# 各采集器对比



火车头在业内发展时间最长，但是不适用于没有编程基础的小白用户。其他几款采集器比如八爪鱼、神箭手，可以提供智能采集模式和官方制定好的规则，降低了采集难度。

后羿优于爬山虎：后羿不用注册就可以试用，且没有采集、导出限制；

爬山虎免费版只能导出为TXT。

八爪鱼由于集搜客：虽然集搜客不能批量提取已知网址，但八爪鱼经常提取失败。

火车优于其余：火车是老字号了，而且界面也好看一点点。

[后羿](http://www.houyicaiji.com/)：跨平台，Windows、 Mac、 Linux



智能识别数据，小白神器

智能模式：基于人工智能算法，只需输入网址就能智能识别列表数据、表格数据和分页按钮，不需要配置任何采集规则，一键采集。

自动识别：列表、表格、链接、图片、价格、电话号码、邮箱等

可视化点击，简单上手

流程图模式：只需根据软件提示在页面中进行点击操作，完全符合人为浏览网页的思维方式，简单几步即可生成复杂的采集规则，结合智能识别算法，任何网页的数据都能轻松采集。

可模拟操作： 输入文本、点击、移动鼠标、下拉框、滚动页面、等待加载、循环操作和判断条件等。

支持多种数据导出方式

采集结果可以导出到本地，支持TXT、EXCEL、CSV和HTML文件格式，也可以直接发布到数据库(MySQL、MongoDB、SQL Server、PostgreSQL)和网站(Wordpress、Discuz等)，提供丰富的发布插件供您使用。

功能强大，提供企业级服务

后羿采集器提供丰富的采集功能，无论是采集稳定性或是采集效率，都能够满足个人、团队和企业级采集需求。

丰富的功能： 定时采集，智能防屏蔽，自动导出，文件下载，加速引擎，按组启动和导出，Webhook，RESTful API，智能识别SKU和电商大图等。

云端账号，方便快捷

创建后羿采集器账号并登录，您的所有采集任务都将自动同步保存到后羿的云端服务器，无需担心采集任务丢失。 后羿采集器对账号没有终端绑定限制，您切换终端时采集任务也会同步更新，任务管理方便快捷。

全平台支持，无缝切换

同时支持Windows、Mac和Linux全操作系统的采集软件，各平台版本完全相同，无缝切换。

[神箭手云采集](http://www.shenjianshou.cn/)：跨平台



**简介：**神箭手云是一个大数据应用开发平台，为开发者提供成套的数据采集、数据分析和机器学习开发工具，为企业提供专业化的数据抓取、数据实时监控和数据分析服务。功能强大，涉及云爬虫、API、机器学习、数据清洗、数据出售、数据订制和私有化部署等。

**优点：**

* 纯云端运行，跨系统操作无压力，隐私保护，可隐藏用户IP。
* 提供云爬虫市场，零基础使用者可直接调用开发好的爬虫，开发者基于官方的云端开发环境开发并上传出售自己的爬虫程序；
* 领先的反爬技术，例如直接接入代理IP和自动登录验证码识别等，全程自动化无需人工参与；
* 丰富的发布接口，采集结果以丰富表格化形式展现；

**八爪鱼采集器**：

市场占有率大

一款网页采集软件，可以从不同的网站获取规范化数据，帮助客户实现数据自动化采集，编辑，规范化，从而降低成本，提高效率。

可应对各种网页的复杂结构（瀑布流等）和防采集措施（登录、验证码、封IP），实现百分之九十九的网页数据抓取。

特点：容易上手，模拟人浏览网页的操作，通过输入文字、点击元素、选择操作项等一些简单操作，即可完成规则配置，无需编写代码，完全可视化图形操作；

内置可扩展的OCR接口，支持解析图片中的文字；

采集任务自动运行，可以按照指定的周期自动采集。

火车头采集器：

一款互联网数据抓取、处理、分析，挖掘软件，可以抓取网页上散乱分布的数据信息，并通过一系列的分析处理，准确挖掘出所需数据。

特点：采集不限网页，不限内容；

分布式采集系统，提高效率；

支持PHP和C#插件扩展，方便修改处理数据。

狂人采集器：

一套专业的网站内容采集软件，支持各类论坛的帖子和回复采集，网站和博客文章内容抓取，分论坛采集器、CMS采集器和博客采集器三类。

特点：支持对文章内容中的文字、链接批量替换和过滤；

可以同时向网站或论坛的多个版块一起批量发贴；

具备采集或发帖任务完成后自动关机功能；

三人行采集器：

一套可以把别人网站、论坛、博客的图文内容轻松采集到自己的网站、论坛和博客的站长工具，包括论坛注册王、采集发帖王和采集搬家王三类软件。

特点：以采集需要注册登陆后才能查看的论坛帖子；

可以同时向论坛的多个版块一起批量发贴；

支持对文章内容中的文字、链接批量替换和过滤。

集搜客：

一款简单易用的网页信息抓取软件,能够抓取网页文字、图表、超链接等多种网页元素，提供好用的网页抓取软件、数据挖掘攻略、行业资讯和前沿科技等。

特点： 可以抓取手机网站上的数据；

支持抓取在指数图表上悬浮显示的数据；

会员互助抓取，提升采集效率。

