



[Home](#) > [Tutorials](#) > How To Build Massive Site Lists For GSA SER

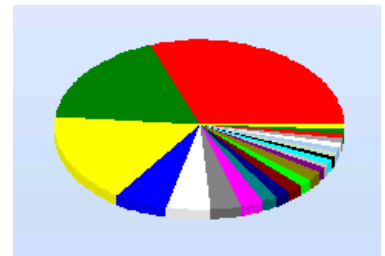
How To Build Massive Site Lists For GSA SER

Neil November 25, 2013 Tutorials 16 Comments



If you have read my [GSA SER Review](#), then you might have been wondering how you can create massive lists to run through GSA to get thousands of links per minute. In this tutorial I am going to show you how you can take an original seed list to expand it into thousands of sites.

This will allow you get a ton of verified links without having to scrape for them yourself (which can be slow) and get rid of buying any lists.



Download A Printable PDF of This Guide

Just share this post to download a printable PDF of this guide to keep on your computer!



WHAT YOU WILL LEARN

- Getting a seed list
- Extracting the internal links from the seed list
- Extracting external links
- Sorting your list

RESOURCES NEEDED

- [GSA SER \(my review\)](#)
- [Scrapebox](#)
- [GSA Captcha Breaker](#) or [Captcha Sniper](#)

THE MAIN IDEA

If you are using GSA SER to post to different blogs, then I am assuming that other users of the software are probably posting to these sites as well. Some GSA users have absolutely massive campaigns for their tiered projects and are blasting hundreds of thousands of links.

What we are going to do is scrape all the internal links of the blogs that you are posting to. This will give you a big list of different posts that have been destroyed by GSA SER.

We will then extract all the external links from the blog post list we have gotten, giving us a list of targets for GSA that other people are using.

GETTING YOUR SEED LIST

The first thing you are going to need to do is find a list. If this is your first time then go to the verified folder of GSA SER and grab all the URLs from all the blog comments you have made. It does not really matter if you have a little list, you will see that even a small list can be expanded to thousands of URLs.

After the first time you do this, you can then just use verified comments from the last project that you successfully used this technique with. If you want a complete site list, then you can grab my starter list below. This is not just blog comments though, it is all platforms.



Please share to get access to my 19k verified domain URL list to use with GSA SER



When you finally have the list, import that list to Scrapebox, trim to the root, get rid of duplicate URLs, and then save the list as "seed list" and move on to the next step.

EXTRACTING THE INTERNAL LINKS

Open up the plugin in Scrapebox called "Link Extractor" and load the seed list you just created into it. Set the mode for the plugin to internal and use as many connections as possible. Start it up and let it run.

Once the process is complete, import the list you just ran through to Scrapebox and close the Link Extractor plugin.

Take a look through the list and check out if there are any categories, comment (#comment), tag links or anything like that. You can then use the "Remove URLs containing" option to get rid of these types of links.

You are looking for just a list that is made up of blog posts. This would be ideal. Save the list as "internal links" and go on to the next step.

Optional: Set a project in GSA just to feed in these lists. This will filter out the trash from the blog posts you are looking for. This also can be a great way to build an auto accept list.

EXTRACTING THE EXTERNAL LINKS

I should note that in this step, your original list you use for this is going to be expanded to hundreds of thousands or millions of URLs. This means that you should probably split up your internal links list into smaller files of no more than 10k links per list. This step will usually create 200 times more URLs than you start with and Scrapebox does not seem to link lists more than 1 million URLs as much as I wish it did.

Tip: Use the "Dup Remover" plugin within Scrapebox to split the files.

Okay, so now that you have the internal links list split into smaller files, open up the link extractor plugin and let's get going.

Load in the internal links file, set your mode to "external" and click start.

This may take a while, so enjoy a break while the software does its job. When it is done, transfer the list to ScrapeBox and deduplicate it again if you need to.

Save the list as something such as: "Needs Sorted".

Do this over again until all your smaller link lists from your internal links lists are done.

SORT YOUR LIST

You should have some pretty big lists on your hands. They need to be sorted! Good news for you, GSA SER can do this!

You will want to setup a new project in GSA SER just like this one. Copy the settings of this picture!

The screenshot shows the 'Project Data - SITECHECK 1' window in GSA SER. It is divided into two main sections: '[Where to Submit]' on the left and '[Settings]' on the right.

[Where to Submit]

- ☒ Article
- ☒ Blog Comment
- ☒ Directory
- ☒ Document Sharing
- ☐ Exploit
- ☒ Forum
- ☒ Guestbook
- ☒ Image Comment
- ☐ Indexer
- ☒ Microblog
- ☐ Pingback
- ☐ Referrer
- ☒ Social Bookmark
- ☒ Social Network
- ☒ Trackback
- ☒ URL Shortener
- ☒ Video
- ☒ Video-Adult
- ☒ Web 2.0
- ☒ Wiki

[Settings]

Buttons: Data, Options (selected), E-Mail verification, Notes

[-] How to submit / verify

- ☐ Pause the project after 150 submissions reached in a day ☐ per URL
- ☒ Ask 1st service to fill captchas ☐ Custom Retry 5 times
- If a form field can't be filled: ☒ Skip hard to solve captchas
- ☐ Enable Custom Mode
- ☒ Verified Links must have exact URL
- When to verify: Custom time 1440 minutes ☒ Don't remove URLs (check hint) ☐ Remove after 1st verification try
- ☐ Also re-verify existing backlinks
- ☐ Send verified Link to Indexer Services ☐ GSA SEO Indexer ☒ Other Indexer(s)
- ☐ Use character spinning where possible (fighting duplicate content)
- ☐ Randomly use typos/spell errors in descriptions (fight duplicate content) 10 %
- ☐ Try to always place an URL with an anchor text in descriptions/comments
- ☐ Continuously try to post to a site even if failed before
- For TAGs use: ☒ Anchor texts ☐ Keywords

[-] How to get Target URLs

Search Engines to use

Name
<input type="checkbox"/> 360EÑE+
<input type="checkbox"/> Acoon
<input type="checkbox"/> Ananzi Search Engine
<input type="checkbox"/> AOL
<input type="checkbox"/> AOL CA
<input type="checkbox"/> Arianna
<input type="checkbox"/> Ask.com

0/1002 Checked

- ☐ Always use keywords to find target sites
- ☐ Use URLs from global site lists if enabled ☐ Identified ☐ Submitted ☐ Verified ☐ Failed
- ☐ Also analyse and post to competitor backlinks

[-] Filter URLs

☐ Skip sites with more than 100 outgoing links on one page

☒ Skip sites with a PR below 0 ☐ And with a PR above 9

☒ Skip also unknown PR use PR of SubDomain

Type of Backlinks to create

☒ Anchor text ☐ Try to skip creating Nofollow links

☒ Article ☐ Avoid posting URL on same domain twice

☒ Article-Wiki ☐ Avoid posting to subdomain sites

☒ Bookmark-Anchor text ☒ Avoid posting to sites with just IP and no domain

☒ Comment-Anchor text ☐ At least 1 of my keywords must be present on the site

☒ Comment-Anchor text - Domain

☒ Comment-Contextual

☒ Comment-Contextual/Anchor text

☒ Comment-URL

☒ Directory-Anchor text

☒ Directory-Contextual

☒ Skip sites where the following words appear

☒ sex

☒ cunt

☒ dick

☒ boobs

☒ boob

☒ Skip sites with the following words in URL/Domain

☒ sex

☒ cunt

☒ dick

☒ boobs

☒ boob

☐ Skip sites from the following countries (detection could fail)

Country

☐ Afghanistan

☐ Albania

☐ Algeria

☐ American Samoa

** With these settings, all unindexed sites will be filtered out. If you want to get as many links as possible and do not care about the PR of the pages, then untick the box for PR filters and it will make your project faster and skip the PR check step.

** Also untick all the search engines and the site lists. What you want is this project to only post to the target URL list that you have just imported.

Be sure that you are using one of the main captcha solving software. If you do not use a software you are going to be paying a LOT of money, so pick up one of these two softwares: [GSA Captcha Solver](#) or [Captcha Sniper](#).

Next, get some spun content that is relevant to your niche. I prefer to use [Kontent Machine](#).

Once you have everything set up, import your "Needs sorted" list as the target URLs for the project.

BAM! The project will then go through the list and you will see all the successful and verified links!

Tip: Make sure to save the identified, verified and Submitted sites in GSA SER!

You can now repeat the process on the already massive list you just ran through GSA SER to create and even BIGGER list!

QUICK RECAP

- Grab a list of blog comments that have been verified on GSA
- Extract all the internal links
- Extract all the external links (from your internal links list)
- Import the list to a GSA project and let it sort it out for you
- Rinse and repeat with the verified links you just made!

Enjoy building your free list of sites that you can use yourself or sell to others. Go get those rankings!

Related Posts



[Buying Deleted/Expired High PR Domains for their Backlinks](#)



[How To Host A Private Blog Network For Free](#)



[How to Improve WordPress Security](#)

About The Author



Neil

has been in the internet marketing game for 3 years now. I am here to help you make money online based on my own experiences and experiments. There will be no fluff or BS, just what is working now.

16 Comments



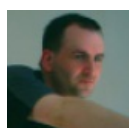
Nick JANUARY 15, 2014

Hi Neil, thanks for taking the time to put this together. I have tried a few methods to scrape site lists for GSA but this guide is yielding some massive lists! Once its done I'll how many verifieds I can get.

Cheers

Nick

[Reply](#)



FC-DH FEBRUARY 21, 2014

Are you the same guy that posted this on BHW?

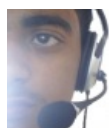
[Reply](#)



Neil MARCH 1, 2014

No, but we are good friends =)

[Reply](#)



Irfan Saleem MAY 23, 2014

That's actually a good guide :). Me too wrote a complete newbie guide to GSA SER and this is something that deserves a OBL from me 😊

[Reply](#)



Neil MAY 26, 2014

Awesome write up! I have been meaning to write up a set up tutorial but have not gotten around to it. You did a great job, keep it up XD

[Reply](#)



iOnGeek MAY 25, 2014

Hi, this is a very informative post.

I followed your process and was able to harvest about 6 million targets. That is the biggest harvest I have been able to achieve since I started using scrapebox for GSA.

Unfortunately got just about 20K GSA verified from the 6 million. Do you have the same experience? Just want to be sure I am not doing something wrong. Thanks.

[Reply](#)





Neil MAY 26, 2014

I am glad you have had some success from this! In my experience, it is about the same. Make sure all your proxies are working and you have all platforms checked in GSA. Keep on harvesting!

[Reply](#)



mya JUNE 18, 2014

Here are two questions:

1. Don't the seed list have to be unique towards my keyword? Therefore, I would have to first scrape my own Seed List even before I pop them into Link Extractor?
2. What do you mean, "Set a project in GSA just to feed in these lists."?
3. What if I didn't dupe my list? It's a 25mb text file, but can I run this enormous link through SB install of breaking it apart and running each of the smaller files?
4. When repeating the whole process again, you mean to use the newly generated extraction list?

[Reply](#)



Neil JUNE 21, 2014

This technique I have mostly used for mass spam so my seed list was just a list of verified URLs for a previous spam campaign. If you want it to be geared towards your niche, then yes, you would need to scrape your own seed list of links that are relevant to your site before you put them into the link extractor.

What I meant by set up a project in GSA to feed in the lists is to create a project in GSA and then run the list through it. When GSA runs, it will not be able to post to any link that is a category page or a tag page so it will essentially get rid of those links for you when GSA has verified all the URLs.

For your third question, you do not need to dupe the list or split it apart. It just makes everything a lot easier to work with and to create bigger lists. Test it out yourself and see what is working best for you. When repeating it again, you are correct to use the newly generated extraction list to keep building much bigger lists. Cheers!

[Reply](#)



Tyler Bailey JULY 27, 2014

Great guide, I have been doing it a similar way except I use scrapebox's automator plugin to add a few steps/cut down on time a bit.

[Reply](#)



Vikram JULY 29, 2014

So with this method, you are essentially NOT using GSA's search functionality but instead you are exclusively using Scapebox's scraping functionality to create your lists?

[Reply](#)



Neil JULY 29, 2014

That is correct =)

[Reply](#)



Reven AUGUST 25, 2014

I like you article layout, everything is sorted out 😊 Easy to follow and read.

I have only 1 question regarding "Sort Your List with GSA" by doing this do you mean that we need to run a campaign before we actually get AA list or Approved list for future campaign?

Oh actually one more question. You must know how serlists(dot)com or other people who sale similar list.

What are there method to scrape links?

Thank You for your article...

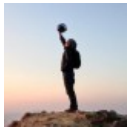
[Reply](#)



Neil AUGUST 26, 2014

By running the list through GSA, only the verified links will be your list so essentially GSA does the work of sorting out the bad links for you. As for serlists, I have never used them but I would assume that they are using something like gscraper running 24/7 on a VPS.

[Reply](#)



milan SEPTEMBER 5, 2014

So this will just build list for the "blog comments" right ? Can't we build site list for other platforms as well ? Like article directories, wikis etc ???

[Reply](#)



Neil SEPTEMBER 18, 2014

Not necessarily, you are extracting links from blog comments to get the lists of other people, and they may have done much more than just blog comments and GSA would be able to sort that out. Give it a try and see what all kinds of links you can get.

[Reply](#)

Leave a Reply

*Your email address will not be published. Required fields are marked **

[Post Comment](#)

☐ Notify me of followup comments via e-mail. You can also [subscribe](#) without commenting.

FOLLOW US



Search the site

Search

RECENT POSTS



[How To Host A Private Blog Network For Free](#)

September 3, 2014 • 3

[Comments](#)



[\(Time Sensitive\) How To Get Unlimited Free Website Content To Build Authority Sites](#)

August 9, 2014 • No Comment



[How To Make Money With Fiverr - \\$X,XXX Startup Money](#)

December 28, 2013 • 8

[Comments](#)

[\\$100 A Day Adsense Journey \(Part 1\)](#)

December 3, 2013 • 7 [Comments](#)



[How To Build Massive Site Lists For GSA SER](#)

November 25, 2013 • 16

[Comments](#)

CATEGORIES

- [Case Studies](#)
- [Internet Marketing Tricks](#)
- [Reviews](#)
- [Tips](#)
- [Tutorials](#)

