

# 数据标注

## 职业技能等级标准

标准代码：510049  
(2021年2.0版)

哈工大大数据集团有限公司 制定  
2021年12月 发布

# 目 次

前言.....	1
1 范围.....	2
2 规范性引用文件.....	2
3 术语和定义.....	2
4 适用院校专业.....	2
5 面向职业岗位（群） .....	3
6 职业技能要求.....	3
参考文献.....	14

# 前 言

本标准按照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。

本标准起草单位：哈工大大数据集团有限公司、语仓科技（北京）有限公司、北京百智天拓教育有限公司、小牛思拓（北京）科技有限公司、辽宁盘石数据科技有限公司、哈尔滨工业大学、北京语言大学、苏州大学、郑州电力高等专科学校。

本标准主要起草人：张熠天、王会珍、于东、饶高琦、李正华、郑爽、冯振杰、喻宙。

声明：本标准的知识产权归属于哈工大大数据集团有限公司与语仓科技（北京）有限公司共同所有，未经哈工大大数据集团有限公司与语仓科技（北京）有限公司双方同意，不得印刷、销售。

## 1 范围

本标准规定了数据标注职业技能等级对应的工作领域、工作任务及职业技能要求。

本标准适用于数据标注职业技能培训、考核与评价，相关用人单位的人员聘用、培训与考核可参照使用。

## 2 规范性引用文件

下列文件对于本标准的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本标准。凡是不注日期的引用文件，其最新版本适用于本标准。

- GB/T 5271.31-2006 信息技术 词汇 第31部分:人工智能 机器学习
- GB/T 5271.34-2006 信息技术 词汇 第34部分:人工智能 神经网络
- GB/T 38548.4-2020 内容资源数字化加工 第4部分:元数据
- GB/T 38548.3-2020 内容资源数字化加工 第3部分:加工规格
- GB/T 38548.2-2020 内容资源数字化加工 第2部分:采集方法
- GB/T 38548.1-2020 内容资源数字化加工 第1部分:术语
- GB/T 38548.6-2020 内容资源数字化加工 第6部分:应用模式

## 3 术语和定义

### 3.1

**数据 data**

对事实、概念和指令的一种形式化表示，适于人工或自动方式进行通信、解释或处理。

[GB/T 18391.4-2001, 定义 3.3 ]

### 3.2

**数据标注 data annotation**

为机器学习建立模型，提供人工数据的过程，是数据到知识转化过程。语言学 linguistics

语言学是研究人类语言的科学。语言学的主要任务是研究语言的结构和功能，追溯语言发展演变的历史，研究语言与其他相关现象的关系。

[语言学概论，华中师范大学出版社，2002]

### 3.3

**人工智能 artificial intelligence**

表现出与人类智能（如推理和学习）相关的各种功能的功能单元的能力。

[GB/T 5271.28-2001, 定义 28.01.02]

### 3.4

**知识表示 knowledge representation**

对知识的一种描述或一组约定，一种计算机可以接受的用于描述知识的数据结构。

[GB/T/ 39475-2020, 定义 3.1]

## 4 适用院校专业

### 4.1 参照原版专业目录

中等职业学校：计算机应用、电子与信息技术、电子技术应用、计算机网络技术、数字媒体技术应用、网络营销、旅游服务与管理、计算机平面设计、软件与信息服务、客户信息服务、计算机速录、文秘、商务助理、商务英语、商务韩语等相关专业

高等职业学校：计算机应用技术、人工智能技术应用、数字图文信息技术、软件与信息服务、大数据技术与应用、信息统计与分析、小学教育、语文教育、英语教育、文秘、建筑工程技术、建筑智能化工程技术、会计信息管理、会计、云计算技术与应用、信息安全与管理、电子商务等相关专业

高等职业教育本科学校：计算机应用工程、网络工程、软件工程、数字媒体技术、大数据技术与应用、区块链技术与应用专业

应用型本科学校：计算机应用工程、智能制造工程、网络工程、统计学、应用统计学、

数据科学与大数据技术、大数据技术与应用、大数据管理与应用、信息资源管理、商务英语、新闻学、教育学、工商管理、电子商务、软件工程等相关专业

#### 4.2 参照新版职业教育专业目录

中等职业学校：计算机应用、电子信息技术、电子技术应用、计算机网络技术、数字媒体技术、网络营销、旅游服务与管理、大数据技术应用、计算机平面设计、软件与信息服务、客户信息服务、文秘、商务助理、商务英语、商务韩语等相关专业

高等职业学校：计算机应用技术、人工智能技术服务、印刷数字图文技术、软件技术、大数据技术、统计与大数据分析、小学教育、小学语文教育、小学英语教育、现代文秘、建筑工程技术、建筑智能化工程技术、会计信息管理、大数据与会计、云计算技术应用、信息安全技术应用、电子商务等相关专业

高等职业教育本科学校：计算机应用工程、网络工程技术、软件工程技术、数字媒体技术、大数据工程技术、区块链技术、人工智能工程技术

应用型本科学校：计算机科学与技术、智能科学与技术、人工智能、机器人工程、网络工程、统计学、应用统计学、数据科学与大数据技术、电子信息工程、电子与计算机工程、大数据管理与应用、智能制造工程、信息资源管理、英语、新闻学、教育学、工商管理、电子商务、软件工程等相关专业

### 5 面向职业岗位（群）

数据标注职业技能等级证书主要适用于已在人工智能、NLP、模型算法、大数据、舆情分析等行业从业或有志于在这些行业从业的标注人员、数据处理人员或人工智能训练师。通过数据标注职业技能等级认证后，可以从事如下工作：

【数据标注职业技能等级证书】（初级）：主要面向数据标注、数据处理、工业互联网、自动驾驶等行业，从事医疗影像标注、语音转写、交通场景标注、人脸标注、数据采集加工、数据审核、数据分类、格式转换、排版、数据可视化等工作。

【数据标注职业技能等级证书】（中级）：主要面向人工智能行业，NLP和算法领域，除可从事医疗影像标注、语音转写、交通场景标注、人脸标注、数据采集加工、数据审核、数据分类等工作外，还可从事语义分割、关键点标注、音频分析、文本分类、实体关系标注、信息抽取、关键词标注等自然语言处理类工作，或按照需求完成标注指南制定、标注项目组织与实施等部分管理工作。

【数据标注职业技能等级证书】（高级）：主要面向人工智能、大数据行业，担任人工智能训练师、数据分析、舆情分析等岗位，对标人工智能训练师能力，从事数据集产品设计、标注知识体系设计、标注平台设计、标注项目管理、技术对接、模型优化方案设计等工作，对标大数据工程师，从事数据预处理、数据采集、舆情分析、行业报告撰写、多语言信息处理、行业调研、信息检索、人工翻译等工作。

### 6 职业技能要求

#### 6.1 职业技能等级划分

数据标注职业技能等级分为三个等级：初级、中级、高级，三个级别依次递进，高级别涵盖低级别职业技能要求。

【数据标注】（初级）：对数据标注有基本的认识和理解，掌握简单任务的目标及基本标注原则和方法，能够通过学习给定标注指南，完成特定的数据标注任务。

【数据标注】（中级）：对数据标注具有较深入的理解，掌握常见标注任务的目标及基本标注方法，不仅可以快速学习给定标注指南，高质量地完成数据标注任务，还要能够承担一定的项目管理职责，辅助管理者或培训师实施培训或答疑解惑并能够在这一过程中，对标注指南提出改进意见，和项目管理者进行有效沟通，提升整个项目的质量和效率。行业内大部分从业人员都将处于这一级别。

【数据标注】（高级）：深入了解标注行业和需求，掌握一定的人工智能理论和实践经验，具备一定的算法意识，能够在标注平台设计中给出优化建议，能够按照规范实施标注任务并达到免检水平，可以针对标注任务进行有效的需求分析，把握标注原则并针

对特殊情况作出恰当处理，同时能够主持或参与标注指南和标注体系的制定和完善。高级数据标注人员主要从事标注全流程的把控以及标注项目的管理工作。

## 6.2 职业技能等级要求描述

表 1 数据标注职业技能等级要求（初级）

工作领域	工作任务	职业技能（能力）要求
1. 文本标注	1.1原始数据处理	<p>1.1.1 能在项目经理的指导下，利用普通办公软件（Word、Excel 等）完成明显格式错误、错别字等的修改和调整，准确率达到90%以上；</p> <p>1.1.2 能够用办公软件或 notepad++等实现共性数据的查找和替换，达到 90%正确率；</p> <p>1.1.3 能够按照使用说明将给定的已处理完成的原始数据导入到标注实训平台，导入成功率100%；</p> <p>1.1.4 能够按照要求利用采集工具或检索平台等完成简单的原始数据采集和整理工作，不合格率不超过20%。</p>
	1.2标注任务实施	<p>1.2.1 能按照指南要求和指导，利用实训平台完成简单的单级分类和意图标注，准确率在80%以上；</p> <p>1.2.2 能够按照指南要求和指导，利用实训平台完成 8 类通用类别的实体标注任务，准确率在80%以上；</p> <p>1.2.3 能够按照质检意见对实训平台中已实施的标注任务结果做全面准确的修改，修改后的正确率不低于90%；</p> <p>1.2.4 能够按照指导和要求，在实训平台内创建全流程的分类和实体标注项目，创建成功率达到100%。</p>
	1.3标注指南理解与制定	<p>1.3.1 能够在项目经理或培训视频的辅助下读懂并理解标注指南中的任务目标，单纯按照指南的试标准确率达70%以上；</p> <p>1.3.2 能按照实际标注情况与标注指南作对比，发现并反馈标注指南与实际标注违背之处，反馈被采纳率不低于60%；</p> <p>1.3.3 能够通过仔细阅读发现标注指南中明显的矛盾点或逻辑错误并及时反馈，被采纳率不低于 80%；</p> <p>1.3.4 能够按照标注指南中的描述总结标注处理原则，从而在必要时按照大原则给出标注结果，误差率不超过30%。</p>

2. 图像标注	2.1原始数据处理	<p>2.1.1 能在项目经理的指导下，利用常用图像查看器对原始图像的本身特征等进行判别和筛查，准确率达到90%以上；</p> <p>2.1.2 能够利用常用软件等对图像数据进行打包或压缩等操作，成功率100%；</p> <p>2.1.3 能够按照使用说明将给定的已处理完成的原始图像数据导入到标注实训平台，导入成功率100%；</p> <p>2.1.4能够按照要求利用采集工具或到给定来源中完成原始图像数据的下载和采集工作，不合格率不超过20%。</p>
	2.2标注任务实施	<p>2.2.1 根据指南要求和指导，利用实训平台从事图像的2D拉框标注，合格率不低于90%；</p> <p>2.2.2 根据指南要求和指导，利用实训平台从事图像分类标注，合格率不低于90%；</p> <p>2.2.3 能够按照视频指导或安装说明，从给定来源下下载并安装标注工具和软件，成功率达100%；</p> <p>2.2.4 能够按照指导和要求，在实训平台内创建全流程的2D拉框和图像分类标注项目，创建成功率达到100%。</p>
	2.3标注指南理解与制定	<p>2.3.1 能够在培训视频的辅助下读懂并理解标注指南中的任务目标，按照指南的试标准准确率90%以上；</p> <p>2.3.2 能按照实际标注情况与标注指南作对比，发现并反馈标注指南与实际标注违背之处，反馈被采纳率不低于60%；</p> <p>2.3.3 能够通过仔细阅读发现标注指南中明显的矛盾点或逻辑错误并及时反馈，被采纳率不低于80%；</p> <p>2.3.4 能够按照标注指南中的描述总结标注处理原则，从而在必要时按照大原则给出标注结果，误差率不超过30%。</p>
3. 语音标注	3.1原始数据处理	<p>3.1.1 能在项目经理的指导下，利用常用音频软件对原始音频的本身特征等进行判别和筛查，准确率达到90%以上；</p> <p>3.1.2 能够利用常用软件等对音频数据进行打包或压缩等操作，成功率100%；</p> <p>3.1.3 能够按照使用说明将给定的已处理完成的原始语音数据导入到标注实训平台，导入成功率100%；</p> <p>3.1.4 能够按照要求利用采集工具或到特定来源中完成语音数据的下载和采集工作，不合格率不超过20%；</p> <p>3.1.5 能够利用给定的音频编辑软件对原始语音数据做基本分割和时长统计等工作，误</p>

		差不超过10%。
	3.2标注任务实施	3.2.1 根据指南要求和指导，利用标注实训平台完成语音切割和转写任务，准确率在85%以上； 3.2.2 根据指南要求和指导，利用标注实训平台完成语音校对任务，准确率在90%以上； 3.2.3 按照说明或指导从指定来源中下载安装辅助标注工具，成功率达100%； 3.2.4 能够按照指导和要求，在实训平台内创建全流程的语音转写和校对标注项目，创建成功率达到100%。
	3.3标注指南理解与制定	3.3.1 能够在培训视频的辅助下读懂并理解标注指南中的任务目标，按照指南的试标准准确率达90%以上； 3.3.2 能按照实际标注情况与标注指南作对比，发现并反馈标注指南与实际标注违背之处，反馈被采纳率不低于60%； 3.3.3 能够通过仔细阅读发现标注指南中明显的矛盾点或逻辑错误并及时反馈，被采纳率不低于80%； 3.3.4 能够按照标注指南中的描述总结标注处理原则，从而在必要时按照大原则给出标注结果，误差率不超过30%。

表2 数据标注职业技能等级要求（中级）

工作领域	工作任务	职业技能（能力）要求
1. 文本标注	1.1标注数据处理	1.1.1 能够利用给定的可执行脚本工具将文档处理成可导入标注系统的文件格式，准确率达100%； 1.1.2 能够通过文本编辑工具等对标注结果做一致性筛查和把握，准确率在90%以上； 1.1.3 能够运用现成的执行工具对原始数据做筛查和清洗，误差率在10%以内； 1.1.4 能够通过工具或办公软件等，对原始数据进行分类，误差率在10%以内。
	1.2 标注任务实施	1.2.1 根据指南要求和指导，熟练地利用标注工具或平台从事实体和分类标注任务，准确率达到90%以上； 1.2.2 根据指南要求和指导，利用标注工具或平台从事中文关系标注任务，准确率达到85%； 1.2.3 根据指南要求和指导，利用标注工具



		<p>或平台从事事件标注任务，准确率达到85%；</p> <p>1.2.4 根据指南要求和指导，利用标注工具或平台从事关键词标注任务，准确率达到90%；</p> <p>1.2.5 能够按照指导和要求，在实训平台内创建全流程的关系、事件、关键词标注项目，创建成功率达到100%。</p>
	1.3 标注项目管理	<p>1.3.1 能够承担实体、分类、关系、事件、关键词标注任务的质检或审核工作，质检准确率不低于90%；</p> <p>1.3.2 熟悉项目管理配置、项目实施、人员管理配置以及数据管理配置方法，能够按照既定方案在标注系统中建立、配置标注项目，参数设置准确率为100%；</p> <p>1.3.3 能够独立地按照培训视频和说明利用标注工具和平台操作任务并顺利通过测试，初次测试准确率达到80%以上；</p> <p>1.3.4 能够按照项目经理的要求和指导，辅助实施指导和测试，通过有效率达80%以上；</p> <p>1.3.5 能够按照项目经理指导通过办公软件和系统完成测试题筛选和评阅工作，评阅准确率在80%以上。</p>
	1.4 标注指南理解与制定	<p>1.4.1 能够快速理解标注指南，按照指南的要求准确地理解标注目标，对目标复述的准确率不低于85%；</p> <p>1.4.2 能够在理解指南的基础上发现指南的缺陷并提出修改和补充意见，意见被采纳率不低于70%；</p> <p>1.4.3 能够按照项目经理的指导对指南做简单的修正，准确率在85%以上；</p> <p>1.4.4 能够按照项目经理的指导对指南做改写，改写后的指南应至少有助于80%标注人员理解。</p>
2. 图像标注	2.1 标注数据处理	<p>2.1.1 能够利用给定的可执行程序将原始图片处理成可导入标注系统的文件格式，准确率达100%；</p> <p>2.1.2 能够利用看图工具或脚本等对标注结果做一致性筛查和把握，准确率在90%以上；</p> <p>2.1.3 能够运用现成的执行工具对原始数据做筛查和清洗，误差率在10%以内；</p> <p>2.1.4 能够通过工具或办公软件等，对原始数据进行分类，误差率在10%以内。</p>

	2.2标注任务实施	<p>2.2.1 能够根据指南要求和指导，熟练地利用标注工具或平台完成2D拉框和图片分类任务，准确率在95%以上；</p> <p>2.2.2 能够根据指南要求和指导，利用标注工具或平台从事2D图片的多边形标注任务，准确率在95%以上；</p> <p>2.2.3 能够根据指南要求和指导，利用标注工具或平台完成关键点标注任务，准确率达到95%以上；</p> <p>2.2.4 能够根据指南要求和指导，利用标注工具或平台完成语义分割标注任务，准确率达到90%以上；</p> <p>2.2.5 能够按照系统说明书在实训平台创建全流程的多边形、关键点、语义分割标注项目，创建成功率达100%。</p>
	2.3标注项目管理	<p>2.3.1 能够利用系统或平台完成2D拉框、多边形标注、关键点标注以及语义分割标注任务的质检或审核工作，质检准确率不低于95%；</p> <p>2.3.2 熟悉项目管理配置、项目实施、人员管理配置以及数据管理配置方法，能够按照既定方案在标注系统中建立、配置标注项目，参数设置准确率为100%；</p> <p>2.3.3 能够独立地按照培训视频和说明利用标注工具和平台操作任务并顺利通过测试，初次测试准确率达到95%以上；</p> <p>2.3.4 能够按照项目经理的要求和指导，辅助实施指导和测试，通过有效率达90%以上；</p> <p>2.3.5 能够按照项目经理指导通过办公软件和系统完成测试题筛选和评阅工作，评阅准确率在90%以上。</p>
	2.4 标注指南理解与制定	<p>2.4.1 能够快速理解标注指南，准确地理解标注目标，对目标复述的准确率不低于85%；</p> <p>2.4.2 能够在理解指南的基础上发现指南的缺陷并提出修改和补充意见，意见被采纳率不低于70%；</p> <p>2.4.3 能够按照项目经理的指导对指南做简单的修正，准确率在85%以上；</p> <p>2.4.4 能够按照项目经理的指导对指南做改写，改写后的指南应至少有助于80%标注人员理解。</p>

3. 语音标注	3.1 标注数据处理	<p>3.1.1 能够利用给定的可执行程序将原始语音处理成可导入标注系统的文件格式，准确率达100%；</p> <p>3.1.2 能够利用辅助工具或脚本等对标注结果做一致性筛查和把握，准确率在90%以上；</p> <p>3.1.3 能够运用现成的执行工具筛选有效的原始数据，误差率在10%以内；</p> <p>3.1.4 能够通过工具或办公软件等，对原始数据进行分类，误差率在10%以内。</p>
	3.2 标注任务实施	<p>3.2.1 能够按照标注指南的要求，熟练地利用标注平台或工具实施语音转写和校对任务，准确率不低于95%；</p> <p>3.2.2 能够按照标注指南的要求，利用标注平台或工具实施语音分类标注，准确率不低于95%；</p> <p>3.2.3 能够按照标注指南的要求，利用普通录音软件实施普通话语音采集任务，阅读准确率不低于90%；</p> <p>3.2.4 能够按照系统说明在标注平台内创建全流程的语音分类、采集项目，成功率达100%。</p>
	3.3 标注项目管理	<p>3.3.1 能够利用系统或平台完成语音转写、校对、分类及采集任务的质检或审核工作，质检准确率不低于95%；</p> <p>3.3.2 熟悉项目管理配置、项目实施、人员管理配置以及数据管理配置方法，能够按照既定方案在标注系统中建立、配置标注项目，参数设置准确率为100%；</p> <p>3.3.3 能够独立地按照培训视频和说明利用标注工具和平台操作任务并顺利通过测试，初次测试准确率达到95%以上；</p> <p>3.3.4 能够按照项目经理的要求和指导，辅助实施指导和测试，通过有效率达95%以上；</p> <p>3.3.5 能够按照项目经理指导通过办公软件和系统完成测试题筛选和评阅工作，评阅准确率在95%以上。</p>
	3.4 标注指南理解与制定	<p>3.4.1 能够快速地理解标注指南，准确地理解标注目标，对目标复述的准确率不低于85%；</p> <p>3.4.2 能够在理解指南的基础上发现指南的缺陷并提出修改和补充意见，意见被采纳率不低于70%；</p> <p>3.4.3 能够按照项目经理的指导对指南做简单的修正，准确率在85%以上；</p> <p>3.4.4 能够按照项目经理的指导对指南做改</p>

		写，改写后的指南应至少有助于80%标注人员理解。
--	--	--------------------------

表3数据标注职业技能等级要求（高级）

工作领域	工作任务	职业技能（能力）要求
1. 文本标注	1.1标注数据处理	1.1.1 能够利用简单的规则、公式等对数据做简单筛选和清洗，准确率不低于95%； 1.1.2 能够利用简单的规则对所有标注结果做一致性筛查和把握，误差率不超过5%； 1.1.3 具备数据安全和敏感意识，能够利用普通办公软件对标注结果做简单的脱敏或加密处理，脱敏误差率不超过5%； 1.1.4 能够通过既有工具或脚本进行数据清洗，准确率不低于95%。
	1.2标注任务实施	1.2.1 能够按照标注指南和要求熟练地运用标注系统完成初级和中级的文本标注任务，达到免检水平；或者 1.2.2 能够对某一门外语或某一领域较为精通，完成多语言或专业领域的上述标注任务，准确率不低于90%； 1.2.3 能够按照标注指南和要求，利用标注系统完成问句复述任务，达到免检水平； 1.2.4 能够按照标注指南和要求，利用标注系统完成问答对标注任务，达到免检水平； 1.2.5 能够按照系统说明在标注平台内创建全流程问句复述或问答对标注项目，成功率达100%。
	1.3标注项目管理	1.3.1 能够利用办公软件独立制定项目实施和管理方案，方案可行性达100%； 1.3.2 能够独立带队或指导他人在系统内组织实施标注项目，项目按期完成率达100%； 1.3.3 能够利用录屏、录像等软件独立录制培训讲解视频并在线上线下组织培训活动，培训视频在项目内的普及率达到100%； 1.3.4 能够独立领导、安排或实施测试题筛选、评阅等工作，准确率达100%。
	1.4标注指南理解与制定	1.4.1 能够独立制定并改写标注指南，指南经过简单修改后能够100%可用； 1.4.2 能够按照需求方的需求，独立制定标注知识体系，所给知识点定义的准确率达到100%； 1.4.3 能够按照实际需求通过标注工具和过程对制定完成的知识体系及指南进行验证和修改，保证按照该知识体系和指南标注的结果能够100%提升训练效果； 1.4.4 能够针对需求方的标注指南提出有效的修改建议，被采纳率达到80%以上。

2. 图像标注	2.1 标注数据处理	<p>2.1.1 能够利用简单的脚本等对图像数据做特征分析和分类，准确率不低于95%；</p> <p>2.1.2 能够利用简单的规则或脚本对所有标注结果做一致性筛查和把握，误差率不超过5%；</p> <p>2.1.3 具备数据安全意识，能够利用普通办公软件对标注结果做加密处理，加密后的文件100%不可见；</p> <p>2.1.4 能够通过既有工具或脚本进行数据清洗，准确率不低于95%。</p>
	2.2 标注任务实施	<p>2.2.1 能够根据指南要求和指导，利用标注系统熟练地从事初级、中级的图像标注任务，达到免检水平；或者</p> <p>2.2.2 能够掌握一门外语，利用标注系统完成多语言图像标注任务，达到90%以上准确率；</p> <p>2.2.3 能够根据指南要求和指导，利用标注系统完成3D点云标注任务，达到90%以上准确率；</p> <p>2.2.4 能够按照系统说明在平台内创建全流程的3D点云标注项目，成功率100%。</p>
	2.3 标注项目管理	<p>2.3.1 能够利用办公软件独立制定项目实施和管理方案，方案可行性达100%；</p> <p>2.3.2 能够独立带队或指导他人在系统内组织实施标注项目，项目按期完成率达100%；</p> <p>2.3.3 能够利用录屏、录像等软件独立录制培训讲解视频并在线上线下组织培训活动，培训视频在项目内的普及率达到100%；</p> <p>2.3.4 能够独立领导、安排或实施测试题筛选、评阅等工作，准确率达100%。</p>
	2.4 标注指南理解与制定	<p>2.4.1 能够独立制定并改写标注指南，指南经过简单修改后能够100%可用；</p> <p>2.4.2 能够按照需求方的需求，独立制定标注知识体系，所给知识点定义的准确率达到100%；</p> <p>2.4.3 能够按照实际需求通过标注工具和过程对制定完成的知识体系及指南进行验证和修改，保证按照该知识体系和指南标注的结果能够100%提升训练效果；</p> <p>2.4.4 能够针对需求方的标注指南提出有效的修改建议，被采纳率达到80%以上。</p>
3. 语音标注	3.1 标注数据处理	<p>3.1.1 能够利用简单的脚本等对语音数据做特征分析和统计，准确率不低于95%；</p> <p>3.1.2 能够利用简单工具或脚本对所有标注结果做一致性筛查和把握，误差率不超过5%；</p> <p>3.1.3 具备数据安全意识，能够利用普通办</p>

		<p>公软件对标注结果做加密处理，加密后的文件100%不可见；</p> <p>3.1.4 能够通过既有工具或脚本对原始语音数据进行自动切割，成功率达到100%；</p> <p>3.1.5 能够通过既有工具或脚本进行数据清洗，准确率不低于95%。</p>
	3.2 标注任务实施	<p>3.2.1 能够按照指南要求，熟练地利用标注工具或平台从事初级和中级的语音标注任务，并达到免检水平；或者</p> <p>3.2.2 能够掌握某门外语的听写能力，利用标注平台完成多语言的语音转写或校对等任务，准确率达到95%；</p> <p>3.2.3 能够按照指南要求，利用标注工具或平台完成拼音标注任务，准确率达到95%；</p> <p>3.2.4 能够按照系统说明在标注平台内创建全流程拼音标注项目，成功率达到100%。</p>
	3.3 标注项目管理	<p>3.3.1 能够利用办公软件独立制定项目实施和管理方案，方案可行性达100%；</p> <p>3.3.2 能够独立带队或指导他人在系统内组织实施标注项目，项目按期完成率达100%；</p> <p>3.3.3 能够利用录屏、录像等软件独立录制培训讲解视频并在线上线下组织培训活动，培训视频在项目内的普及率达到100%；</p> <p>3.3.4 能够独立领导、安排或实施测试题筛选、评阅等工作，准确率达100%。</p>
	3.4 标注指南理解与制定	<p>3.4.1 能够独立制定并改写标注指南，指南经过简单修改后能够100%可用；</p> <p>3.4.2 能够按照需求方的需求，独立制定标注知识体系，所给知识点定义的准确率达到100%；</p> <p>3.4.3 能够按照实际需求通过标注工具和过程对制定完成的知识体系及指南进行验证和修改，保证按照该知识体系和指南标注的结果能够100%支持标注过程；</p> <p>3.4.4 能够针对需求方的标注指南提出有效的修改建议，被采纳率达到80%以上。</p>

## 参考文献

- [1] 中等职业学校专业教学标准（2017年）
- [2] 高等职业学校专业教学标准（2019年）
- [3] 普通高等学校本科专业类教学质量国家标准（2018年）
- [4] 中华人民共和国职业分类大典(2015年)
- [5] GB/T 5271-2006 信息技术
- [6] GB/T 38548-2020 内容资源数字化加工
- [7] GB/T 18391.4-2001 信息技术 数据元的规范与标准化
- [8] GB/T39475-2020 通用制造工艺知识表示
- [9] 邢福义、吴振国主编. 语言学概论[M]. 武汉: 华中师范大学出版社. 2002.
- [10] 教育部关于印发《职业教育专业目录（2021年）》的通知（教职成〔2021〕2号）
- [11] 《教育部关于公布2019年度普通高等学校本科专业备案和审批结果的通知》（教高函〔2020〕2号）
- [12] 《教育部关于公布2020年度普通高等学校本科专业备案和审批结果的通知》（教高函〔2021〕1号）