

文本处理服务

本章介绍的模块提供了广泛的字符串操作和其他文本处理服务。

在 [二进制数据服务](#) 之下描述的 `codecs` 模块也与文本处理高度相关。此外也请参阅 Python 内置字符串类型的文档 [文本序列类型 --- str](#)。

- [string --- 常见的字符串操作](#)
 - 字符串常量
 - 自定义字符串格式化
 - 格式字符串语法
 - 格式规格迷你语言
 - 格式示例
 - 模板字符串 (\$-字符串)
 - 辅助函数
 - [string.Template --- 对模板字符串字面值的支持](#)
 - 模板字符串
 - 类型
 - 辅助函数
 - [re --- 正则表达式操作](#)
 - 正则表达式语法
 - 模块内容
 - 标志
 - 函数
 - 异常
 - 正则表达式对象（正则对象）
 - 匹配对象
 - 正则表达式例子
 - 检查对子
 - 模拟 `scanf()`
 - `search()` vs. `match()`
 - 制作一个电话本
 - 文字整理
 - 查找所有副词
 - 查找所有的副词及其位置
 - 原始字符串标记
 - 写一个词法分析器
- [difflib --- 计算差异的辅助工具](#)
 - SequenceMatcher 对象
 - SequenceMatcher 的示例
 - Differ 对象

- Differ 示例
- difflib 的命令行接口
- ndiff 示例
- `textwrap` --- 文本自动换行与填充
- `unicodedata` --- Unicode 数据库
- `stringprep` --- 因特网字符串预处理
- `readline` --- GNU readline 接口
 - 初始化文件
 - 行缓冲区
 - 历史文件
 - 历史列表
 - 启动钩子
 - Completion
 - 示例
- `rlcompleter` --- 用于 GNU readline 的补全函数