

文本审核 产品简介



腾讯云

【 版权声明 】

©2013–2024 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【 商标声明 】

及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

文档目录

产品简介

产品概述

产品优势

应用场景

产品简介

产品概述

最近更新时间：2023-09-22 14:55:02

文本审核（Text Auditing Service，TAS）由腾讯云 [数据万象（Cloud Infinite，CI）](#) 提供，基于腾讯云天御内容安全防护算法实现，为用户提供文本内容安全智能审核服务，可帮助用户有效识别色情低俗、违法违规、恶心头反感等违禁内容，规避运营风险。

监管部门对内容监管日益严格，内容违规可能为平台带来约谈、整顿、罚款、关停等风险，且影响平台使用者体验，接入文本审核服务可以有效杜绝内容违规风险。文本审核服务包括 Web 管理控制台、服务 API，无需开发，您可以快速接入使用。

产品功能

文本审核功能可帮助用户有效检测文本中的违禁内容，可以检测涉黄、违法违规、广告、谩骂等可能令人不适的内容。目前支持机器自动审核和专业团队人工复审，全方位保障平台的安全。

涉黄检测

精准检测色情文本，能检测各种变体，如拼音、拆字、谐音变体等，检测算法结合海量的违规文本数据进行训练建模，具有很好的抗干扰能力。

广告检测

检测垃圾广告内容，包括恶意刷屏广告、微商广告、牛皮癣广告、色情引流广告等，支持检测符号、图标、拼音、表情包等变异手法。

📌 说明

开启文本审核功能后，您可通过控制台方式获取检测标签、违规内容及置信度，可以直接采用置信度高的结果；针对置信度低的结果，您可选择人工复审，腾讯云天御人工审核团队提供专业审核服务，保障内容安全。

产品优势

最近更新时间：2021-11-26 17:09:52

高效可用

服务毫秒级响应，每秒超过万级并发，支持海量的数据样本，可快速动态扩容，灵活应对高并发需求，无需担心性能损耗。

方便易用

支持 API 调用和控制台配置，可视化页面操作，详尽的操作指南，增量内容自动触发，帮助您快速接入使用，免除繁杂的开发成本。

功能丰富

支持文本等多媒体数据，提供丰富全面的模型学习样本，覆盖各个行业数百种违规类型，用户可根据实际业务按需适配。

成本最优

使用文本审核，您无需前期投资，无需人力维护，自动化智能识别为您大幅度降低自建人工审核成本。文本审核按实际使用量收取费用，您无需预先支付任何预留处理的费用。如您的用量较大，我们也提供资源包形式的预付费方式，价格相比后付费更加优惠。

安全可靠

文本数据存储于腾讯云 [对象存储（Cloud Object Storage, COS）](#) 中，数据跨多架构、多设备冗余存储，提供异地容灾和资源隔离，保证业界领先的数据持久性。

持久可用

高性能的缓存层，搭配高性能的处理集群，快速返回文本审核结果。针对大文件处理场景，增加了异步处理队列逻辑，提高审核处理的成功率。

应用场景

最近更新时间：2021-11-26 17:09:54

电商平台

电商类平台中存在大量的用户评论，用户的评论内容多种多样，很容易存在谩骂、广告等令人不适内容。文本审核可以覆盖各种文本内容类型，检测用户发出的内容，拦截评论中的违规内容，降低人工审核成本。

社交平台

文本审核可以广泛应用于 BBS、博客，以及有用户 UGC 内容的各类网站，包括发帖、回帖及站内信等场景，检测文本中的不适宜内容。通过控制台配置实现增量内容自动触发，服务毫秒级响应，有效保障用户的浏览体验。

视频平台

视频直播中，存在大量用户弹幕、用户评论等场景。传统审核依靠人工审核标注去除违规内容，审核周期较长。文本审核可以精准检测令人反感、不安全或不适宜的内容，快速定位，自动禁封，有效提高审核效率，保障平台安全。