

R

ICS 35.240.20

XX/T

中国人工智能产业发展联盟

## OCR 服务智能化分级技术要求和评估方法

The technical requirements and evaluation method for optical character recognition service intelligent classification

XXXX-XX-XX发布

XXXX-XX-XX实施

中国人工智能产业发展联盟 发布



# 目 次

前言 .....	II
1 范围 .....	3
2 规范性引用文件 .....	3
3 术语、定义和缩略语 .....	3
3.1 术语和定义 .....	3
3.2 缩略语 .....	4
4 OCR 服务评估准则概述 .....	4
5 企业和服务基本信息验证 .....	5
5.1 企业基本信息验证 .....	5
5.2 OCR 服务基本信息验证 .....	5
6 OCR 服务技术要求和测试方法 .....	5
6.1 技术指标要求 .....	5
6.2 测试方法 .....	7

## 前 言

本标准按照GB/T 1.1-2009给出的规则起草。

本标准由中国人工智能产业发展联盟提出并归口。

本标准起草单位：中国信息通信研究院、南京新一代人工智能研究院、华为技术有限公司…。

本标准主要起草人：石霖、梅舒欢、符海芳…。

## 1 范围

本评估规范规定了OCR服务在功能、性能、安全等方面的技术要求以及评估方法。本评估规范适用于各类提供OCR服务的平台和产品。

## 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。

## 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

**光学字符识别** optical character recognition

将图片或扫描件中的文字识别成可编辑的文本。

#### 3.1.2

**OCR 服务** OCR service

OCR服务以公有云、私有云、边缘服务等为基础，提供OCR算法和模型，组合预处理、检测、识别、后处理等模块，用户可以基于该服务直接将图片、PDF、视频中的文字识别出来。

#### 3.1.3

**最长公共子序列，即 Longest Common Subsequence, LCS**

一个序列S任意删除若干个字符得到新序列T，则T叫做S的子序列；两个序列X和Y的公共子序列中，长度最长的那个，定义为X和Y的最长公共子序列。

$$LCS(X_m, Y_n) = \begin{cases} LCS(X_{m-1}, Y_{n-1}) + x_m & \text{当 } x_m = y_n \\ \max\{LCS(X_{m-1}, Y_n), LCS(X_m, Y_{n-1})\} & \text{当 } x_m \neq y_n \end{cases}$$

#### 3.1.4

**字段召回率**

$$Recall = N_p/N$$

$N_p$ 表示OCR服务预测正确的字段数量， $N$ 表示OCR服务需要提取的字段数量。

### 3.1.5

#### 字符准确率

通过计算预测结果和ground truth的lcs，所有lcs长度相加并乘以2，除以所有预测结果长度和ground truth长度的和。

### 3.1.6

#### 字段准确率

如果预测结果和ground truth的完全一致，即视为一个字段正确，正确字段的个数乘以2，除以所有预测结果和ground truth个数的和。

## 3.2 缩略语

下列缩略语适用于本文件。

OCR 光学字符识别 Optical Character Recognition  
LCS 最长公共子序列 Longest Common Subsequence

## 4 OCR 服务评估准则概述

OCR服务的评估内容主要包括企业和服务基本信息以及OCR服务技术要求。

企业和服务基本信息评估主要采用材料审查的方式，对参评企业和服务的基本信息进行验证。其中企业基本信息包括企业执照、规模、资金等，服务基本信息包括架构名称、版本号、基本描述、上市时间等内容。必选项目，企业必须提交材料进行审查；可选项目，企业可以根据自身情况提交相应材料进行审查。

OCR服务技术要求分为基本和增强两级要求，根据OCR服务的相关协议和条款承诺，对每项功能进行符合性测试及验证，实际能力与承诺指标能力相符。

OCR服务等级评估分为基本和增强两级：

a) 基本级通过准则如下：

- 应通过企业和服务基本信息验证所有必选项；
- 应通过OCR服务技术要求中基础性能指标；
- 应通过OCR服务技术要求中3项上基础类功能指标；
- 应通过OCR服务技术要求中2项以上性能指标。

b) 增强级通过准则如下：

- 应通过企业和服务基本信息验证所有必选项；
- 应通过企业和服务基本信息验证6项以上可选项；
- 应通过OCR服务技术要求中增强性能指标；
- 应通过OCR服务技术要求中6项以上基础类功能指标；
- 应通过OCR服务技术要求中3项以上增强类功能指标；
- 应通过OCR服务技术要求中2项以上性能指标；
- 应通过OCR服务技术要求中增强类安全指标。

注：未明确手写体识别的技术要求项均为印刷体文字识别。

## 5 企业和服务基本信息验证

### 5.1 企业基本信息验证

企业基本信息从以下几个方面进行验证：

- a) 营业执照，公司企业法人营业执照（必选）
  - 参评企业需向测试机构提供公司的合法营业执照副本，并加盖公章。
- b) 规模，公司整体社保缴纳信息证明（必选）
  - 参评企业需向测试机构证明公司人员规模，即提供社保缴纳信息证明副本，并加盖公章。
- c) 资金，验资报告（必选）
  - 参评企业需向测试机构证明资金规模，即提供验资报告副本，并加盖公章。
- d) 已经通过的认证（可选）
  - 参评企业需向测试机构提交已经通过的相关认证。

### 5.2 OCR 服务基本信息验证

OCR服务基本信息从以下11个方面进行验证：

- a) OCR服务名称（必选）
  - 企业需向测试机构提供本次OCR服务的名称。
- b) OCR服务版本号（必选）
  - 企业需向测试机构提供本次评测OCR服务的版本号。
- c) OCR服务基本描述（必选）
  - 企业需向测试机构提供本次评测OCR服务的基本描述。
- d) OCR服务平台上线时间（必选）
  - 企业需向测试机构提供本次评测OCR服务平台上线时间的证明材料。
- e) OCR服务功能（必选）
  - 企业需向测试机构提供OCR服务完整的功能说明文档。
- f) 行业实施案例（可选）
  - 企业需向测试机构提供真实的行业实施案例。
- g) 后期维护及响应方式（必选）
  - 企业需向测试机构提供质保时间、质保方式（如远程、现场、7\*24小时等）信息。
- h) 收费和支付方式（必选）
  - 企业需向测试机构提供收费方式、价格等信息
- i) 预测数据的管理要求（必选）
  - 企业需向用户明确提出上传的预测数据的管理要求，已经采取的保护措施。
- j) 服务端安全要求（可选）
  - 服务端在内存态处理原始信息，处理后即删除，不存储用户原始信息、中间结果、最终结果。

## 6 OCR 服务技术要求和测试方法

### 6.1 技术指标要求

### 6.1.1 基础类功能指标

OCR服务应支持以下功能要求，支持程度与承诺相符，且字符准确率在95%以上，字段准确率在90%以上：

- a) 提供卡证类文字识别服务，如以下类型卡证：身份证、银行卡、营业执照、驾驶证、行驶证、护照、道路运输证等。
- b) 提供票据类文字识别服务，如以下类型票据：增值税发票、出租车发票、火车票、机票行程单、定额发票、车辆通行费发票、机动车销售发票等。
- c) 提供通用表格文字识别服务，支持恢复表格行列信息以及文字内容。
- d) 提供通用图片文字识别服务。
- e) 服务提供语言或平台SDK，如以下类型SDK：如Java、Python、iOS、Android、Node.js、C++、C#。
- f) 票据类服务支持盖章检测。
- g) 支持中英文识别。

### 6.1.2 增强类功能指标

OCR服务宜支持以下增强功能要求，支持程度与承诺相符，且字符准确率在95%以上，字段准确率在90%以上：

- a) 提供卡证类文字识别服务，如：身份证、银行卡、营业执照、驾驶证、行驶证、护照、道路运输证等。至少支持六类卡证文字识别。
- b) 提供票据类文字识别服务，如以下类型票据：增值税发票、出租车发票、火车票、机票行程单、定额发票、车辆通行费发票、机动车销售发票等。至少支持六类票据文字识别。
- c) 提供通用自然场景文字识别服务，如指示牌、菜单、车牌等，至少支持两类如上列举的自然场景文字识别。
- d) 提供非规范票据文字识别服务，如购物小票文字识别服务。
- e) 所有服务支持任意角度进行检测识别，如正面、侧面、弱光等场景。
- f) 支持多语言识别，支持程度与承诺相符，并至少支持十种语言识别（包括中英文），中文包含GB 2312常用汉字6763个，日文包含常用假名和2010年11月日本政府公布的《改订常用汉字表》中的2136个汉字。其他语言为本语言特有字符。
- g) 支持手写体中英文识别。
- h) 支持票据类票据内文字和盖章弧形文字同时识别，可支持盖章真伪检测。
- i) 支持一图多票能力，提供一图多票API。
- j) 支持自定义模板开发。

### 6.1.3 性能指标

性能指标按照证件类和票据类两类进行测试，共14项：

- a) 标准证件类保证字段召回率60%以上的基础上，以lcs计算的字符准确率达到99%，字段准确率达到94%。与承诺相符。证件类数据集包括：
  - 1) 身份证
  - 2) 银行卡
  - 3) 营业执照



- 4) 驾驶证
  - 5) 行驶证
  - 6) 护照
  - 7) 道路运输证
- b) 标准票据类保证字段召回率60%以上的基础上，以lcs计算的字符准确率达到98%，字段准确率达到93%。与承诺相符。票据类数据集包括：
- 1) 增值税发票
  - 2) 出租车发票
  - 3) 火车票
  - 4) 机票行程单
  - 5) 定额发票
  - 6) 车辆通行费发票
  - 7) 机动车销售发票

#### 6.1.4 增强类安全指标

服务宜支持对用户信息的隐私保护，对原始信息进行处理，网络传输和服务端无完整有序信息。

## 6.2 测试方法

### 6.2.1 测试材料和账户要求

评估对象：OCR服务

评估方法：主要基于材料审查和技术测试的方式，包括：

- 材料审查，参评企业应提供与承诺相符的功能性相关材料，包括：参评企业应能提供用户有关使用其服务的操作指导/说明资料，以及不同类型用户的功能差异化说明、第三方接入 API、SDK 和文档等。
- 技术测试，由参评企业提供测试账号或测试环境进行测试。

### 6.2.2 基础类功能指标评估方法

参评企业提供 OCR 服务相关材料，完成材料审查的同时，基于技术指标要求和材料完成如下指标测试，指标都应满足字符准确率 95%，字段准确率 90%，且能力与证明材料承诺相符：

- a) 提供卡证类文字识别服务  
准备待测图片，按照服务输入参数格式，调用API或SDK进行识别，返回参数正确，且能力与证明材料承诺相符即视为通过。
- b) 提供票据类文字识别服务  
准备图片，按照服务输入参数格式，调用API或SDK进行识别，返回参数正确，且能力与证明材料承诺相符即视为通过。
- c) 提供通用表格文字识别服务  
准备图片，按照服务输入参数格式，调用API或SDK进行识别，返回参数正确，且能力与证明材料承诺相符即视为通过。
- d) 提供通用图片文字识别服务

准备图片，按照服务输入参数格式，调用API或SDK进行识别，返回参数正确，且能力与证明材料承诺相符即视为通过。

e) 服务提供语言或平台SDK

准备图片，按照服务输入参数格式，调用所承诺的API或SDK分别进行识别，返回参数正确，且能力与证明材料承诺相符即视为通过。

f) 票据类服务支持盖章检测

准备具备盖章的票据类图片，按照服务输入参数格式，调用API或SDK进行识别，返回参数正确，且能力与证明材料承诺相符即视为通过。

g) 支持中英文识别

准备中文、英文的卡证和票据类图片，按照服务输入参数格式，调用API或SDK分别进行识别，返回参数正确，且能力与证明材料承诺相符即视为通过。

### 6.2.3 增强类功能指标评估方法

参评企业提供OCR服务相关材料，完成材料审查的同时，基于技术指标要求和材料完成如下指标测试：

a) 提供卡证类文字识别服务，

准备待测图片，如以下类型卡证：身份证、银行卡、营业执照、驾驶证、行驶证、护照、道路运输证等，按照服务输入参数格式，调用API或SDK进行识别，返回参数正确，能力与证明材料承诺相符，能够识别六类以上则视为通过。

b) 提供票据类文字识别服务支持程度与承诺相符，并至少支持六类票据文字识别。

准备待测图片，如以下类型票据：增值税发票、出租车发票、火车票、机票行程单、定额发票、车辆通行费发票、机动车销售发票等。按照服务输入参数格式，调用API或SDK进行识别，返回参数正确，能力与证明材料承诺相符，能够识别六类以上则视为通过。

c) 提供通用自然场景（支持指示牌、菜单、车牌文字）文字识别服务

准备指示牌图片，按照服务输入参数格式，调用API或SDK进行识别，返回参数正确，且能力与证明材料承诺相符，能够识别两类以上则视为通过。

d) 提供非规范票据文字识别服务

准备购物小票图片，按照服务输入参数格式，调用API或SDK进行识别，返回参数正确，且能力与证明材料承诺相符即视为通过。

e) 所有服务支持任意角度进行检测识别。支持程度与承诺相符。

准备卡证、票据、表格、菜单任意角度图片，按照服务输入参数格式，调用API或SDK分别进行识别，返回参数正确，且能力与证明材料承诺相符即视为通过。

f) 支持多语言识别

准备多种语言票据、卡证图片，调用API或SDK分别进行识别，返回参数正确，且能力与证明材料承诺相符，并至少支持十种语言识别（包括中英文）即视为通过。

g) 支持手写体中英文识别

准备手写体中英文图片，调用API或SDK分别进行识别，返回参数正确，且能力与证明材料承诺相符即视为通过。

h) 支持票据类盖章弧形文字识别

准备票据类盖章图片，调用API或SDK分别进行识别，返回票据上的文字和盖章内弧形文字参数正确，且能力与证明材料承诺相符即视为通过。

j) 支持一图多票能力

准备待测试图片，一张图中包含多个票据信息，调用 API 或 SDK 进行识别，返回参数正确，且能力与证明材料承诺相符即视为通过。

k) 支持自定义模板

准备待测试的图片，自定义上传识别的模板，再调用 API 或 SDK 分别进行识别，返回参数正确，且能力与证明材料承诺相符即视为通过。

#### 6.2.4 性能指标评估方法

参评企业提供OCR服务相关材料，完成材料审查的同时，基于技术指标要求和材料完成如下指标测试：

a) 标准证件类保证字段召回率 60% 以上的基础上的 lcs 字符准确率达到 99%，字段准确率达到 94%。

准备证件类图片 300 张及以上，调用 API 或 SDK 进行识别，再基于返回值与实际值进行比较，计算出的准确率数据与证明材料承诺相符，且字符准确率达到 99%，字段准确率达到 94%，即视为通过。种类包括如下几种，分别进行测试，单独记录通过情况：

- 1) 身份证
- 2) 银行卡
- 3) 营业执照
- 4) 驾驶证
- 5) 行驶证
- 6) 护照
- 7) 道路运输证

b) 标准票据类数据集保证字段召回率 60% 以上的基础上的 lcs 字符准确率达到 98%，字段准确率达到 93%。与承诺相符。

准备票据类图片 300 张及以上，调用 API 或 SDK 进行识别，再基于返回值与实际值进行比较，计算出的准确率数据与证明材料承诺相符且字符准确率达到 98%，字段准确率达到 93% 即视为通过。种类包括如下几种，分别进行测试，单独记录通过情况：

- 1) 增值税发票
- 2) 出租车发票
- 3) 火车票
- 4) 机票行程单
- 5) 定额发票
- 6) 车辆通行费发票
- 7) 机动车销售发票

#### 6.2.5 增强类安全指标评估方法

参评企业提供OCR服务相关材料，完成材料审查的同时，基于技术指标要求和材料完成如下指标测试：

服务宜支持对用户信息的隐私保护

准备待测试的图片，先调用抠图的接口，将图片中待识别区域裁剪出来，本地通过增加干扰的方法，生成多于待识别区域个数的裁剪图片，然后调用API或SDK进行识别，在本地将识别

的结果进行组合，生成最后的结果数据正确，本地增加的裁剪图片个数多于待识别区域个数，从生成的裁剪图片，有明显的加密干扰信息，且测试准确率与证明材料承诺相符即视为通过。