流行语数据填写指南

目录

快速开始

第一步：打开数据文件

编辑 `buzzwords.json` 文件

第二步：添加新词条

在 `buzzwords` 对象中添加新的词条，词条ID使用小写字母：

```
{
  "buzzwords": {
    "词条id": { ...词条数据... },
    "新词条id": { ...新词条数据... }
  }
}
```

第三步：更新主页搜索列表

编辑 `js/main.js` 文件（第97行附近）：

```
this.buzzwords = ["brat", "yyds", "躺平", "内卷", "新词条id"];
```

数据结构说明

每个流行语词条包含以下主要部分：

```
词条
├── 基本信息（headword, language, partOfSpeech等）
├── 发音信息（pronunciation）
├── 义项数组（senses）
│   ├── 流行义（trending）
│   ├── 现代主流义（common）
│   │   └── 子义项（subsenses）
│   └── 其他义（other）
│       └── 子义项（subsenses）
├── 词频数据（frequencyData）
├── 历史信息（history）
├── 来源（source）
└── 相关词汇（relatedTerms）
```

字段详细说明

1. 基本信息字段

headword（必需）

- **类型**：字符串
- **说明**：词条的标题
- **示例**：`"brat"`, `"YYDS"`, `"躺平"`

language（必需）

- **类型**：对象
- **说明**：来源语言，中英双语
- **格式**：

```
"language": {
  "en": "English",
  "zh": "英语"
}
```

partOfSpeech（必需）

- **类型**：对象
- **说明**：词类，中英双语
- **示例**：

```
"partOfSpeech": {
  "en": "noun",
  "zh": "名词"
}
```

常用词类：

- noun/名词
- verb/动词
- adjective/形容词
- adverb/副词
- abbreviation/缩写

pronunciation（必需）

- **类型**：对象
- **说明**：发音信息，支持多种语言

英语词条格式：

```
"pronunciation": {
  "uk": "/brat/",
  "us": "/bræt/",
  "audio": {
    "uk": "assets/audio/brat-uk.mp3",
    "us": "assets/audio/brat-us.mp3"
  }
}
```

中文词条格式：

```
"pronunciation": {
  "pinyin": "tǎng píng",
  "audio": {
    "zh": "assets/audio/tangping-zh.mp3"
  }
}
```

俄语词条格式：

```
"pronunciation": {
  "ipa": "/ˈtɔvarʲɪɕ/",
  "romanization": "tovarisch",
  "audio": {
    "ru": "assets/audio/tovarisch-ru.mp3"
  }
}
```

德语词条格式：

```
"pronunciation": {
  "ipa": "/ˈaɪ̯tʃaɪ̯t/",
  "audio": {
    "de": "assets/audio/zeitgeist-de.mp3"
  }
}
```

日语词条格式：

```
"pronunciation": {
  "romaji": "kawaii",
  "hiragana": "かわいい",
  "audio": {
    "ja": "assets/audio/kawaii-ja.mp3"
  }
}
```

韩语词条格式：

```
"pronunciation": {
  "romanization": "hallyu",
  "hangul": "한류",
  "audio": {
    "ko": "assets/audio/hallyu-ko.mp3"
  }
}
```

发音字段说明：

字段	用途	适用语言
`uk`	英式发音（IPA）	英语
`us`	美式发音（IPA）	英语
`pinyin`	拼音	中文
`ipa`	国际音标	所有语言（通用）
`romanization`	罗马化拼写	俄语、韩语等
`romaji`	罗马字	日语
`hiragana`	平假名	日语
`hangul`	韩文	韩语

音频文件语言代码：

代码	语言	文件名示例
`uk`	英国英语	`word-uk.mp3`
`us`	美国英语	`word-us.mp3`
`zh`	中文	`word-zh.mp3`
`ru`	俄语	`word-ru.mp3`
`de`	德语	`word-de.mp3`
`fr`	法语	`word-fr.mp3`
`ja`	日语	`word-ja.mp3`
`ko`	韩语	`word-ko.mp3`
`es`	西班牙语	`word-es.mp3`
`ar`	阿拉伯语	`word-ar.mp3`

注意事项：

- 音频文件是可选的，如果没有音频，可以省略 `audio` 字段
- 至少要提供一种发音表示方式（IPA、拼音、罗马化等）

- 可以同时提供多种发音方式（如IPA + romanization）
- 如果不确定使用哪个字段，统一使用 `ipa` 字段（国际音标）最通用

firstRecorded（可选）

- **类型**：对象
- **说明**：首次记录时间

```
"firstRecorded": {
  "en": "2021, Chinese internet forums",
  "zh": "2021年网络论坛"
}
```

trendingPeriod（必需）

- **类型**：字符串
- **说明**：流行时间
- **示例**：`"2024"`, `"2021-2023"`

numberOfSenses（必需）

- **类型**：数字
- **说明**：义项总数（包括所有子义项）
- **示例**：`7`, `3`

---

2. 义项（senses）

义项类型（type）

- `"trending"` - 流行义
- `"common"` - 现代主流义
- `"other"` - 其他义

流行义示例

```
{
  "type": "trending",
  "title": {
    "en": "Trending Sense",
    "zh": "流行义"
  },
  "label": {
    "en": "Confident, Independent Person or Attitude",
    "zh": "自信独立者"
  },
  "definition": {
    "en": "A confident, independent, and unapologetic person...",
    "zh": "指一种自信、独立、无拘无束的个体..."
  },
  "examples": [ ...例句数组... ],
  "etymology": {
    "en": "Originally emerged from...",
    "zh": "最初出现在..."
  }
}
```

注意: `etymology` 字段可以为 `null`，如果没有词源信息。

带子义项的义项示例

```json
{
  "type": "common",
  "title": {
    "en": "Common Modern Sense",
    "zh": "现代主流义"
  },
  "subsenses": [
    {
      "label": {
        "en": "Spoiled or Mischievous Child",
        "zh": "被宠坏的孩子"
      },
      "definition": {
        "en": "A child, especially one who is ill-mannered...",
        "zh": "指无礼、顽皮或被宠坏的孩子..."
      },
      "examples": [ ...例句数组... ]
    }
  ],
  "etymology": {
    "en": "The origin is uncertain...",
    "zh": "词源不确定..."
  }
}
```

3. 例句 (examples)

每个例句包含完整的句子和位置信息:

```json
{
  "year": 2024,
  "sentence": {
    "en": "To be a brat is to be confident and unapologetic...",
    "zh": "\"做一个 brat\"意味着要自信且毫不道歉..."
  },
  "source": {
    "author": "Gwen Tam",
    "title": "BRAT: a shiny, lime green breakdown",
    "url": "https://berkeleybside.com/brat-a-shiny-lime-green-breakdown/"
  },
  "keyword": "brat",
  "keywordPosition": {
    "start": 5,
    "end": 9
  }
}
```

计算 keywordPosition

方法：在英文句子中找到关键词的位置

示例句子：`"To be a brat is to be confident..."`

1. 找到关键词 `"brat"` 的位置
2. 从0开始计数字符
3. `"To be a "` = 8个字符，所以 start = 8
4. `"brat"` = 4个字符，所以 end = 8 + 4 = 12

快速计算工具 (Python)：

```python
sentence = "To be a brat is to be confident..."
keyword = "brat"
start = sentence.find(keyword)
end = start + len(keyword)
print(f"start: {start}, end: {end}")
```

注意：

- 如果没有URL，可以设置为空字符串 `""`
- `source` 字段不能省略，至少要有 author

---

4. 词频数据（frequencyData）

记录每年的使用频次：

```json
"frequencyData": [
  {"year": 2020, "count": 150},
  {"year": 2021, "count": 300},
  {"year": 2022, "count": 800},
  {"year": 2023, "count": 2500},
  {"year": 2024, "count": 15000}
]
```

注意：

- 至少提供3个年份的数据
- 年份按时间顺序排列
- count 为整数

---

5. 历史信息（history）

```json
"history": {
  "en": "The term 'brat' has undergone a significant semantic shift...",
  "zh": "\"brat\"一词在2024年经历了显著的语义转变..."
}
```

---

6. 来源 (source)

```
"source": {
  "en": "Global Buzzwords Report 2024",
  "zh": "《2024年全球流行语报告》"
}
```

7. 相关词汇 (relatedTerms)

```
"relatedTerms": ["aesthetic", "confidence", "authenticity", "rebellion"]
```

建议：3-6个相关词汇

填写示例

示例1：简单的中文流行语

```
"摆烂": {
  "headword": "摆烂",
  "language": {
    "en": "Chinese",
    "zh": "中文"
  },
  "partOfSpeech": {
    "en": "verb",
    "zh": "动词"
  },
  "pronunciation": {
    "pinyin": "bǎi làn"
  },
  "firstRecorded": {
    "en": "2022, Chinese internet",
    "zh": "2022年网络流行"
  },
  "trendingPeriod": "2022-2024",
  "numberOfSenses": 1,
  "senses": [
    {
      "type": "trending",
      "title": {
        "en": "Trending Sense",
        "zh": "流行义"
      },
      "label": {
        "en": "Give Up and Let Things Fall Apart",
        "zh": "放弃努力，破罐破摔"
      },
      "definition": {
        "en": "To give up trying and let a situation deteriorate, often used to describe a passive-aggressive response to difficulties.",
        "zh": "放弃努力，任由事态恶化，常用于描述面对困难时的消极态度。"
      },
```

```json
      "examples": [
        {
          "year": 2023,
          "sentence": {
            "en": "I'm just going to give up trying. (Wo jiu bai lan le.)",
            "zh": "我就摆烂了。"
          },
          "source": {
            "author": "Social Media User",
            "title": "Weibo Post",
            "url": ""
          },
          "keyword": "摆烂",
          "keywordPosition": {
            "start": 2,
            "end": 4
          }
        }
      ],
      "etymology": {
        "en": "Originally from gaming culture, meaning to intentionally play poorly.",
        "zh": "原本来自游戏文化，指故意打得很差。"
      }
    }
  ],
  "frequencyData": [
    {"year": 2022, "count": 5000},
    {"year": 2023, "count": 25000},
    {"year": 2024, "count": 18000}
  ],
  "history": {
    "en": "The term originated in gaming communities and spread to describe general life attitudes.",
    "zh": "该词源于游戏社区，后扩展用于描述一般的生活态度。"
  },
  "source": {
    "en": "Chinese Internet Slang Report 2023",
    "zh": "《中国网络流行语报告2023》"
  },
  "relatedTerms": ["躺平", "佛系", "放弃", "破罐破摔"]
}
```

示例2：英语流行语

```json
"rizz": {
  "headword": "rizz",
  "language": {
    "en": "English",
    "zh": "英语"
  },
  "partOfSpeech": {
    "en": "noun",
    "zh": "名词"
  },
  "pronunciation": {
    "uk": "/rɪz/",
```

```json
    "us": "/rɪz/"
  },
  "firstRecorded": {
    "en": "2021, TikTok",
    "zh": "2021年，TikTok"
  },
  "trendingPeriod": "2023-2024",
  "numberOfSenses": 1,
  "senses": [
    {
      "type": "trending",
      "title": {
        "en": "Trending Sense",
        "zh": "流行义"
      },
      "label": {
        "en": "Charisma or Charm",
        "zh": "魅力，吸引力"
      },
      "definition": {
        "en": "Charisma or charm, especially in a romantic or social context. Short for 'charisma'.",
        "zh": "魅力或吸引力，尤其在浪漫或社交场合。是charisma的缩略。"
      },
      "examples": [
        {
          "year": 2023,
          "sentence": {
            "en": "He's got so much rizz, everyone loves him.",
            "zh": "他太有魅力了，所有人都喜欢他。"
          },
          "source": {
            "author": "TikTok User",
            "title": "Viral Video",
            "url": ""
          },
          "keyword": "rizz",
          "keywordPosition": {
            "start": 17,
            "end": 21
          }
        }
      ],
      "etymology": null
    }
  ],
  "frequencyData": [
    {"year": 2021, "count": 100},
    {"year": 2022, "count": 2000},
    {"year": 2023, "count": 50000},
    {"year": 2024, "count": 30000}
  ],
  "history": {
    "en": "Popularized on TikTok and became Oxford Word of the Year 2023.",
    "zh": "在TikTok上流行，并成为2023年牛津年度词汇。"
  },
  "source": {
    "en": "Oxford Dictionary 2023",
```

```
    "zh": "《牛津词典2023》"
  },
  "relatedTerms": ["charisma", "charm", "game", "vibe"]
}
```

常见问题

Q1: 如果没有音频文件怎么办?

A: 省略 `audio` 字段即可:

```
"pronunciation": {
  "uk": "/brat/",
  "us": "/bræt/"
}
```

Q2: 如果没有词源信息怎么办?

A: 将 `etymology` 设为 `null`:

```
"etymology": null
```

Q3: 子义项没有例句怎么办?

A: 将 `examples` 设为空数组:

```
"examples": []
```

Q4: 如何确定义项数量?

A: 统计所有义项和子义项的总数:

- 1个trending义项
- 1个common义项(包含2个subsense)
- 1个other义项(包含3个subsense)
- 总数 = 1 + 1 + 2 + 1 + 3 = 8

Q5: keywordPosition 计算错误会怎样?

A: 关键词高亮位置会错误,但不影响整体功能。建议使用Python脚本计算。

检查清单

在提交数据前,请检查:

- ☐ 词条ID使用小写字母
- ☐ 所有必需字段都已填写
- ☐ 多语言字段包含 `en` 和 `zh` 两个键
- ☐ 例句的 `keywordPosition` 计算正确
- ☐ `frequencyData` 至少包含3个年份

- [ ] `numberOfSenses` 数量正确
- [ ] JSON 格式正确（无语法错误）
- [ ] 已更新 `js/main.js` 中的搜索列表

JSON 验证

在线验证工具：https://jsonlint.com/

本地验证（Python）：

```python
import json

with open('buzzwords.json', 'r', encoding='utf-8') as f:
    data = json.load(f)
    print("☑ JSON格式正确！")
    print(f"共有 {len(data['buzzwords'])} 个词条")
```