

fuscan

极速通用文件扫描器

用户手册 · 版本 0.1.7

适用对象：初级用户 · GUI 图形界面

目录

1. 软件简介
 2. 安装与启动
 3. 主界面总览
 4. 第一次扫描
 5. 查看扫描结果
 6. 详情区与关键词高亮
 7. 导出扫描结果
 8. 规则管理
 9. 设置
 10. 扫描历史
 11. 快捷键
 12. 常见问题
-

1. 软件简介

fuscan 能做什么：

- 多格式扫描：PDF、Word (.docx)、Excel (.xlsx)、PowerPoint (.pptx)、ODF (.odt/.ods)、WPS、纯文本，以及 ZIP/RAR 压缩包内文件
- 规则灵活：按文件名、文件内容、路径三类目标匹配，支持 contains/equals/regex 等模式，可 AND/OR/NOT 逻辑组合
- 内置规则：随包分发 8 条通用安全规则，开箱即用
- 实时反馈：扫描进度、当前文件、命中数、错误数实时显示
- 结果可视化：按规则/严重等级分组，关键词高亮定位，一键跳转
- 导出归档：扫描结果可导出为 CSV 或 JSON

2. 安装与启动

安装

```
pip install fuscan
```

或使用 uv：

```
uv add fuscan
```

注意：GUI 功能需要 PySide2，仅在 Python 3.8~3.10 可用。

若使用 Python 3.11+，会自动使用 PySide6。

启动 GUI

命令行输入：

```
fuscan gui
```

即可打开图形界面。

3. 主界面总览

主界面采用 GitHub Desktop 风格，从上到下分为五个层次：

1. 菜单栏：文件 / 扫描 / 视图 / 帮助
2. 顶部操作区：扫描模式选择、扫描路径、规则加载、扫描按钮
3. 列表区（Tab 切换）：扫描结果 / 规则文件 / 扫描历史
4. 详情区：选中文件后展示文件信息、命中规则表、内容预览
5. 状态栏：扫描统计、进度条、当前文件名

左侧另设侧边栏，用于在工作流阶段之间切换并高亮当前阶段。

工作流三阶段

主界面通过整页切换表达三个阶段，侧边栏同步高亮当前阶段：

1. 配置页：选择扫描模式、加载规则、点击"开始扫描"
2. 扫描中页：显示进度、当前文件、命中列表、跳过目录
3. 结果页：浏览命中结果、查看详情、导出报告

顶部 Tab

顶部有三个 Tab 按钮：

- 扫描：主扫描工作流（默认）
- 规则：规则文件管理
- 历史：扫描历史记录

4. 第一次扫描

步骤 1：选择扫描模式

在配置页的"扫描模式"下拉框选择：

- 全盘扫描：扫描所有盘符（耗时较长）
- 盘符扫描：选择具体盘符（如 C:、D:）
- 文件夹扫描：选择具体文件夹（最常用）

选择"文件夹扫描"后，点击右侧文件夹图标按钮选择目录。

步骤 2：加载规则（可选）

软件已内置通用规则，默认启用。如需使用自定义规则：

1. 点击"加载规则"按钮
 2. 选择 YAML 规则文件（.yaml 或 .yml）
 3. 规则文件会出现在规则列表中，可加载多个
- 内置规则与用户规则会自动合并，用户规则可覆盖内置规则中的同名规则。

步骤 3：开始扫描

点击右上角的"开始扫描"按钮（或菜单：扫描 → 开始扫描，快捷键 Ctrl+S）。

扫描期间：

- 状态栏显示进度条、当前文件名、已扫描/命中/错误数、已用时
- "扫描中"页实时显示命中文件列表与跳过目录列表
- 可点击"暂停扫描"临时暂停，"继续扫描"恢复
- 可点击"取消扫描"终止本次扫描（已扫描结果会保留）

步骤 4：查看结果

扫描完成后自动跳转到"结果页"。详见下一节。

5. 查看扫描结果

结果树结构

结果区以树形展示命中：

- 不分组：每个命中文件为顶层项，其下展开命中的规则
- 按规则分组：规则名为顶层项，命中的文件为子项
- 按严重等级分组：等级为顶层项（严重/警告/信息），文件为子项

通过分组下拉框切换分组模式。

严重等级

命中结果按严重等级着色：

等级	颜色	含义
严重 (CRITICAL)	红色	高危，需立即处理
警告 (WARNING)	橙色	中等风险，建议关注
信息 (INFO)	蓝色	提示性信息

严重等级的整行会以浅色背景高亮，便于快速识别高危项。

筛选

结果区顶部提供三个筛选框：

- 路径筛选：按文件路径关键词过滤（支持部分匹配）
- 规则筛选：下拉选择具体规则，只显示该规则的命中

- 分组模式：切换不分组/按规则/按严重等级

选中查看详情

单击结果树中的文件项，下方详情区会显示该文件的详细信息与内容预览。

双击文件项会在独立窗口打开详情对话框，便于放大查看。

6. 详情区与关键词高亮

详情区位于结果区下方，包含三部分：

文件信息

显示文件路径、大小、命中数、可切换位置数等元信息。

命中规则表

列出该文件命中的所有规则：

列	含义
规则	规则名称
严重	严重等级
次数	该规则在文件中的匹配次数
位置	在预览中可高亮跳转的位置数（"-" 表示仅匹配文件名，无内容位置）
详情	规则命中的详细说明

点击命中规则表某行，预览区会自动跳转到该规则对应的高亮位置。

内容预览与关键词高亮

预览区显示文件内容（支持多格式提取），命中关键词以黄色背景高亮。

当前选中的位置以橙色背景突出显示。

命中导航：

- 上一条：点击"上一条"按钮或按 Shift+F3
- 下一条：点击"下一条"按钮或按 F3
- 导航标签显示当前位置：如 "3 / 12"

若文件内容超过 100KB，预览会截断并提示"内容已截断"。

若规则仅匹配文件名（如"敏感文件名检测"），预览区会提示无内容关键词可高亮。

7. 导出扫描结果

扫描完成后，可将结果导出为四种格式的文件：

1. 在结果页点击"导出"按钮

2. 在弹出的格式选择对话框中选取目标格式

3. 选择保存位置与文件名（默认 `fuscan_report.<■■■■>`）

CSV 格式（.csv）：表格形式，适合用 Excel 打开分析

JSON 格式（.json）：结构化数据，适合程序化处理

PDF 格式（.pdf）：含标题、扫描统计与命中文件表格，适合汇报与归档

Excel 格式（.xlsx）：双工作表，"扫描汇总"含统计、"命中明细"含命中记录并按严重等级着色

命令行同样支持：

```
fuscan scan /path -o pdf -f report.pdf
fuscan scan /path -o excel -f report.xlsx
```

导出仅包含命中的文件，不含未命中的文件。

8. 规则管理

切换到顶部"规则"Tab 进入规则管理界面。

规则列表

左侧显示已加载的规则文件列表，右侧显示当前合并后的规则集：

- 规则名称、严重等级、适用扩展名

加载规则文件

点击"加载规则"按钮或菜单：文件 → 加载规则（快捷键 Ctrl+O），
选择 YAML 文件后加入列表。

调整规则优先级

多个规则文件按顺序链式合并，后加载的规则文件覆盖先加载的同名规则。

通过右键菜单或快捷键调整顺序：

- 上移：右键 → 上移
- 下移：右键 → 下移
- 移除：右键 → 移除（快捷键 Delete）

编辑规则

点击"编辑规则"按钮或菜单：文件 → 编辑规则，打开规则编辑器对话框，
可修改当前加载的规则文件内容。

9. 设置

通过菜单：文件 → 设置（快捷键 Ctrl+,）打开设置对话框。

常用设置项：

- 缓存：启用扫描缓存可大幅提升重复扫描速度（推荐开启）

- 缓存路径：默认在用户目录下，可自定义
- 扫描压缩包：是否扫描 ZIP/RAR 内文件
- 压缩包密码：扫描加密压缩包时使用的密码
- 并发线程数：多线程扫描，0 表示单线程，建议 4-8
- 扫描深度：限制目录递归深度，空表示不限
- 忽略目录：不扫描的目录名（如 .git、node_modules）
- 忽略扩展名：不扫描的文件扩展名
- 显示网络驱动器：盘符列表是否显示网络驱动器

设置修改后自动保存，下次启动恢复。

10. 扫描历史

切换到顶部"历史"Tab 查看扫描历史。

- 历史列表记录最近扫描过的路径（最多 20 条）
- 双击历史项可快速切换到文件夹模式并扫描该路径
- 路径不存在时会提示

11. 快捷键

快捷键	功能
Ctrl+O	加载规则文件
Ctrl+S	开始扫描
Ctrl+,	打开设置
F3	跳转到下一个命中位置
Shift+F3	跳转到上一个命中位置
Delete	移除选中的规则文件

12. 常见问题

Q: 扫描很慢怎么办？

- 启用缓存（设置 → 启用扫描缓存）
- 缩小扫描范围（用文件夹扫描代替全盘扫描）
- 适当增加并发线程数（设置 → 并发线程数：4-8）
- 忽略无关目录（设置 → 忽略目录：如 .git、node_modules、__pycache__）

Q: 扫描时界面卡顿？

- 进度回调已做节流，正常情况不应卡顿
- 若仍卡顿，可能是命中文件列表过长，可关闭"扫描中"页的列表实时更新

Q: 找不到某些文件？

- 检查文件扩展名是否在"忽略扩展名"列表中
- 检查所在目录是否在"忽略目录"列表中
- 检查扫描深度是否受限
- 若在压缩包内，确认"扫描压缩包"已开启

Q: 加密压缩包扫描不出内容？

- 在设置中填写"压缩包密码"
- 注意：fscan 仅支持 ZIP 标准加密格式

Q: 规则不生效？

- 切换到"规则"Tab 检查规则是否正确加载
- 检查规则的 `file_extensions` 是否包含目标文件扩展名
- 尝试菜单：文件 → 编辑规则，确认规则语法正确

Q: 如何查看历史扫描结果？

fscan 目前不保存历史扫描结果，仅记录扫描路径。每次扫描会覆盖上次结果。

若需归档，请扫描完成后导出为 CSV 或 JSON。

Q: 内置规则有哪些？

内置 8 条通用安全规则，覆盖：

- 敏感文件名检测（password、secret、key 等）
- AWS 密钥泄露检测
- 数据库连接字符串
- 私钥文件检测
- 其他常见敏感信息

可在"规则"Tab 查看，或通过设置关闭"使用通用规则"仅用自定义规则。

如需更多信息，请参阅项目 README

或通过菜单：帮助 → 关于 查看版本信息。