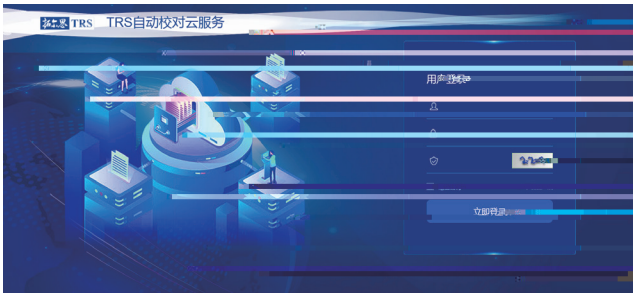
面向文字校审中常见的五大类错误，提供智能化、自动化的文字校审服务，包括：



字词类错误校对

识别由各种编码输入法、语音识别、机器翻译等导致的字词类错误，包括字音相似、字型相似、人名错字、人名多字、称谓错误、地名错误、机构错误等。

- 采用知识库匹配和深度学习相结合的技术，囊括单词和文本语义，文本纠错效果更好。
- 纠错词库、知识库的种类齐全，提供5大类、20多种纠错功能。
- 提供网页端与Microsoft Word插件客户端，并提供Restful API接口规范可与业务系统快速集成。
- 对纠错候选集进行了语言模型的权重计算，使得纠错误报率降低。
- 单机支持100并发，每秒钟30k以上的吞吐能力。
- 针对行业语料，可以用已有词库、知识库自动构建训练集，训练行业纠错模型。
- 支持大部分的国产芯片和操作系统，包括x86平台的海光、arm的鲲鹏和飞腾、龙芯等国产化硬件环境，以及深度、中标麒麟等国产化操作系统。

语法类错误校对

识别多种句法类错误，包括：句式杂糅错误和结构助词错误等。

常识类错误校对

识别多种语义知识错误，包括：地名错误、地域归属错误、国家首都搭配错误、洲国归属错误、组织成员国错误、单位指标错误、缩略词错误以及特定事件错误等。

政治敏感类错误校对

识别多种政治类常识错误，包括：职务称谓错误、人物职务搭配错误、宗教、民族、政治敏感词错误等。

格式类错误校对

识别多种文档格式错误，包括：标点符号错误、日期格式错误、百分比格式错误等。

公文校对

在公文稿件写作的各个环节中，文稿校对是非常重要的一环。利用自动校对技术，校审文稿内容是否合乎党和国家的方针、政策、法律、法规，避免在政治、法律、思想、道德等方面，可能对社会产生的影响和后果。拓尔思自动校对云服务利用AI自动校对技术，从内容准确性、表述规范性、敏感信息检测等多个方面出发，辅助校审人员对公文稿件进行自动校审工作。

新闻媒体内容采编

加强媒体内容审核检测，不仅有利于提升新闻舆论工作的引导力和公信力，更有利于吸引更多的用户和创作者关注平台，保证平台良好的发展势头。然而人工审核无法应对铺天盖地的网络信息，且无法保证在信息时效性内完成相关内容的审核。拓尔思自动校对云服务通过与采编平台的集成，可实现内容发布前的内容校对；通过与内容巡检平台的集成，可实现对网站、新媒体等内容发布后的检测，做到及时发现及时整改。

OCR识别与语音识别后校审

在OCR识别场景中，原始图片信息由于清晰度不够或污损情况将导致识别出现偏差，造成文字错误的情况；同样在语音识别中，由于方言或语速因素也会造成识别不准确现象。借助拓尔思自动校对云服务可以快速定位疑似问题，解决大批量电子文件的自动快速校审问题。