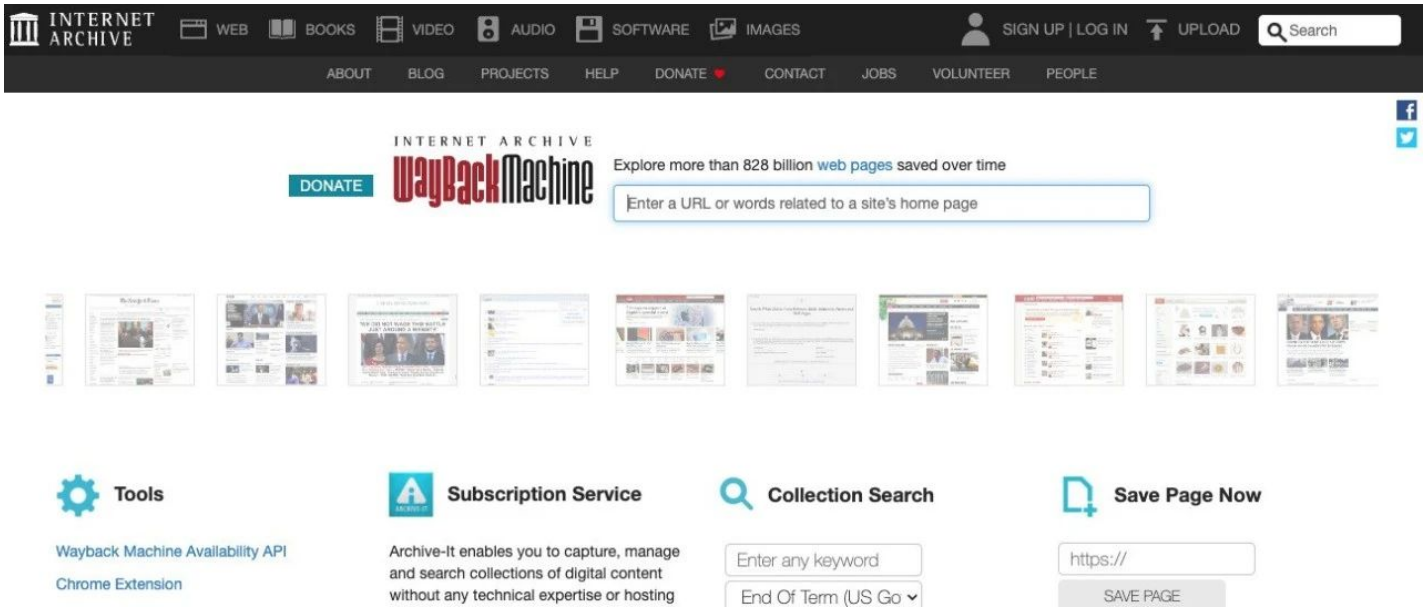
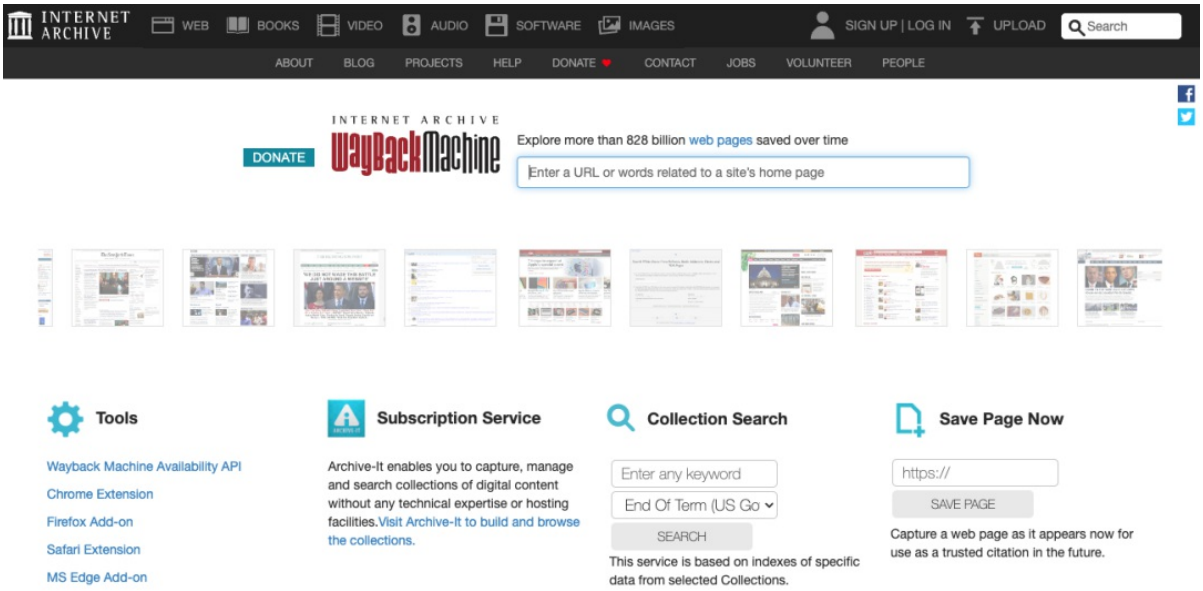



如何浏览一个网站 10 年前的样子?

原创 我是哥飞 哥飞




哥飞今天在V2EX看到一个帖子，问如何浏览一个网站 10 年前的样子，下面几个V友都推荐了 archive 的 Wayback Machine。

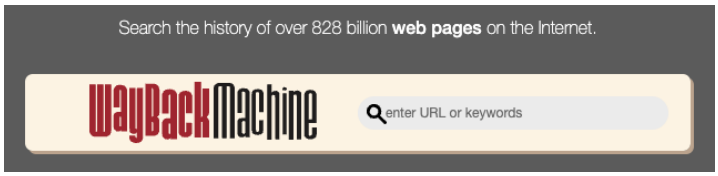



其实哥飞在之前的文章中多次提到了 archive：

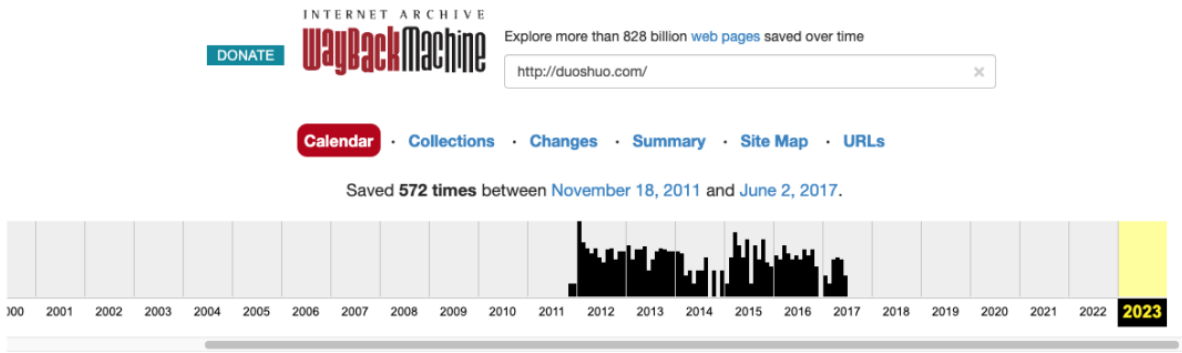
> 哥飞为了写这篇文章，特意又从 archive.org 上把当时网站界面截图下来了给大家看。
>
> 我是哥飞，公众号：哥飞
> 我如何通过SEO和Adsense优化，把一个网站广告收入从每月八百多美元提升到每月两千多美元，经验全分享
> 我查了一下，这个域名居然是7月17日注册的，再查 archive 发现网页是7月22日被收录的，也就是说，建站到今天，也才20天。
>
> 我是哥飞，公众号：哥飞
> 【5000字调查分析】建站20天拿下480万访问量，俄罗斯版的妙鸭相机是怎么做到的?
> 再查 Archive.org 获取收录的历史页面
>
> 我是哥飞，公众号：哥飞
> 人人都能学会的挖掘 web 产品需求之从出站域名发现新需求新产品

还有更多相关文章，大家可以关注哥飞后，搜索查看。
今天，哥飞就给大家正式介绍一下这个工具，以及用法。
记住，网址是 https://web.archive.org/ ，当然用 https://archive.org/ 也可以打开。从不同的入口进去，搜索框界面不太一样，但是功能是相同的。
注意搜索框上面这句话，archive 已经收录了8280亿个网页了，不愧是互联网档案馆。



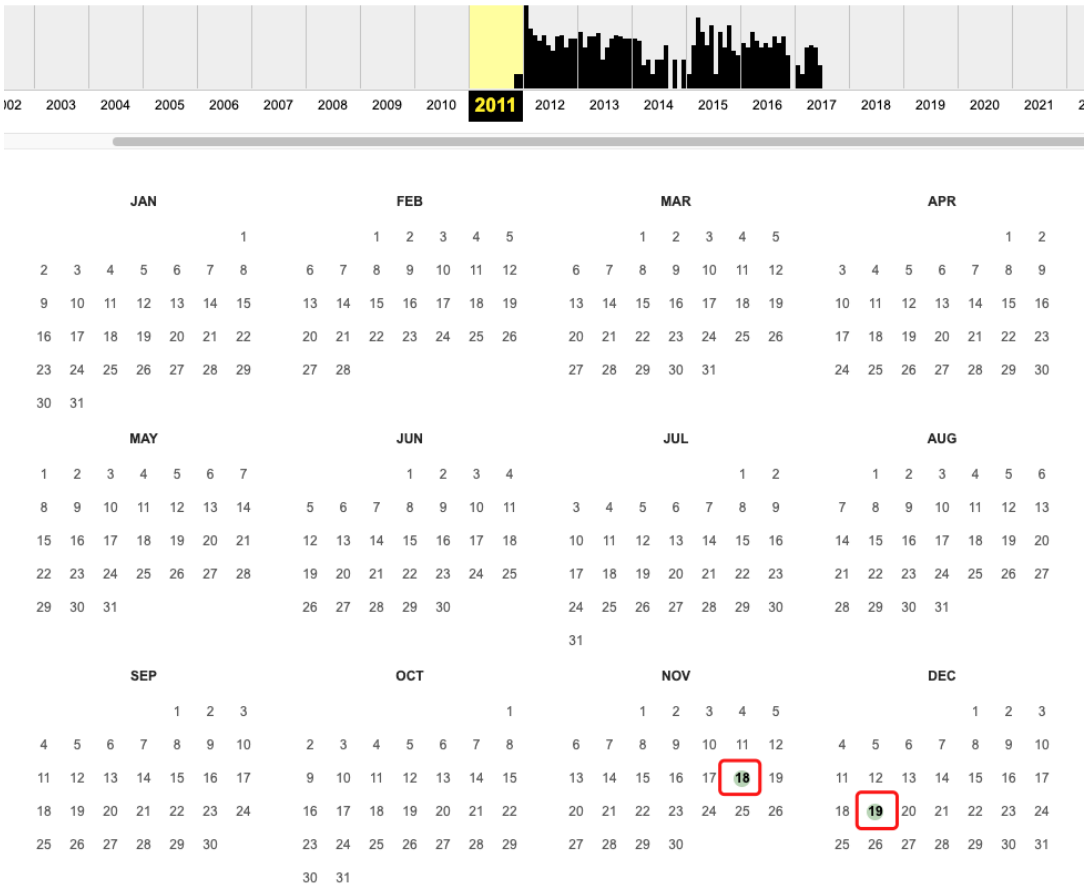




用法也很简单，你可以在输入框中粘贴任意网址进去，点击按钮或者敲下回车键，就可以开启时光机，回到过去，查看这个网页的任意时刻。
刚好今天郭宇讲到说他之前被收购进字节时所在公司是多说，那我们就用多说来演示一下用法。
你现在已经无法打开 http://duoshuo.com/ 了，但是通过 archive ，我们还是可以找到这个网站曾经的模样。
输入网址并提交搜索后，打开了下面这个界面，这是 archive 最经典的界面之一，上半部分是导航、数据汇总和时间轴。
我们只关注导航中的"Calendar"就好了，其它的大家可以自己点开看一下，哥飞平常不使用剩下的这些。

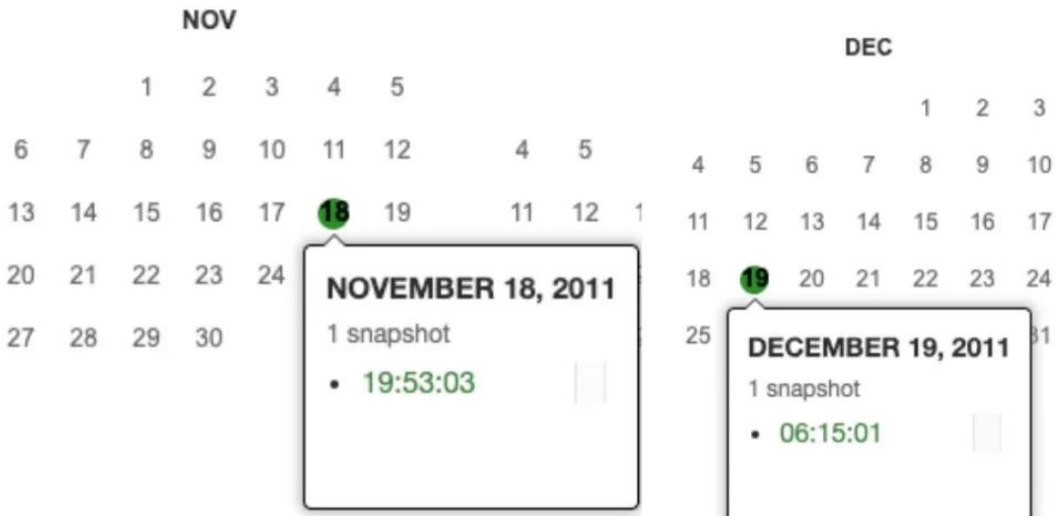



我们主要看时间轴，可以看到这个网站最早是2011年开始收录的，然后到2017年中之后就没有收录了。

很多时候，我们可以几乎把开始收录时间当作网站上线时间。
现在，我们点击2011年，会看到，这一年，只收录了2次，一次是11月18日，一次是12月19日。




把鼠标放在有颜色的日期上面，或者点击也行，会弹出一个小窗口，显示这一天收录的所有次数，有时候一天会收录多次。



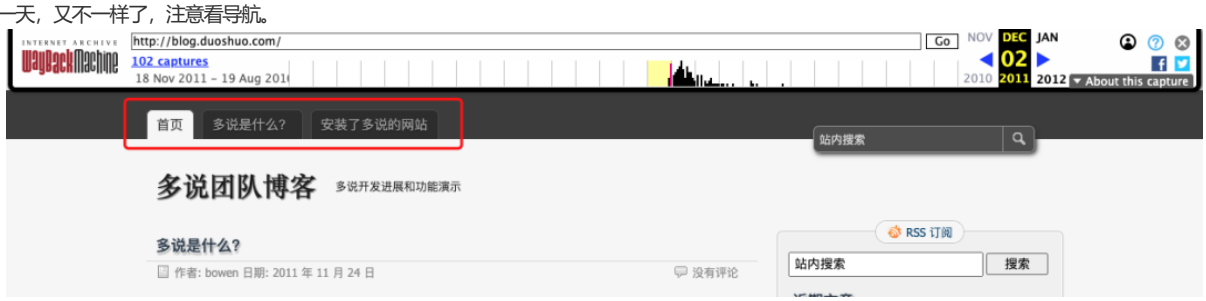

点击时间，此时 archive 就会带我们回到那时那刻，这里我们点击11月18日的19:53:03，在新打开的页面，的确可以看到那时的多说。
哥飞强烈建议大家不要直接点时间，而是右键选择在浏览器 新标签页打 开，至于为什么，你实际跟着哥飞的操作试一下就立马能知道。




可以看到，实际上多说第一版是用Wordpress博客作为官网首页的，第一篇博文发布于11月16日，但是2天之后，也就是11月18日才被 archive 收录。
看完了这一天，我们可以点击页面右上角的三角形▶□按钮，看下一天。
每次都有新变化，时隔一个月，多说已经发布0.33版本了，上架到wordpress官方插件商店了。




再点到下一天，又不一样了，注意看导航。



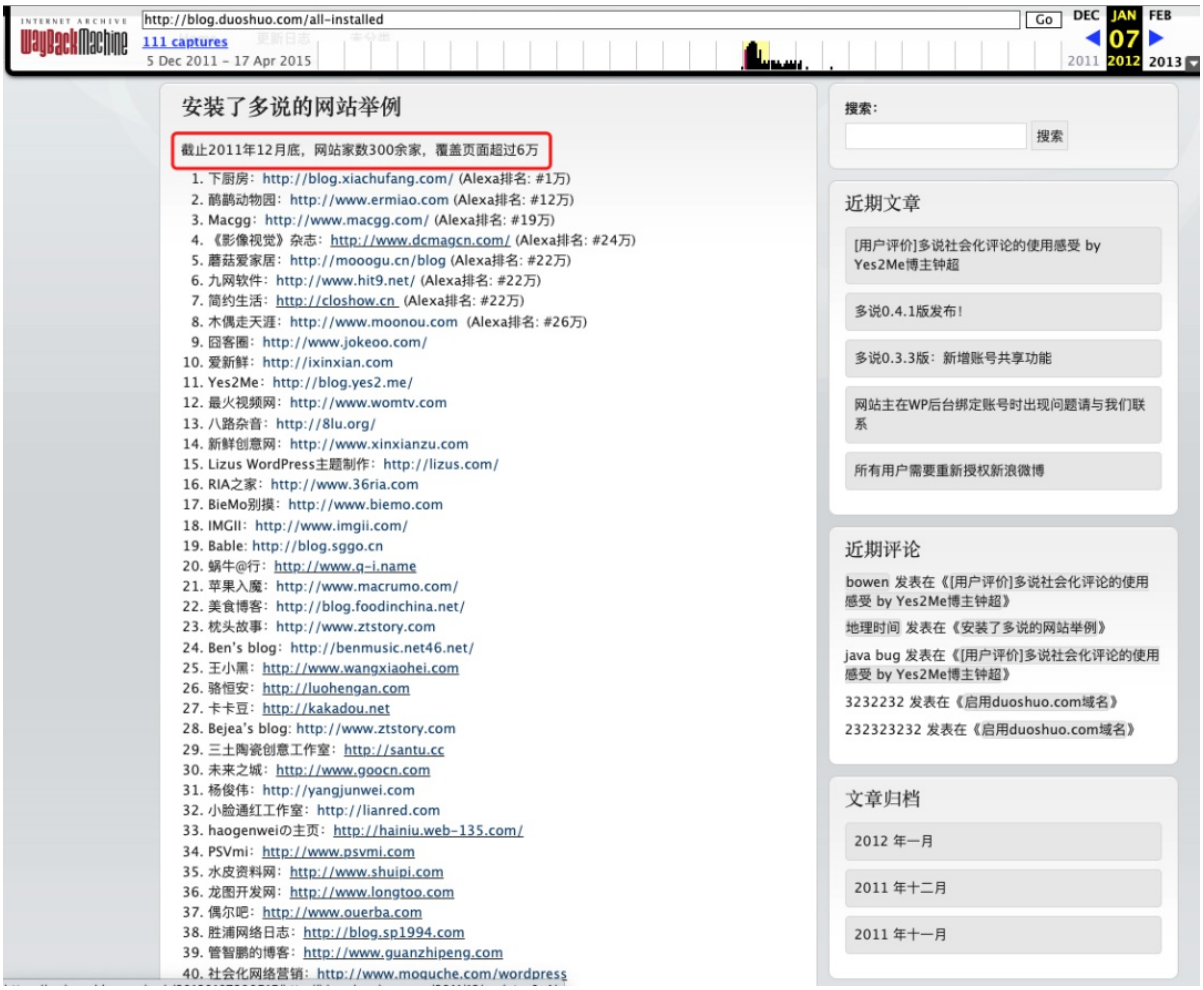

这里哥飞也跟大家说下，每一天的内页，其实也可以点击的，不是只能看首页。如我们可以点击“安装了多说的网站”，可以看到那时候，只有3个网站安装了多说。



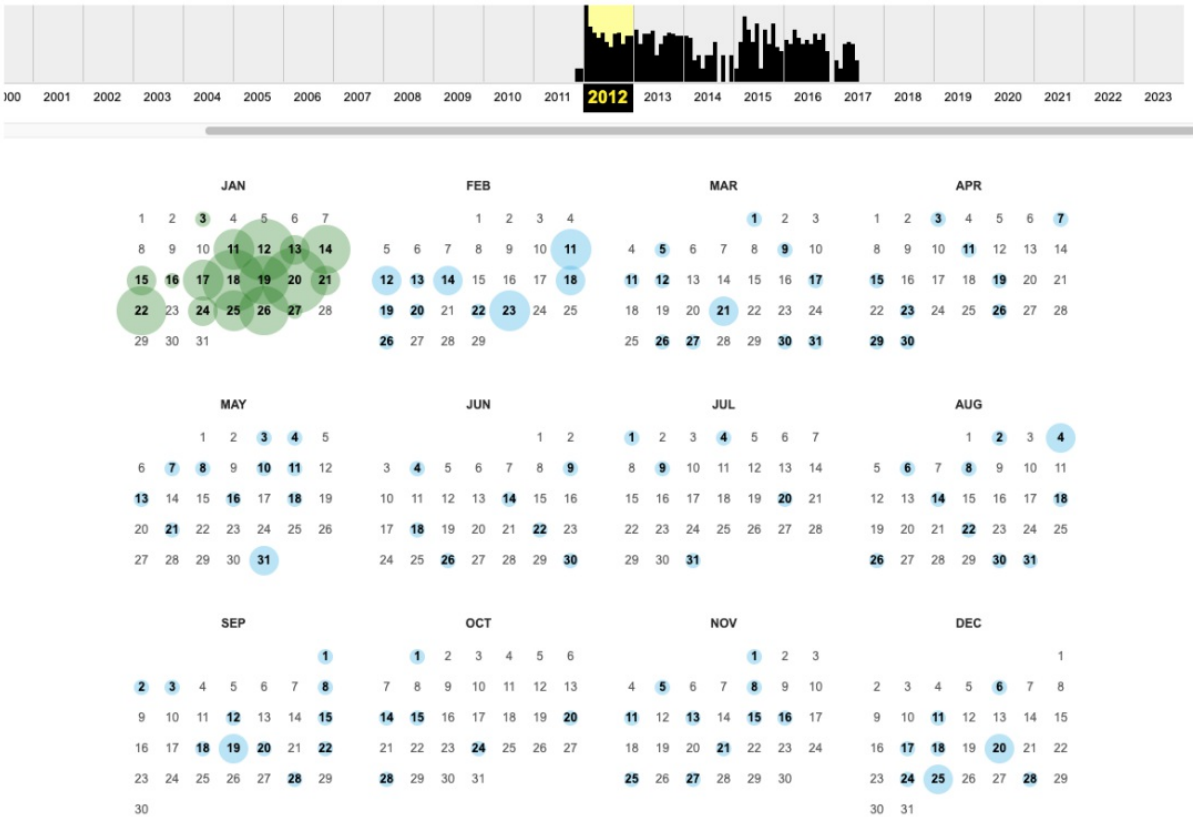

我们甚至可以点击外链，如点击爱新鲜链接打开页面如下。




继续往后面看，2012年1月7日收录的这个页面显示，截止到2011年11月，已经有三百多家网站接入多说了。 这个发展速度还是不错的。




我们回到 archive 的Calendar界面，看下2012年的收录就更频繁了。同一天，收录次数越多，圆圈直径越大，颜色也会不同，根据这个就很容易判断哪些天收录多。




archive 的基础用法就是这些了，你可以放进任何网页地址进去体验一下时光机。甚至被删掉的网页，说不定也可以在这里看到。

那么除了这样看，archive 这个网站的数据，对于我们还有什么作用呢?

其实 arvhive 是有 api 的，通过 api ，我们传入任意一个域名进去，都能找到这个域名的被收录信息，我们就可以用来判断网站的上线时间。

你可能会说，判断上线时间不能用域名注册日期吗?

哥飞告诉你，有些网站注册了域名后并没有及时上线，或者是从别人手里买过来的域名，换了网站。

如前几天哥飞这篇文章介绍的一个食谱网站，域名注册时间和上线时间中间就隔了8年。

> tasty.co 每个月720万访问量，域名是2010年注册的，但实际网站从2018年开始做的

> 我是哥飞，公众号：哥飞
> 出海不知道做什么网站？来看看别人的成功案例

昨天分享的这个网站也是一样，域名注册于2008年，但是2020年才开始做千篇国学网站。

> 这个域名虽然注册于2008年，但实际上千篇国学网站，是2020年5月份才开始做的。
>
> 我是哥飞，公众号：哥飞
> 分享一个谷歌和百度都优化得不错的网站

有了真实的上线时间，我们就可以更好的定位到运营团队的一些推广节点。

如哥飞前几天的这篇调查分析报告，里边就用了 archive 来判断网站上线时间。

> 我查了一下，这个域名居然是7月17日注册的，再查 archive 发现网页是7月22日被收录的，也就是说，建站到今天，也才20天。
>
> 我是哥飞，公众号：哥飞
> 【5000字调查分析】建站20天拿下480万访问量，俄罗斯版的妙鸭相机是怎么做到的？

刚才哥飞说了， 其实 arvhive 是有 api 的，那么具体 api 地址是什么呢？请加哥飞微信 qiayue 领取。

