；
- b) 算法客观性。算法客观性评价互联网信息推荐算法的推荐客观程度，以推荐内容商业化程度、推荐内容类型多样性程度、推荐内容所属账号分散程度综合评价；
- c) 算法风险性。算法风险性评价互联网信息推荐算法的推荐传播风险，以推荐内容潜在传播风险和影响范围综合评价；
- d) 茧房系数。茧房系数评价互联网信息推荐算法推荐分发的信息内容导致用户形成信息茧房的可能性程度，以时间茧房系数和访问茧房系数综合评价；
- e) 收敛系数。收敛系数评价互联网信息推荐算法推荐分发的信息内容导致用户形成信息茧房的可能用时，以时间茧房收敛极值和访问茧房收敛极值综合评价。

4.3 影响分析原理

推荐技术影响分析原理如图1所示：

推荐技术影响分析涉及算法导向性、算法客观性、算法风险性、茧房系数、收敛系数等基本要素。各要素包含的具体指标及量化方法见表1—推荐技术影响评估指标量化表。

影响分析的主要内容为：

- a) 经评估对象授权，获取待评估的目标数据；
- b) 在目标数据范围内识别影响要素，对目标数据进行必要的预处理；
- c) 对已识别的影响要素开展对应各指标的影响量化；
- d) 综合各指标的影响量化结果；
- e) 依据综合量化结果划分影响分级。

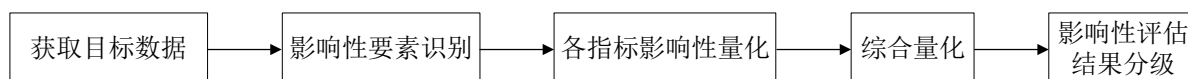


图1 推荐技术影响评估原理图

4.4 影响评估流程

影响评估实施划分为评估准备、推荐技术影响要素识别计算、影响结果判定与影响风险处理四个阶段。其中，评估准备阶段工作是有效实施评估的保证；影响要素识别计算阶段是针对评估活动中的各类影响要素进行识别和量化；影响结果判定阶段是对识别计算阶段中获得的各类影响要素量化结果进行综合判定，整体评估推荐技术影响；影响风险处理阶段针对评估结果，提出相应的整改处置建议，按照处置建议实施算法调整改进。

推荐技术影响评估流程如图2所示：

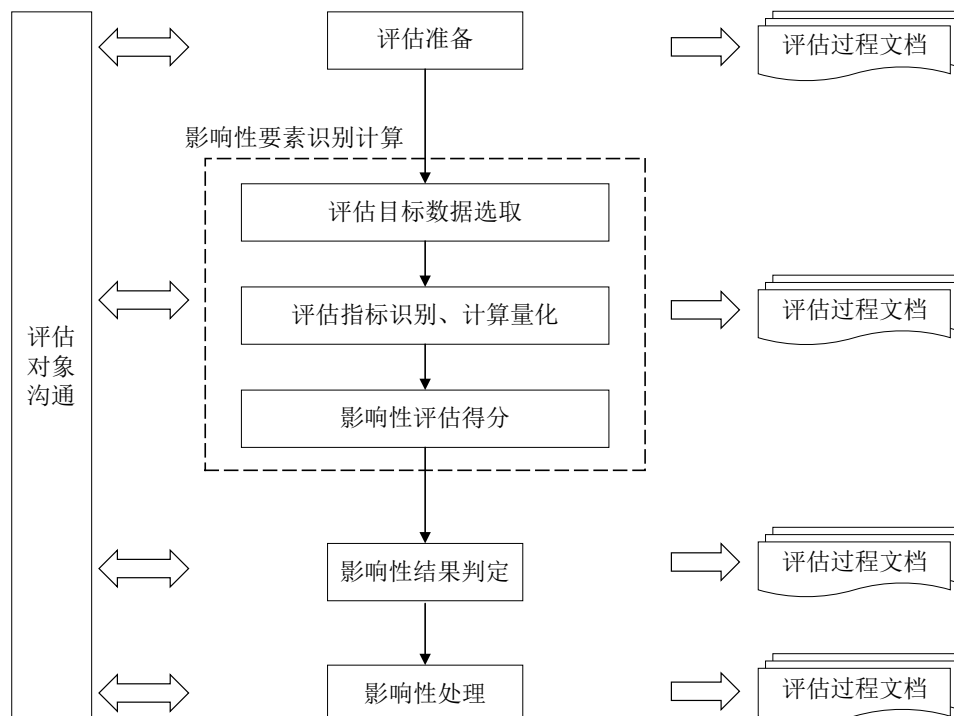


图2 影响评估流程图

5 互联网信息推荐技术影响评估方法

5.1 评估模型概述

推荐技术影响评估模型由推荐算法导向性、推荐算法客观性、推荐算法风险性、推荐算法茧房系数、推荐算法收敛系数共5项一级指标构成，一级指标下，各分别设置相应二级指标（共11项二级指标），量化因子（共12项量化因子）用于计算各评估指标对应的影响数值，各指标影响数值之和即推荐算法综合影响数值（综合影响数值介于0—100之间，数值大小与影响大小成正比例关系）。

表 1 推荐技术影响评估指标量化表

一级指标	二级指标	量化因子	参数 u_i	权重 w_i
算法导向性 (15%)	政治导向	推送非政务新媒体发布量占比	$\log_2(X_1+1)$	5%
	主题情感导向	推送灾害事故信息占比	$\log_2(X_2+1)$	5%
		推送煽动性信息占比	$\log_2(X_3+1)$	5%
算法客观性 (15%)	推荐商业化程度	推送广告信息占比	$\log_2(X_4+1)$	5%
	推荐内容多样性程度	推送主题类型数量占比	$\log_2(1-X_5+1)$	5%
	推荐内容所属账号分散程度	推送信息所属账号数量占比	$\log_2(1-X_6+1)$	5%
算法风险性 (20%)	传播风险	负面信息数量占比	$\log_2(X_7+1)$	10%
	传播范围	发布负面信息的主体粉丝数占比	$\log_2(X_8+1)$	10%
茧房系数 (25%)	时间茧房	时间茧房系数极差	$\log_2(X_9+1)$	12.5%
	访问茧房	访问茧房系数极差	$\log_2(X_{10}+1)$	12.5%
收敛系数 (25%)	时间茧房收敛极值	最小收敛时间	$\log_2\left(\frac{1}{X_{11}}+1\right)$	12.5%
	访问茧房收敛极值	最小收敛时间	$\log_2\left(\frac{1}{X_{12}}+1\right)$	12.5%
推荐算法综合影响数值= $\sum_{i=1}^{12} u_i \times w_i \times 100$				

5.2 算法导向性

5.2.1 政治导向性

5.2.1.1 量化因子

在推荐技术影响评估数据集（此数据集指任意连续N天内（N一般取7天的整数倍，本标准建议N取35，保证时间跨度的一致性和评估数据的广泛性），由推荐算法推送分发的各采样用户接收信息）中，非政务新媒体发布（非行政机关或非承担行政职能的事业单位及其内设机构发布）的数量占比。

$$X_1 = \frac{C_{\text{非政务}}}{C_{\text{数据集}}}$$

其中，量化因子 X_1 为非政务新媒体发布数量占比； $C_{\text{非政务}}$ 为数据集内非政务新媒体发布数量； $C_{\text{数据集}}$ 为数据集信息总量。

5.2.1.2 影响数值

政治导向影响数值 $V_{\text{政治导向}}$ 由量化因子 X_1 决定，计算关系为：

$$V_{\text{政治导向}} = [\log_2(X_1 + 1)] \times w_1 \times 100$$

其中， $V_{\text{政治导向}}$ 为政治导向影响数值， X_1 为量化因子， w_1 为权重，权重分配情况详见表1。

5.2.2 主题情感导向性

5.2.2.1 量化因子

a) 推送灾害事故信息占比

在推荐技术影响评估数据集（此数据集指任意连续N天内，由推荐算法推送分发的各采样用户接收信息）中，涉及地质灾害、重大事故信息占比。

$$X_2 = \frac{C_{\text{灾害事故}}}{C_{\text{数据集}}}$$

其中，量化因子 X_2 为灾害事故信息占比； $C_{\text{灾害事故}}$ 为数据集内涉及灾害事故信息数量； $C_{\text{数据集}}$ 为数据集信息总量。

b) 推送煽动性信息占比

在推荐技术影响评估数据集（此数据集指任意连续N天内，由推荐算法推送分发的各采样用户接收信息）中，煽动性信息（包括但不限于传播抢购、囤积、抛售、挤兑、恶意预测、虚假宣传、人身攻击、地域歧视等煽动性信息）占比。

$$X_3 = \frac{C_{\text{煽动信息}}}{C_{\text{数据集}}}$$

其中，量化因子 X_3 为煽动性信息占比； $C_{\text{煽动信息}}$ 为数据集内煽动性信息数量； $C_{\text{数据集}}$ 为数据集信息总量。

5.2.2.2 影响数值

主题情感导向影响数值 $V_{\text{主题情感导向}}$ 由量化因子 X_2 、 X_3 决定，计算关系为：

$$V_{\text{主题情感导向}} = [\log_2(X_2 + 1)] \times w_2 \times 100 + [\log_2(X_3 + 1)] \times w_3 \times 100$$

其中， $V_{\text{主题情感导向}}$ 为政治导向影响数值， X_2 、 X_3 为量化因子， w_2 、 w_3 为权重，权重分配情况详见表1。

5.3 算法客观性

5.3.1 商业化程度

5.3.1.1 量化因子

在推荐技术影响评估数据集（此数据集指任意连续N天内，由推荐算法推送分发的各采样用户接收信息）中，推送广告信息数量占比。

$$X_4 = \frac{C_{\text{广告信息}}}{C_{\text{数据集}}}$$

其中，量化因子 X_4 为广告信息数量占比； $C_{\text{广告信息}}$ 为数据集内广告信息数量； $C_{\text{数据集}}$ 为数据集信息总量。

5.3.1.2 影响数值

商业化程度影响数值 $V_{\text{商业化程度}}$ 由量化因子 X_4 决定，计算关系为：

$$V_{\text{商业化程度}} = [\log_2(X_4 + 1)] \times w_4 \times 100$$

其中， $V_{\text{商业化程度}}$ 为商业化程度影响数值， X_4 为量化因子， w_4 为权重，权重分配情况详见表1。

5.3.2 推荐内容多样性程度

5.3.2.1 量化因子

在推荐技术影响评估数据集（此数据集指任意连续N天内，由推荐算法推送分发的各采样用户接收信息）中，任意采样用户推荐列表中的主题类型数量在该数据集内全部主题数量中的占比。

$$X_5 = \frac{C_{\text{任意用户主题类型}}}{C_{\text{主题类型}}}$$

其中，量化因子 X_5 为任意采样用户接收的主题类型占比； $C_{\text{任意用户主题类型}}$ 为任意采样用户推荐列表中的主题类型数量， $C_{\text{主题类型}}$ 为数据集内全部信息主题类型数量。

5.3.2.2 影响数值

多样性程度影响数值 $V_{\text{多样性程度}}$ 由量化因子 X_5 决定，计算关系为：

$$V_{\text{多样性程度}} = [\log_2(1 - X_5 + 1)] \times w_5 \times 100$$

其中， $V_{\text{多样性程度}}$ 为多样性程度影响数值， X_5 为量化因子， w_5 为权重，权重分配情况详见表1。

5.3.3 推荐内容所属账号分散程度

5.3.3.1 量化因子

在推荐技术影响评估数据集（此数据集指任意连续N天内，由推荐算法推送分发的各采样用户接收信息）中，任意采样用户接收推送账号数量在该数据集内全部推送账号数量中的占比。

$$X_6 = \frac{C_{\text{任意用户接收账号数量}}}{C_{\text{账号数量}}}$$

其中，量化因子 X_6 为任意采样用户接收的账号数量占比； $C_{\text{任意用户接收账号数量}}$ 为任意采样用户接收推送的账号数量， $C_{\text{账号数量}}$ 为数据集内全部推送账号数量。

5.3.3.2 影响数值

账号分散程度影响数值 $V_{\text{账号分散程度}}$ 由量化因子 X_6 决定，计算关系为：

$$V_{\text{账号分散程度}} = [\log_2(1 - X_6 + 1)] \times w_6 \times 100$$

其中， $V_{\text{账号分散程度}}$ 为账号分散程度影响数值， X_6 为量化因子， w_6 为权重，权重分配情况详见表1。

5.4 算法风险性

5.4.1 传播风险

5.4.1.1 量化因子

在推荐技术影响评估数据集（此数据集指任意连续N天内，由推荐算法推送分发的各采样用户接收信息）中，负面信息（涉黄、涉暴、涉恐、涉赌、涉诈、反宣以及其他对社会稳定造成负面影响的信息）数量占比。

$$X_7 = \frac{C_{\text{负面信息}}}{C_{\text{数据集}}}$$

其中，量化因子 X_7 为非负面信息占比； $C_{\text{负面信息}}$ 为数据集内负面信息数量； $C_{\text{数据集}}$ 为数据集信息总量。

5.4.1.2 影响数值

传播风险影响数值 $V_{\text{传播风险}}$ 由量化因子 X_7 决定，计算关系为：

$$V_{\text{传播风险}} = [\log_2(X_7 + 1)] \times w_7 \times 100$$

其中， $V_{\text{传播风险}}$ 为传播风险影响数值， X_7 为量化因子， w_7 为权重，权重分配情况详见表1。

5.4.2 传播范围

5.4.2.1 量化因子

在推荐技术影响评估数据集（此数据集指任意连续N天内，由推荐算法推送分发的各采样用户接收信息）中，发布负面信息的主体粉丝数占比。

$$X_8 = \frac{C_{\text{负面信息主体粉丝数}}}{C_{\text{数据集主体粉丝总量}}}$$

其中，量化因子 X_7 为非负面信息占比； $C_{\text{负面信息主体粉丝数}}$ 为数据集内负面信息主体粉丝数； $C_{\text{数据集主体粉丝总量}}$ 为数据集包含的所有账号主体的粉丝总量。

5.4.2.2 影响数值

传播范围影响数值 $V_{\text{传播范围}}$ 由量化因子 X_8 决定，计算关系为：

$$V_{\text{传播范围}} = [\log_2(X_8 + 1)] \times w_8 \times 100$$

其中， $V_{\text{传播范围}}$ 为传播范围影响数值， X_8 为量化因子， w_8 为权重，权重分配情况详见表1。

5.5 茧房系数

5.5.1 时间茧房

5.5.1.1 量化因子

在推荐技术影响评估数据集（此数据集指任意连续N天内，由推荐算法推送分发的各采样用户接收信息）中，计算任意采样用户u在第t天的时间茧房系数 $Co_t(r, u)$ 。

$$Co_t(r, u) = \text{norm}\left(\sum_1^K ((P_k^u)^2)\right)$$

其中，t—第t天；r—recommenders，即目标互联网内容提供商的特定推荐算法；u—目标用户。

$P_k^u = \text{Count}_k^u / \text{Count}_{\text{Total}}^u$ 为用户u在第t天接收的k种类型推荐信息的概率；式中 Count_k^u 为推荐算法推荐的k种主题类型信息数量， $\text{Count}_{\text{Total}}^u$ 为推荐算法推荐的信息总量。（当k=1时，时间茧房系数为1，表示用户沉溺于某一特定信息类型）

norm—normalize, 将概率平方和进行归一化处理，使得时间茧房系数落在[0-1]数值区间。

$$X_9 = Co_t(r, u)_{\text{max}} - Co_t(r, u)_{\text{min}}$$

其中，量化因子 X_9 为时间茧房系数极差； $Co_t(r, u)_{\text{max}}$ 为时间茧房系数最大值； $Co_t(r, u)_{\text{min}}$ 为时间茧房系数最小值。时间茧房系数极差为连续7天内（在上述任意连续N天内截取连续7天）的时间茧房系数最大值与最小值之差。

5.5.1.2 影响数值

时间茧房影响数值 $V_{\text{时间茧房}}$ 由量化因子 X_9 决定，计算关系为：

$$V_{\text{时间茧房}} = [\log_2(X_9 + 1)] \times w_9 \times 100$$

其中， $V_{\text{时间茧房}}$ 为时间茧房影响数值， X_9 为量化因子， w_9 为权重，权重分配情况详见表1。

5.5.2 访问茧房

5.5.2.1 量化因子

在推荐技术影响评估数据集（此数据集指任意连续N天内，由推荐算法推送分发的各采样用户接收信息，宜包括任意采样用户点击的不少于2500个推送信息）中，计算用户u完成100个点击的访问茧房系数 $Co_c(r, u)$ 。

$$Co_c(r, u) = \text{norm} \left(\sum_{k=1}^K ((P_k^u)^2) \right)$$

其中，c—clicks，即点击次数；r—recommenders，即目标互联网内容提供商的特定推荐算法；u—目标用户。

$P_k^u = \text{Count}_k^u / \text{Count}_{Total}^u$ 为用户u完成100次点击k类型推荐信息的概率；式中 Count_k^u 为用户完成100次推荐信息点击过程中，推荐算法推荐的k种主题类型信息数量， Count_{Total}^u 为推荐算法推荐的信息总量（即100个）。

norm—normalize, 将概率平方和进行归一化处理，使得时间茧房系数落在[0-1]数值区间。

$$X_{10} = Co_c(r, u)_{max} - Co_c(r, u)_{min}$$

其中，量化因子 X_{10} 为访问茧房系数极差； $Co_c(r, u)_{max}$ 为访问茧房系数最大值； $Co_c(r, u)_{min}$ 为访问茧房系数最小值。访问茧房系数极差为用户u完成N个100次推荐信息点击（N宜取5）的访问茧房系数最大值与最小值之差。

5.5.2.2 影响数值

访问茧房影响数值 $V_{访问茧房}$ 由量化因子 X_{10} 决定，计算关系为：

$$V_{访问茧房} = [\log_2(X_{10} + 1)] \times w_{10} \times 100$$

其中， $V_{访问茧房}$ 为访问茧房影响数值， X_{10} 为量化因子， w_{10} 为权重，权重分配情况详见表1。

5.6 收敛系数

5.6.1 时间茧房收敛极值

5.6.1.1 量化因子

分别计算任意采样用户连续N天各天的时间茧房系数，比较时间茧房系数连续增长的时间区间（时间茧房系数在该区间内由最小值连续增长到最大值所用天数），若干时间区间中历时最短的即为最小收敛时间 T_{t-min} ，量化因子取最小收敛时间。

$$X_{11} = T_{t-min}$$

5.6.1.2 影响数值

时间茧房收敛极值影响数值 $V_{时间茧房收敛极值}$ 由量化因子 X_{11} 决定，计算关系为：

$$V_{时间茧房收敛极值} = \left[\log_2 \left(\frac{1}{X_{11}} + 1 \right) \right] \times w_{11} \times 100$$

其中， $V_{\text{时间茧房收敛极值}}$ 为时间茧房收敛极值影响数值， X_{11} 为量化因子， w_{11} 为权重，权重分配情况详见表1。

5.6.2 访问茧房收敛极值

5.6.2.1 量化因子

分别计算任意采样用户u的25个访问茧房系数（单个访问茧房系数是通过100次点击量计算获得，25个访问茧房系数是顺序时间段内的2500次点击分别计算获得），比较访问茧房系数连续增长的时间区间（访问茧房系数从该区间内由最小值连续增长到最大值所用天数），若干时间区间中历时最短的即为最小收敛时间 C_{t-min} ，量化因子取最小收敛时间。

$$X_{12} = C_{t-min}$$

5.6.2.2 影响数值

时间茧房收敛极值影响数值 $V_{\text{访问茧房收敛极值}}$ 由量化因子 X_{12} 决定，计算关系为：

$$V_{\text{访问茧房收敛极值}} = \left[\log_2 \left(\frac{1}{X_{12}} + 1 \right) \right] \times w_{12} \times 100$$

其中， $V_{\text{访问茧房收敛极值}}$ 为访问茧房收敛极值影响数值， X_{12} 为量化因子， w_{12} 为权重，权重分配情况详见表1。

6 互联网信息推荐技术影响评估实施

6.1 影响评估准备

6.1.1 确定目标

根据评估对象在其业务开展过程中使用的信息推荐技术及其应用中产生的互联网信息内容导向、信息茧房、信息安全等方面存在的问题，评估推荐技术对信息传播效果的影响。

6.1.2 组建团队

根据影响评估的工作形式（自评估或检查评估）确定影响评估实施团队。评估实施团队应召开影响评估工作启动会议，做好评估前的表格、文档、检测工具等各项准备工作，进行影响评估技术培训和保密教育，制定影响评估过程管理相关规定。可根据被评估方要求签署保密协议。

6.1.3 前期调研

前期调研是了解评估对象的业务形式、用户规模、信息传播规模、信息类型等基本业务情况的过程，影响评估实施团队应进行充分调研，为影响评估的精准实施奠定基础。前期调研可采取问卷调查、现场面谈相结合的方式。调研内容包括但不限于：

- a) 业务及相关流程，具体管理和技术部门及其相关人员；

- b) 业务采用的相关信息推荐技术；
- c) 主要的业务功能；
- d) 业务数据和信息；
- e) 数据的敏感性；
- f) 支持和使用业务后台的人员；
- g) 评估对象信息安全管理制度；
- h) 相关法律法规。

6.1.4 制定方案

影响评估方案的目的是为后续的影响评估实施活动提供一个总体计划，用于指导实施方开展后续工作。影响评估方案应得到评估对象的确认和认可。影响评估方案内容包括但不限于：

- a) 影响评估工作框架：包括评估目标、评估范围、评估依据等；
- b) 评估团队组织：包括评估团队成员、组织结构、角色、责任等；
- c) 评估工作计划：包括各阶段工作内容、工作形式、工作成果等；
- d) 风险规避：包括保密协议、评估工作环境要求、评估方法、工具选择、应急预案等；
- e) 风险接受：制定影响风险接受准则，根据影响风险接受准则进行影响风险处理；
- f) 时间进度安排：评估工作实施的时间进度安排。

6.1.5 获得支持

上述所有内容确定后，应形成较为完整的影响评估实施方案，得到评估对象的支持、批准。同时，须对评估对象相关人员、评估团队成员等进行传达，在组织范围内就影响评估相关内容进行培训，以明确相关人员在影响评估中的任务和注意事项。

6.2 影响要素识别和计算

参照第5章—推荐技术影响评估方法进行实施。

6.3 影响结果判定

表 2 影响等级划分表

等级	影响分值	标识	描述
5	80 以上	很高	应用的推荐算法对社会公众的具有极大的影响，存在较严重的导向性、主观性、风险性问题，极易导致信息茧房的形成
4	61—80	高	应用的推荐算法对社会公众的具有较大的影响，存在导向性、主观性、风险性问题，存在导致信息茧房的潜在风险

等级	影响分值	标识	描述
3	41—60	中等	应用的推荐算法对社会公众的具有一定的影响，可能存在导向性、主观性、风险性问题，可能存在导致信息茧房的潜在风险
2	21—40	低	应用的推荐算法对社会公众的影响较小
1	0—20	很低	应用的推荐算法对社会公众的影响极小

- a) 影响结果判定为1级、2级的，属低影响等级，推荐算法的影响在可控范围内，风险可接受；
- b) 影响结果判定为3级的，属中等影响等级，推荐算法的影响存在一定风险，可适当采取整改措施加以调整；
- c) 影响结果判定为4级、5级的，属高影响等级，推荐算法的影响风险较高，需采取必要的整改措施加以调整，并对整改后的推荐算法进行复评估。

6.4 影响评估文档记录

6.4.1 影响评估文档记录要求

记录影响评估过程的相关文档，应符合但不限于以下要求：

- a) 确保文档发布前是得到批准的；
- b) 确保文档的更改和现行修订状态是可识别的（有版本控制措施）；
- c) 确保文档的分发得到适当控制；
- d) 防止作废文档的非预期适用，若因任何目的需保留作废文档时，应对这些文档进行适当标识。

6.4.2 影响评估文档

影响评估文档是指在整个影响评估过程中产生的评估过程文档和评估结果文档，包括但不限于：

- a) 影响评估方案：阐述影响评估的目标、范围、人员、评估方法、工作计划和实施进度等；
- b) 影响评估程序：明确评估的目的、职责、过程、相关的文档要求，以及实施本次评估所需的各种信息数据；
- c) 目标数据清单：影响评估所需要的业务数据；
- d) 影响评估报告：对整个影响评估过程和结果进行总结，详细说明评估对象基本情况、影响评估方法、评估对象所采用的信息推荐技术算法的总体表现情况、影响要素识别结果、影响分析和结论等内容；
- e) 影响风险处理计划：对评估结果中影响较大的风险制定影响风险处理计划，选择适当的控制目标及算法调优措施，明确责任、进度，通过对影响较大的要素进行二次评估，以确定所选择调优措施的有效性；

- f) 影响评估记录：根据影响评估程序，要求影响评估过程中的各种现场记录可复现评估过程，并作为产生歧义后解决问题的依据。

6.5 影响风险处理

将影响评估结果与预先制定的影响风险接受准则进行对照，识别存在的可接受风险和不可接受风险。对不可接受的影响风险应根据导致该影响风险的影响要素制定影响风险处理计划。影响风险处理计划中应明确采取的算法调优措施、预期效果、实施条件、进度安排、责任部门等。宜针对不同影响等级开展风险处理：

- a) 影响等级1级：风险完全可控，影响风险极低；
- b) 影响等级2级：风险基本可控，影响风险较低；
- c) 影响等级3级：具有一定的影响风险，宜制定影响风险处理计划；
- d) 影响等级4级：具有较高的影响风险，需制定影响风险处理计划，根据计划进行算法优化和效果跟踪，定期进行影响评估；
- e) 影响等级5级：具有极高的影响风险，需立即制定影响风险处理计划，指定责任部门组织实施相关整改，整改完成后需再次进行影响评估。

7 互联网信息推荐技术影响评估工作形式

影响评估分为自评估与检查评估两种形式。影响评估应以自评估为主，自评估和检查评估相结合、互为补充。

7.1 自评估

自评估是ICP等采用互联网信息推荐技术的单位或组织对其使用的推荐技术进行影响评估。自评估应在本标准的指导下，结合评估对象自身的实际需求进行实施。周期性进行的自评估可以在评估流程上适当简化，重点针对影响波动较大的要素进行识别和评估，以便邻近两次评估结果的对比。若评估对象所采用的信息推荐技术发生重大变化时，应根据本标准进行完整的评估。

自评估可由评估对象自主实施或委托信息安全评估机构支撑实施。由评估对象自主实施的评估可以降低实施费用、提高内部相关人员的信息安全意识，但容易受到组织内部各种因素的影响，其评估结果的客观性易受影响。委托信息安全评估机构实施的评估，过程较规范、评估结果客观性较好，评估对象应对其背景与资质、评估过程与结果的保密要求等方面进行控制。

7.2 检查评估

检查评估是ICP等采用互联网信息推荐技术的单位或组织的上级主管部门或国家有关职能部门开展的影响评估，检查评估可委托信息安全评估机构支持实施。检查评估可以依据本标准的要求，实施完整的影响评估，也可在自评估实施的基础上，对关键环节或重点内容实施抽样评估，包括但不限于：

- a) 自评估实施团队及技术人员审查；
- b) 自评估目标数据审查；
- c) 自评估方法的检查；
- d) 自评估过程文档检查；
- e) 自评估影响风险处理措施审查；
- f) 其他相关内容审查。