
中文地址要素解析标注规范

一、概述

1. 中文地址要素解析标注任务需要标注给定地址中要求识别的地址要素（详细列表请查看第二部分）。如：

输入：陕西省渭南市大荔县户家乡边章营村

实体标注：prov=陕西省 city=渭南市 district=大荔县 town=户家乡 community=边章营村

2. 适用范围：中国大陆地址（新疆建设兵团类地址除外），邮件地址。

3. 行政区划标注依据为国家统计局行政区划表：

<http://www.stats.gov.cn/tjsj/tjbz/tjyqhdmhcxhfdm/2020/index.html>

二、标注内容

1、【prov】

解释：省级行政区划，省、自治区、直辖市。注意：港澳台有专有标注规范遇到可整体 others

示例 1：

Query：内蒙古赤峰市锦山镇

实体标注：prov=内蒙古自治区 city=赤峰市 town=锦山镇

示例 2：

Query：陕西省渭南市大荔县户家乡边章营村

实体标注：prov=陕西省 city=渭南市 district=大荔县 town=户家乡 community=边章营村

2、【city】

解释：地级行政区划，地级市、地区、自治州等；

示例 1：

Query：杭州市戴家墩路 91 号东阳诚心木线(富阳店)

实体标注：city=杭州市 road=戴家墩路 roadno=91 号 poi=东阳诚心木线(富阳店)

示例 2：

Query：西藏自治区日喀则地区定日县柑碑村 712 号

实体标注：prov=西藏自治区 city=日喀则地区 district=定日县 community=柑碑村 houseno=712 号

3、【district】

解释：县级行政区划，市辖区、县级市、县等；

示例 1:

Query: 北京市东城区福新东路 245 号

实体标注: city=北京市 district=东城区 road=福新东路 roadno=245 号

4、【devzone】

解释: 广义的上的开发区，除国家级、省级等具备一定行政职能的开发区外，一般性产业园区、度假区等也应标注为开发区。

示例 1:

Query: 内蒙古自治区呼和浩特市土默特左旗金川开发区公元仰山 9 号楼 2 单元 202

实体标注: prov=内蒙古自治区 city=呼和浩特市 district=土默特左旗 devzone=金川开发区 poi=公元仰山 houseno=9 号楼 cellno=2 单元 roomno=202 室

5、【town】

解释: 乡级行政区划，镇、街道、乡等；

示例 1:

Query: 上海 上海市 静安区 共和新路街道 柳营路 669 弄 14 号 1102

实体标注: prov=上海 city=上海市 district=静安区 town=共和新路街道 road=柳营路 road=669 弄 houseno=14 号 roomno=1102

6、【community】

解释: 包含社区、行政村（生产大队、村委会），自然村

示例 1:

Query: 张庆乡北胡乔村

实体标注: town=张庆乡 community=北胡乔村

示例 2:

Query: 五常顾家桥社区河西北 9 号衣服鞋子店

实体标注: town=五常街道 community=顾家桥社区 community=河西北 houseno=9 号 poi=衣服鞋子店

7、【village_group】

解释: 限定 xx 组、xx 队、xx 社 (xx 为数字)。部分村落分设的村民生产队等标注为 community;

示例 1:

Query: 陕西省商洛市镇安县腰庄河村三组

实体标注: prov=陕西省 city=商洛市 district=镇安县 community=腰庄河村 village_group =三组

示例 2:

Query: 沿山村 5 组 6 号

实体标注: community=沿山村 village_group=5 组 houseno=6 号

8、【road】

解释: 有正式名称的道路, 包括隧道、高架、街、弄、巷等。步行街、商业街、统一标注为 road, 除非地图上明显表明是 poi 类别 (只有一栋楼); xx 路段、xx 路南段等整体标注为 road;

示例 1:

Query: 上海市静安区江场三路 238 号 1613 室

实体标注: city=上海市 district=静安区 road=江场三路 roadno=238 号 roomno=1613 室

示例 2:

Query: 江宁区江宁滨江开发区中环大道 10 号环宇人力行政部

实体标注: district=江宁区 devzone=江宁滨江开发区 road=中环大道 roadno=10 号 poi=环宇人力行政部

9、【roadno】

解释: 路号;

示例 1:

Query: 江宁区江宁滨江开发区中环大道 10 号环宇人力行政部

实体标注: district=江宁区 devzone=江宁滨江开发区 road=中环大道 roadno=10 号 poi=环宇人力行政部

10、【poi】

解释: 兴趣点

示例 1:

Query: 阿里巴巴西溪园区

实体标注: poi=阿里巴巴西溪园区

示例 2:

Query: 文一西路阿里巴巴五期

实体标注: road=文一西路 poi=阿里巴巴 subpoi=五期

示例 3:

Query: 丽景天成 2 期仙林苑

实体标注: poi=丽景天成 subpoi=2 期 subpoi= 仙林苑

11、【subpoi】

解释：子兴趣点：

示例 1：

Query：新疆维吾尔自治区 昌吉回族自治州 昌吉市 延安北路街道 延安南路石油小区东门

实体标注：prov=新疆维吾尔自治区 city=昌吉回族自治州 district=昌吉市 town=延安北路街道 road=延安南路 poi=石油小区 subpoi=东门

示例 2：

Query：西湖区新金都城市花园西雅园 10 幢 3 底层

实体标注： district=西湖区 poi=新金都城市花园 subpoi=西雅园 houseno=10 幢 floorno=3 底层

示例 3：

Query：广宁伯街 2 号金泽大厦东区 15 层

实体标注： road=广宁伯街 roadno=2 号 poi=金泽大厦 subpoi=东区 floorno=15 层

12、【houseno】

解释：楼栋号，农村地址的门牌号(包括类似南楼、北楼一类的描述)

示例 1：

Query：阿里巴巴西溪园区 6 号楼小邮局

实体标注： poi=阿里巴巴西溪园区 houseno=6 号楼 subpoi=小邮局

示例 2：

Query：四川省 成都市 金牛区 沙河源街道 金牛区九里堤街道 金府机电城 A 区 3 栋 16 号

实体标注：prov=四川省 city=成都市 district=金牛区 town=沙河源街道 district=金牛区 town=九里堤街道 poi=金府机电城 subpoi=A 区 houseno=3 栋 detail=16 号

示例 3：

Query：竹海水韵春风里 12 幢 3 单元 1001 室

实体标注： poi=竹海水韵 subpoi=春风里 houseno=12 幢 cellno=3 单元 roomno= 1001 室

13、【cellno】

解释：单元号； 包括“东西”“甲乙丙丁”一类的单元号

示例 1：

Query：竹海水韵春风里 12 幢 3 单元 1001 室

实体标注： poi=竹海水韵 subpoi=春风里 houseno=12 幢 cellno=3 单元 roomno= 1001 室

示例 2：

Query：蒋村花园新达苑 18 幢二单元 101

实体标注： poi=蒋村花园 subpoi=新达苑 houseno=18 幢 cellno=二单元 roomno=101 室

14、【floorno】

解释：楼层号；

示例 1：

Query：北京市东城区东中街 29 号东环广场 B 座 5 层信达资本

实体标注：city=北京市 district=东城区 road=东中街 roadno=29 号 poi=东环广场
houseno=B 座 floorno=5 层 person=信达资本

15、【roomno】

解释：房间号、户号；

05 商铺、803 房间、304 寝室、房、东户、西户、中户、门市等

对于结构很明确，如下例子 301 可以标注 roomno。

示例 1：

Query：西溪北苑南区 2 栋 1 单元 301

实体标注：poi=西溪北苑 subpoi=南区 houseno=2 栋 cellno=1 单元 roomno=301

16、【detail】

解释：poi 内部的四层关系（house, cell, floor, room）没明确是哪一层，如 xx-xx-x-x，则整体标注 detail。

示例 1：

Query：竹海水韵春风里 12-3-1001

实体标注：poi=竹海水韵 subpoi=春风里 detail=12-3-1001

17、【assist】

解释：普通辅助定位词，

示例 1：

Query：广西柳州市城中区潭中东路勿忘我网吧附近

实体标注：prov=广西 city=柳州市 district=城中区 road=潭中东路 poi=勿忘我网吧
assist=附近

Query：浙江省 杭州市 滨江区 六和路东 50 米 六和路东信大道口东北自行车租赁点

实体标注：prov=浙江省 city=杭州市 district=滨江区 road=六和路 assist=东
distance=50 米 road=六和路 road=东信大道 intersection=口 assist=东北 poi=自行车
租赁点

18、【distance】

解释：距离辅助定位词，比如“716 县道北 50 米”中的 50 米，具有具体数字；模糊距离，如出现单位词“米”、“千米”等，可以标注 distance，如“几十米”，如只是定性描述，如“不远”、“快到”都标注为 assist

示例 1:

Query: 浙江省 杭州市 滨江区 六和路东 50 米 六和路东信大道口东北自行车租赁点

实体标注: prov=浙江省 city=杭州市 district=滨江区 town=浦沿街道 road=六和路 direction=东 distance=50 米 road=东信大道 intersection=口 direction=东北 poi=自行车租赁点

示例 2:

Query: 浙江省 杭州市 上城区 劳动路与红门局交叉口东 50 米 鹰记皮具定制

实体标注: prov=浙江省 city=杭州市 district=上城区 town=清波街道 road=劳动路 redundant=与 road=红门局 intersection=交叉口 direction=东 distance=50 米 subpoi=鹰记皮具定制

19、【intersection】

解释：道路口，口、交叉口、道路（高速）出入口，一定与 road 同时出现，注意：小区出入口和车库出入口为 poi，“与”“和”两条路中间的修饰词为 redundant

示例 1: 见 distance 的示例

20、【redundant】

解释：非地址元素，如配送提示、配送要求、友情提醒或威胁，等人名、邮编、电话。退货订单信息。胡乱填写。（重复元素正常标注）标点和特殊无用符号为 redundant，无需标注。

示例 1:

a. Query: 浙江省 杭州市 上城区 劳动路与红门局交叉口东 50 米 鹰记皮具定制

b. 实体标注: prov=浙江省 city=杭州市 district=上城区 road=劳动路 redundant=与 road=红门局 intersection=交叉口 assist=东 distance=50 米 poi=鹰记皮具定制 redundant=来时记得打电话（此为示例说明，实际标注系统中无需对此片段标注 redundant 标记，可以一定程度上加快标注速度。）

示例 2:

a. Query: 浙江省 杭州市 滨江区 六和路东 50 米 六和路东信大道口东北自行车租赁点

b. 实体标注: prov=浙江省 city=杭州市 district=滨江区 road=六和路 assist=东 distance=50 米 road=东信大道 intersection=口 assist=东北 poi=自行车租赁点 redundant=有人在那里等候

备注：括号里的内容出现地址正常标注，干扰信息标注 redundant。

21、【others】

解释：以上标签未覆盖的情况，或者无法判断的地址元素（此标签标注的一定是地址元素，如山、岛等，注意与 redundant 的区分），标注 others

如：港澳台地址