

# 《移动端智能交互命名实体标注规范》编制说明

(征求意见稿)

## 一、工作简况

### 1 主要工作过程

#### 1.1 起草（草案、调研）阶段：

2022年5月，成立标准起草工作组，确定主笔人、起草单位，确定工作方法及工作内容。

2022年5月至6月，成立标准编制工作小组，开展课题前期研究工作。

2022年7月，标准起草工作组根据研究结果，形成《移动端智能交互命名实体标注规范》标准初稿。

#### 1.2 征求意见阶段：

2022年7月至8月，组织相关人员对《移动端智能交互命名实体标注规范》标准初稿开展研讨，编制工作组根据研讨意见对文本进行修订、完善，形成《移动端智能交互命名实体标注规范》草案。

#### 1.3 送审阶段：

2022年8月，将《移动端智能交互命名实体标注规范》草案送审。

#### 1.4 报批阶段：

2022年9月，编制工作组按照会议审查意见对标准送审稿作了进一步的修改、整理和完善，形成了标准报批稿、编制说明及其它相关文件，报至中国电工技术学会。

## 2 主要参加单位和工作组成员及其所做的工作

本标准由国网信息通信产业集团有限公司、四川中电启明星信息技术有限公司、福建亿榕信息技术有限公司、北京国网信通埃森哲信息技术有限公司、安徽继远软件有限公司、国网重庆市电力公司电力科学研究院、四川大学、国网重庆市电力公司、中国电力科学研究院有限公司、国网四川省电力公司共同负责起草。

主要成员：李强、赵峰、宋卫平、杨帆、邓芷珊、叶林峰、李欢欢、徐小云、王红蕾、向哲宏、孙觉予、高攀、刘迪、邱镇、庄莉、李炳森、廖逍、黄晓光、刘永清、向辉、许中平、谭洪恩、吕跃春、苏少春、杨迎春、周孔均、钟加勇、彭舰、王秋琳、宋立华、黄飞虎、王金策、田鹏、吕小红、厉仄平、苏江文、刘

礼、邢国用、丘志强、禹国印、杨成、王晓东、张琳瑜、崔迎宝、刘璟、宫晓辉、梁翀、李温静、周伟、季知祥。

所做的工作：标准起草工作组主要成员查阅并收集了移动端智能交互命名实体标注相关资料，通过整理分析，确立了主要规范内容，同时根据专家建议完成对标准的完善。

## 二、标准编制原则和主要内容

### 1、标准编制原则

1.1 本文件的编写格式按 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》的规定进行编写。与相关现行国家标准和行业标准相协调，同时考虑了移动端智能交互命名实体标注的实际情况。

1.2 本文件的起草符合产业发展的原则，本着先进性、科学性、合理性和可操作性的原则以及标准的目标、统一性、协调性、适用性、一致性和规范性原则来进行本文件的起草工作。

### 2、标准主要内容

本指导性技术文件依据团体标准编制要求进行编制，为业务、生产等移动端智能交互中命名实体数据采集、数据标注提供标准可行的技术指导，为数据集中化、标准化提供理论依据和功能支撑，促进数据信息的共享与交互。

标准规定了规范性引用文件、术语和定义、数据样本基本要求、样本标注流程和标注任务展开。

(1) 前言：标准起草单位和主要起草人

(2) 范围：本部分规定了移动端智能交互训练时命名实体（包括狭义和广义命名实体）标注的基本要求、流程和规范。本部分适用于各单位进行移动端智能交互命名实体模型开发时的样本标注和样本入库的统一管理，包括样本的质量管控、样本标注的技术管控和流程管控。

(3) 规范性引用文件：本指导性技术文件引用了国标、行标、地标等，以保证指导性技术文件条款的可依性和可行性。GB/T 5271.28—2001 信息技术 词汇 第 28 部分 人工智能 基本概念与专家系统、GB/T 18521—2001 我国地名分类与分类代码编制规则、Q/GDW 12118.1~12118.3—2021 人工智能平台架构及技术要求、T/CESA 1040—2019 信息技术 人工智能 面向机器学习的数据标注规程。

(4) 术语和定义：对本指导性技术文件中会使用的术语进行定义，对后续内容描述提供了术语支持。包括命名实体、样本数据、标注、标签、分词、标注

工具等。

(5) 缩略语：包括 BIOES、JSON、WAV、XML。

(6) 样本基本要求：对文本文件存储格式、命名要求、质量要求、描述文件进行规定，以保证移动端智能交互命名实体样本数据的规范性。

(7) 样本标注要求：对单类词词性样本、多类词词性样本、实体抽取样本、实体关系抽取样本、事件抽取样本进行规范要求，以及确定标注文件命名与存储要求。

(8) 样本标注流程：样本标注应包含样本检查、安全管控、标注工具选择、标注任务开展、标注结果收集和标注结果检查环节。对不同环节进行阐述。

(9) 标注任务开展：应根据业务需求和命名实体分类进行标注，包括：人名、职名、地名、组织机构名、产品名、事件、时间结构、数量结构、时间结构、数码标注规范。

### 3、主要技术差异

目前国内外关于移动端智能交互命名实体标注规范还没有已经实施或在编的国标、行标也没有其他社会组织的类似团体标准已在实施或在编。本标准定位为团体标准，是对国家标准的补充，与相关技术领域的国家现行法律、法规、规章、政策及相关标准保持一致。

### 4、解决的主要问题

针对新型电力系统的非线性、不确定性、强耦合性、多变量等特性，国家着力建设电力新一代人工智能开放创新平台。面向新型电力系统人工智能开放创新平台移动端用户输入复杂多变的特性，研究移动端用户多意图识别技术、移动端智能交互关键技术，构造智能对话机器人，通过与业务深度融合，为用户提供智能、友好、易用的交互能力，提升移动智能化水平及工作效率，扩大用户流量，增强应用资源服务能力。其中，对移动端智能交互关键技术对命名实体模块制定标注规范，有助于解决打造人工智能平台、提升技术创新研发实力和基础软硬件开放共享服务能力的过程中涉及的资源整合、全流程开放等问题。统一命名实体标注的规范有助于实现电力人工智能技术资源要素的高效整合，促进电力人工智能创新创业及成果转化平台的全环节、全要素对接和资源优化配置，进行全产业链聚合，全流程开放。

### 三、主要试验（或验证）情况

本标准相关内容已在国网信息通信产业集团有限公司、国网重庆市电力公司电力科学研究院、四川中电启明星信息技术有限公司推广实施，并支撑命名实体

标注规范在移动端智能交互领域的应用。

#### **四、标准中涉及专利的情况**

本标准不涉及专利问题。

#### **五、预期达到的社会效益、对产业发展的作用等情况**

本标准批准发布后，将为移动端智能交互命名实体标注提供统一明确的技术规范，进而推动移动端智能交互命名实体的标准化、规范化管理，有利于实现数据信息共享、融合，对于推进移动端智能交互领域融合具有引领意义，有效推动能源企业业务模式创新和数字化转型。

#### **六、与国际、国外对比情况**

本标准没有采用国际、国外标准。

本标准在制定过程中未查到同类国际标准。

本标准在制定时未对国外的样品、样机进行测试。

本标准的总体水平属于国内领先水平。

#### **七、在标准体系中的位置，与现行相关法律、法规、规章及相关标准，特别是强制性标准的协调性**

本标准与相关法律、法规、规章及相关标准协调一致，没有冲突。

#### **八、重大分歧意见的处理经过和依据**

无

#### **九、标准性质的建议说明**

建议本文件以团体标准发布实施，为推荐性标准。

#### 十、贯彻标准的要求和措施建议

建议本标准批准发布 7 天后实施。

#### 十一、废止现行相关标准的建议

无

#### 十二、其他应予说明的事项

无