



链滴

链接熔炉设计

作者: [88250](#)

原文链接: <https://ld246.com/article/1473238756829>

来源网站: [链滴](#)

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

背景

很多人都有收藏网络资料的习惯，常用的方式有两种：

1. 浏览器书签
2. 笔记软件

对于个人的知识管理这两种方式都比较有效，但是如果要将自己整理好的内容分享给其他人时，可选方式就比较有限了，特别是当要批量分享某个类目时。

除了分享的场景，很多时候还 **需要大众的力量进行资源的完善和维护**。完善意味着挖掘新的资源添进来，维护意味着修正一些不可用的资源。这个场景下最有效的方式是通过 Wiki 进行协作编辑。

Wiki 方式主要站点是 [WikiPedia]，但因为维基百科是一个综合化的知识库，并且对于主题侧重的权威性，所以一些我们觉得有帮助但比较小众的资源链接是整合不到其中的。

目前流行的一种方式是通过 [GitHub] 进行知识资源整合，比如一些关注数非常高的资源导航库（Awesome/资源合集等）。通常的协作流程是通过 Pull Request 发起，库 Owner 同意合并后就能够将 PR 发起者的变更（比如加入一个有价值的链接）同步到主库。这个协作方式略比 Wiki 复杂一些，但在程序员圈子是非常受欢迎的。

现状分析

通过观察 GitHub 上面很多实例我们发现，对于知识类型的资源共享/协同活动的共同点是：**少数人理，多数人关注**，很多 1K+ stars 的库其实贡献者很少。

以开源项目的观点来看的话，如果贡献者太少，那项目的活跃度主要依赖于 Owner，这样的开源项是很难长远的。结合前面我们讨论的资源整理型项目来看，其他人的参与成本是非常低的（比如只需提交链接和一段短小的说明），但是还是鲜有人参与，为什么呢？

原因至少应该有如下几点：

1. PR 流程比较复杂，对于贡献者和 Owner 都存在 **操作门槛**
2. 不能批量化进行变更，手工编辑过于 **繁琐**
3. 大部分人喜欢占有（star/fork），对于 **细节改进并不在意**

另辟蹊径

在 GitHub 上进行资源整合和分享并不是最高效的方式，而 Wiki 也不太适合解决这个需求，我们得明个轮子了——**链接熔炉**：

1. 从浏览器书签导入资源链接：书签都是经过个人整理的，是 **具有实用价值的资源链接**，并且导方式解决了手工、批量等操作繁琐的问题，当然，也支持提交单个链接
2. 自动整理：根据大家提交的链接进行计算整理，按标签进行分类、**链接价值** 进行排序，最终以个单页面（比如程序员领域）展现给浏览者，该页面汇聚了一些标签以及每个标签最具价值的资源链接

技术关键点

- 书签格式：以 Chrome 书签为例，导出后是一个 Netscape Bookmark 格式的 HTML 文件

- 链接页面解析：按爬虫方式访问目标链接，跳过不可用（非 200）的链接，解析页面
- 提取标签：按页面内容进行标签提取（接入第三方服务，例如玻森），并将标签结果和已有的标签做匹配，抽取最适合的标签
- 链接评分：按搜索引擎结果（链接引用数）进行排序，引用越多说明越有价值（简化的 PageRank，还需要探索更精准的算法

结论

链接熔炉是一个自动化的链接整理系统，一旦加入了新的链接就会再计算以产生更具实时性的结果，以后你就只用关注这一个页面了，里面有所有分好类的高质量资源

mile