

# 人机对话智能系统开发

## 职业技能等级标准

(2021年1.0版)

腾讯云计算（北京）有限责任公司 制定  
2021年4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.....    | 4  |
| 5 面向职业岗位（群）..... | 4  |
| 6 职业技能要求.....    | 5  |
| 参考文献.....        | 11 |

# 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：腾讯云计算（北京）有限责任公司、天津中德应用技术大学、陕西工业职业技术学院、重庆电子工程职业技术学院、黄河水利职业技术学院、许昌职业技术学院、福建信息职业技术学院、西安航空职业技术学院、长江工程职业技术学院、中国移动研究院、智赢未来教育科技有限公司、随机数（浙江）智能科技有限公司。

本标准主要起草人：江楠、李沐、马龙、曹云波、王新强、夏东盛、许磊、丁爱萍、王永乐、卓树峰、龚小涛、吴琦、崔航、刘明辉、葛鹏。

**声明：本标准的知识产权归属于腾讯云计算（北京）有限责任公司，未经腾讯云计算（北京）有限责任公司同意，不得印刷、销售。**

## 1 范围

本标准规定了人机对话智能系统开发职业技能等级对应的工作领域、工作任务及职业技能要求。

本标准适用于人机对话智能系统开发职业技能培训、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。凡是未注日期的引用文件，其最新版本适用于本标准。

人机对话智能系统开发职业技能等级标准中应遵循以下操作规范：

GB/T 36464.1-2020 信息技术 智能语音交互系统 第1部分：通用规范

GB/T 1.1-2009 《标准化工作导则》 第1部分：标准的结构和编写

GB/T 5271.29-2006 信息技术. 词汇. 第29部分:人工智能. 语音识别与合成

GB/T 5271.31-2006 信息技术. 词汇. 第31部分:人工智能. 机器学习

## 3 术语和定义

国家、行业标准界定的以及下列术语和定义适用于本标准。

### 3.1 人工智能 Artificial Intelligence

人工智能是建立在计算机应用基础上的先进技术，它指的是计算机“拥有”人类思想，并且进行模拟、延伸以及拓展人类智能的操作，不需要人类过多干涉，这可以算是一种真正的机器自动化。

### 3.2 人机对话 Man-machine Dialogue

人机对话，是指智能设备与用户通过语音的方式进行对话，对话过程中通过 AI 技术来理解用户的意图，最后向用户提供所需的服务。这种方式与以往的

手机屏幕触摸、PC 电脑的鼠标/键盘的交互模式不一样，更贴近人们实际生活沟通场景。

### 3.3 语音识别 Speech Recognition

语音识别是指将人类语音中的词汇内容转换为计算机可读的输入，例如按键、二进制编码或者字符序列。

### 3.4 语音合成 Text To Speech

语音合成是通过机械的、电子的方法产生人造语音的技术。它是将计算机自己产生的、或外部输入的文字信息转变为可以听得懂的、流利的汉语口语输出。

### 3.5 自然语言处理 Natural Language Processing

自然语言处理是计算机科学领域与人工智能领域中的一个重要方向。它研究能实现人与计算机之间用自然语言进行有效通信的各种理论和方法。自然语言处理是一门融语言学、计算机科学、数学于一体的科学。

### 3.6 意图 Intent

意图是指用户说话的目的，用户通过这句话想要表达什么、想要做什么。

### 3.7 语料 Corpus

语料，是用户为了发起意图说出的自然语句。每个意图都包含一些常用的语料。

### 3.8 槽位 Slot

在用户表达意图的句子中，用来准确表达该意图的关键信息，则称为槽位。槽位是一种变量，帮助语义从机器角度来理解人类的意图。每个槽位关联一个或多个实体库。技能下已有的槽位，可以被这个技能的其它意图直接引用。

### 3.9 技能 Skill

技能，类似于一个 app，通过语音对话完成一个或多个的服务，并返回到用户体验。每个技能包括语义和服务两部分，语义上包括该技能有哪些意图、常用问法、槽位等配置信息，服务上包括服务地址和调用。

### 3.10 多轮对话 Multi-Round Dialogue

多轮对话是指技能在与用户交互过程中不断持续的对话过程。对于需要明确意图关键信息的场景，就需要多轮对话。多轮对话也能带给用户更智能的语音交互体验。

## 4 适用院校专业

中等职业学校：物联网技术应用、软件与信息服务、通信技术、电子与信息技术、电子技术应用、计算机应用等专业。

高等职业学校：电子信息工程技术、云计算技术应用、信息安全技术应用、软件技术、计算机应用技术、大数据技术、计算机网络技术、工业互联网技术、人工智能技术应用、嵌入式技术应用等专业。

应用型本科学校：计算机科学与技术、信息安全、软件工程、网络工程、信息与计算科学、数据科学与大数据技术、人工智能、大数据技术与应用、物联网工程等专业。

## 5 面向职业岗位（群）

**【人机对话智能系统开发】（初级）**：主要面向售前售后工程师、语音标注工程师、人工智能技术支持等职业岗位，主要完成语音获取、数据标注、分析、测试等工作，从事语音获取、数据标注、测试等工作。

**【人机对话智能系统开发】（中级）**：主要面向语音处理工程师、语音产品经理、系统开发工程师等职业岗位，主要完成数据处理、产品设计、应用开发、

测试等工作，从事数据处理、应用开发、产品设计等工作。

【人机对话智能系统开发】（高级）：主要面向人工智能高级开发工程师、架构师、算法工程师等职业岗位，主要完成系统架构设计、业务架构设计、应用开发、产品设计等工作，从事系统架构设计、业务架构设计、应用开发等工作。

## 6 职业技能要求

### 6.1 职业技能等级划分

人机对话智能系统开发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。

【人机对话智能系统开发】（初级）：要面向各类与语音处理、语音分析、人机对话等相关的企事业单位，从事语音获取、标注、分析、应用开发、测试等工作，根据人机对话智能系统开发要求，实现一个简单人机对话系统的能力。

【人机对话智能系统开发】（中级）：主要面向各类与语音处理、语音分析、人机对话等相关的企事业单位，从事数据处理、产品设计、应用开发、测试等工作，根据人机对话智能系统开发要求，具备设计、实现和部署实施人机对话系统的能力。

【人机对话智能系统开发】（高级）：主要面向各类有语音处理、语音分析、人机对话等相关的企事业单位，从事系统架构设计、业务架构设计、应用开发等工作，根据人机对话智能系统开发要求，掌握传统软件的架构方法和模式，具备传统软件向对话式 AI 系统迁移的能力。

### 6.2 职业技能等级要求描述

表 1 人机对话智能系统开发职业技能等级要求（初级）

| 工作领域 | 工作任务 | 职业技能要求 |
|------|------|--------|
|------|------|--------|

|         |              |   |
|---------|--------------|---|
| 语音采集与标注 | 1.1 语音数据采集   | <p>1.1.1 能够熟练使用不同场景下的语音工具或产品。</p> <p>1.1.2 熟练掌握Praat等语音处理工具的使用。</p> <p>1.1.3 采集语音准确率达98%以上。</p> <p>1.1.4 高准确率完成语音采集。</p>  |
|         | 1.2 语音清洗     | <p>1.2.1 能够熟练通过技术清洗音频，筛选出符合条件的音频。</p> <p>1.2.2 熟练掌握Praat等语音处理工具的使用。</p> <p>1.2.3 语音清洗准确率达98%以上，助力语音识别模型训练。</p> <p>1.2.4 高准确率完成语音清洗。</p>   |
|         | 语音标注         | <p>1.3.1 熟练使用语音标注工具，完成语音切分和音素标注。</p> <p>1.3.2 能够熟练根据实际需求，对段落、空格、字符、属性分类等相关问题进行整理。</p> <p>1.3.3 熟练掌握Praat等语音处理工具的使用。</p> <p>1.3.4 按照标注规范对音频内容人工转写成对应文本，对标注结果进行质量审核。</p>  |
| 2. 系统开发 | 2.1 功能设计     | <p>2.1.1 熟练掌握人机对话系统的工作原理和实现流程。</p> <p>2.1.2 掌握轻量级技能服务的开发方法。</p> <p>2.1.3 熟练运用Shell, Python等脚本语言，或者精通java、C++、C等任意一门语言。</p> <p>2.1.4 熟悉MySQL, MongoDB, Elasticsearch 等常用数据库。</p> <p>2.1.5 实现基础的人机对话服务，包括轻量级小技能，内容类技能、自定义有限话术的技能服务。</p> |
|         | 2.2 语音平台认知   | <p>2.2.1 了解云厂商提供的语音相关云语音平台。</p> <p>2.2.2 掌握语音平台等相关功能接口和应用场景。</p> <p>2.2.3 熟悉语音平台的相关文档和资源。</p> <p>2.2.4 可以根据实际项目需求，选择合适的云服务接口进行应用开发。</p>   |
|         | 2.3 语音平台基础调用 | <p>2.3.1 了解语音处理相关基础知识。</p> <p>2.3.2 掌握人机对话平台SDK和API的接入使用。</p> <p>2.3.3 会调用学习模型。</p> <p>2.3.4 实现语音识别功能，包括开放录音文件识别、实时语音识别和一句话识别，基本可以完成客服机器人语音平台。</p>  |
| 3. 系统测试 | 3.1 制定测试方案   | <p>3.1.1 掌握人机对话服务的工作原理和实现流程。</p> <p>3.1.2 熟悉软件测试理论及软件测试各项规范。</p> <p>3.1.3 能够根据项目计划与开发人员沟通，制定功能测试文档。</p> <p>3.1.4 对人机对话产品的功能进行人工测评，生成测试计划文档。</p>   |



|  |            |  |
|--|------------|--|
|  | 3.2 编写测试用例 | <p>3.2.1 熟练掌握功能测试用例设计方法。</p> <p>3.2.2 熟悉常用测试工具及Bug跟踪工具。</p> <p>3.2.3 熟练掌握Excel、Word等常用文档编辑工具的使用。</p> <p>3.2.4 完成智能语音对话系统中功能测试用例。</p>               |
|  | 3.3 执行测试   | <p>3.3.1 精通人机对话服务的模拟测试、真机测试、系统测试方法。</p> <p>3.3.2 掌握单元测试，接口测试的方法，掌握黑白盒测试等。</p> <p>3.3.3 能够对开发人员修复bug后提交的新版本进行版本控制。</p> <p>3.3.4 能够根据测试结果撰写测试报告。</p> |

表 2 人机对话智能系统开发职业技能等级要求（中级）

| 工作领域       | 工作任务            | 职业技能要求   |
|------------|-----------------|--|
| 1. 产品分析与设计 | 1.1 语音交互产品的规划   | <p>1.1.1 了解行业现状，熟悉AI领域语音产品形态。</p> <p>1.1.2 理解产品分析与设计的流程和方法。</p> <p>1.1.3 熟悉Visio、思维导图等软件，熟练撰写产品需求文档、设计文档。</p> <p>1.1.4 完成产品相关文档编写。</p>   |
|            | 1.2 语音交互产品VUI设计 | <p>1.2.1 掌握VUI语音交互界面设计技术。</p> <p>1.2.2 掌握人机对话服务技术，包括基础应用设计、多终端交互设计、语音多轮交互设计等。</p> <p>1.2.3 熟练使用axure、ps、visio等产品设计工具。</p> <p>1.2.4 完成产品功能设计、流程优化以及交互功能优化。</p>                      |
|            | 1.3 定义语音交互过程    | <p>1.3.1 根据用户场景定义交互过程，掌握语音交互业务逻辑。</p> <p>1.3.2 对ASR、NLP、TTS语音全链路有一定了解。</p> <p>1.3.3 深入了解ASR、NLP，TTS等相关业务流程、技术能力及限制，以及产品应对方案。</p> <p>1.3.4 完成音频，文本语料收集、预处理及数据分析，根据用户场景定义语音交互过程。</p> |
| 2. 系统开发    | 2.1 语音数据存储      | <p>2.1.1 精通数据结构和算法设计，熟悉系统性能调优的方式，熟悉 Linux 开发环境。</p> <p>2.1.2 熟练掌握语音数据存储体系、数据处理框架构建，掌握服务的云部署和私有部署。</p> <p>2.1.3 了解 HDFS、HBASE等大数据工具。</p> <p>2.1.4 完成语音数据存储及部署。</p>                  |
|            | 2.2 对话实现        | <p>2.2.1 能够准确理解特定语音应用场景，并能够熟练应用行业工具或算法模型，解决行业实际需求。</p>   |

|         |               |   |
|---------|---------------|---|
|         |               | <p>2.2.2 熟悉Linux、UNIX系统,熟悉tensorflow、pytorch或其他开源深度学习框架。</p> <p>2.2.3 熟悉Spring,Ibatis等服务端开源框架。</p> <p>2.2.4 结合业务需求,利用平台已有对话管理能力,优化人机对话系统的多轮对话、语义识别效果。</p>                               |
|         | 2.3语音平台的调用与优化 | <p>2.3.1 熟练掌握人机对话服务SDK和API应用。</p> <p>2.3.2 掌握语音平台智能闲聊API接口、协议及技术内容。</p> <p>2.3.3 掌握人机对话服务中ASR、NLP、TTS的优化、迭代。</p> <p>2.3.4 实现多轮对话,可以完成聊天机器人语音平台。</p>                                     |
| 3. 系统测试 | 3.1制定测试方案     | <p>3.1.1 掌握人机对话服务的工作原理和实现流程。</p> <p>3.1.2 熟悉软件测试理论及软件测试各项规范。</p> <p>3.1.3 能够根据项目计划与开发人员沟通,制定功能测试、效果测试、接口测试、性能测试的内容、规则、环境等测试计划文档。</p> <p>3.1.4 对人机对话产品的功能、性能、安全、易用性等进行人工测评,生成测试计划文档。</p> |
|         | 3.2 编写测试用例    | <p>3.2.1 熟练掌握功能测试用例设计方法。</p> <p>3.2.2 熟练使用常用测试工具及Bug跟踪工具。</p> <p>3.2.3 熟练掌握Excel、Word等常用文档编辑工具的使用。</p> <p>3.2.4 完成智能语音对话系统中多轮对话、语义识别效果的功能测试、效果测试、接口测试、性能测试的测试用例。</p>                    |
|         | 3.3 产品测试      | <p>3.3.1 掌握单元测试,接口测试的方法,掌握黑白盒测试等。</p> <p>3.3.2 熟练掌握至少一种自动化测试框架(pytest,notestests,unittest等)。</p> <p>3.3.3 了解基于Docker、Jenkins等工具的持续集成方案。</p> <p>3.3.4 执行测试并撰写测试报告</p>                    |

表3 人机对话智能系统开发职业技能等级要求(高级)

| 工作领域       | 工作任务       | 职业技能要求   |
|------------|------------|--|
| 1. 产品设计与规划 | 1.1 业务架构设计 | <p>1.1.1 能够深入理解分析系统需求、规划功能设计,实现产品业务架构设计。</p> <p>1.1.2 熟悉主流人机对话平台的实现机制和典型技术。</p> <p>1.1.3 熟悉深度学习基本原理和语言模型。</p> <p>1.1.4 设计支持多轮对话,知识问答、自由聊天等多场</p> |

|         |                |   |
|---------|----------------|---|
|         |                | 景的人机交互产品，跟踪监控产品研发的质量和进度，对产品测试进行管理，组织产品验收和发布。  |
|         | 1.2 产品迁移       | 1.2.1 能够深入掌握语音产品的服务端、客户端技术。<br>1.2.2 熟悉人机对话服务的领域驱动设计。<br>1.2.3 熟悉主流人机对话平台的实现机制和典型技术。<br>1.2.4 实现多个对话平台之间的上层人机对话系统迁移。  |
|         | 1.3 产品优化       | 1.3.1 能够对语音产品的服务端、客户端不断地进行分析与优化，持续提升产品质量与用户体验。<br>1.3.2 熟悉特定应用场景的语音算法模型，能够针对不同的应用场景，选择合适的工具并解决实际问题。<br>1.3.3 熟练使用一种深度学习框架优化语音模型。<br>1.3.4 传统软件系统向人机对话服务的迁移并实现产品优化。                                    |
| 2. 语音处理 | 2.1 语音模型应用与优化  | 2.1.1 熟练语音识别、语音合成产品系统进行数据处理、模型训练、结果分析、实验验证。<br>2.1.2 能够熟悉使用一种开源工具包（如Kaldi, CMUSphinx）。<br>2.1.3 对语音识别、语音合成有深入的理解，熟悉相关算法。<br>2.1.4 能够对语音识别引擎、合成引擎、声纹识别的开发调优。   |
|         | 2.2 自然语言处理模型   | 2.2.1 理解自然语言进行问题分析、短文本处理、语言模型建立。<br>2.2.2 熟悉自然语言语义工具包如NLTK、Stanford。<br>2.2.3 理解自然语言理解技术和生成技术的相关模型、算法。<br>2.2.4 可以将成熟的自然语言处理技术整合到问答系统中。   |
|         | 2.3 自然语言处理模型优化 | 2.3.1 熟练使用自然语言进行语言模型建立。<br>2.3.2 理解知识图谱相关模型及算法。<br>2.3.3 熟练掌握逻辑回归，决策树等常用模型算法的原理。<br>2.3.4 可以将成熟的自然语言处理技术整合到各类多任务、多轮对话系统中，满足业务实际需求的能力。   |
| 3. 系统开发 | 3.1 架构设计与优化    | 3.1.1 具备语音行业应用的设计与架构能力，对工具选择、性能优化问题能提出独立见解。<br>3.1.2 熟练使用Linux各组件模块，在UI/网络/底层存储介质等方面有很深刻的理解。<br>3.1.3 掌握常用NLP和NLG技术典型应用高并发高可用的架构技术。<br>3.1.4 掌握常用的系统架构，如MVC、SOA、微服务等。<br>3.1.5 完成人机对话服务的系统架构设计、实现及优化。 |

|  |            |  |
|--|------------|--|
|  | 3.2 接口设计   | <p>3.2.1 熟悉和了解系统开发接口设计架构及规范。</p> <p>3.2.2 能够使用主流开发语言开发语音方向的专业工具或行业应用制定接口规范。</p> <p>3.2.3 了解常用的各类开源框架、组件或中间件。</p> <p>3.2.4 能够将给定的模型或算法制定接口规范。</p>   |
|  | 3.3语音平台的实现 | <p>3.3.1 熟悉软件工程设计，开发，测试，部署上线等流程。</p> <p>3.3.2熟练掌握逻辑回归，决策树等常用模型算法的原理和适用范围，并能熟练应用到实际场景中。</p> <p>3.3.3 熟悉并行计算基本原理及分布式计算框架，熟悉Hadoop、Spark等分布式开发环境。</p> <p>3.3.4 实现新技术、新方案的产品难点技术预研与突破，可以通过设计实现多任务对话平台。</p> |

## 参考文献

- [1] 人工智能产业人才发展报告（2019-2020版）
- [2] 教育部下发《普通高等学校高等职业教育（专科）专业目录》确定2019年增补专业共9个，含人工智能技术服务
- [3] 人工智能产业人才岗位能力标准（2019-2020版）
- [4] 普通高等学校本科专业目录（2020年）
- [5] 普通高等学校高等职业教育（专科）专业目录（2015年）
- [6] 中等职业学校专业目录（2010年）
- [7] 中等职业学校专业教学标准（2014年试行）
- [8] 高等职业学校专业教学标准（2018年）
- [9] 本科专业类教学质量国家标准（2018年）
- [10] 国家职业技能标准编制技术规程（2018年版）
- [11] 中华人民共和国职业分类大典（2015年版）
- [12] 战略性新兴产业分类（2018）