

语言标记是一些短小的标准字符串，用来命名所使用的语言——比如，fr（法语）和 en-GB（英式英语）。每个标记都由一个或多个称为子标记的部分组成，中间由连字符分隔。16.4 节曾详细介绍了语言标记。

本附录总结了语言标记的规则、标准化的标记以及注册信息。包含下列参考资料：

- G.1 节总结了第一个（主）子标记所用规则；
- G.2 节总结了第二个子标记所用规则；
- 表 G-1 显示了已在 IANA 注册的语言标记；
- 表 G-2 列出了 ISO 639 语言代码；
- 表 G-3 列出了 ISO 3166 国家代码。

G.1 第一个子标记所用规则

如果第一个子标记为：

- 2 个字符长，是 ISO 639¹ 和 639-1 标准的语言代码；
- 3 个字符长，是 ISO 639-2² 标准中列出的语言代码；
- 字母 “i”，语言标记就是在 IANA 中显式注册过的语言标记；
- 字母 “x”，语言标记就是私有的、非标准的扩展子标记。

581 表 G-2 总结了 ISO 639 和 639-2 的名称。

G.2 第二个子标记所用规则

如果第二个子标记为：

- 2 个字符长，就是 ISO 3166³ 定义的国家 / 地区码；
- 3 ~ 8 个字符长，是已在 IANA 注册过的语言标记；
- 1 个字符长，就是非法的。

表 G-3 总结了 ISO 3166 国家代码。

注 1：参见 ISO 标准 639，“Codes for the representation of names of languages”。

注 2：参见 ISO 639-2，“Codes for the representation of names of languages—Part 2: Alpha-3 code”（语言名称表示码——第 2 部分：Alpha-3 代码）。

注 3：ISO 3166 将国家代码 AA、QM-QZ、XA-XZ 和 ZZ 保留作为用户分配的代码。不能用它们来构造语言标记。