Mirroring a Censored WordPress Blog

By Sami Ben Gharbia with Rebekah Heacock & Jeremy Clarke

This guide is for bloggers with self-hosted WordPress blogs who believe their sites may be blocked by government filters. Its goal is to help bloggers use a mirror site to make censored content available to readers despite these filters.

only for WordPress.org blogs

Introduction

When it comes to accessing banned web sites and blogs, most online free speech advocates focus on bypassing censorship using circumvention tools. While this is important, teaching Internet users how to evade censorship is not enough. Many people are not aware of or do not have access to circumvention tools, as countries that filter the Internet also tend to block proxy servers and other circumvention technologies. When circumvention tools are used, they often affect connection speeds, making Internet access even slower in places where connectivity is already poor.

To make your content accessible you cannot simply rely on circumvention technologies or the technical knowledge of your readers.* Taking your users' needs into account and keeping ahead of the blocking efforts of their governments are the first steps toward implementing creative workarounds that make your site's content available to all potential readers, regardless of where they live.

One way to increase access to blocked sites for Internet users in countries that filter online content is through mirroring: duplicating a site's content on another domain name or subdomain. Mirror sites automatically reflect any changes made to the original site, allowing blog authors to get around censorship by providing multiple locations where readers can access their content.

This guide is for bloggers with self-hosted WordPress blogs. A self-hosted WordPress blog is one that is not hosted on the WordPress.com free blogging service, but rather on a separate server using the WordPress.org publishing platform. For more information, see http://wordpress.org/. who believe their sites may be blocked by government filters. Its goal is to help bloggers use a mirror site to make censored content available to readers despite these filters. It contains step-by-step instructions for setting up a mirror for an original ("source") WordPress blog.

Contents

* For a description of the strengths and weakness of circumvention tools, see: Ethan Zuckerman, "Internet Freedom: Beyond Circumvention", My Heart's in Accra, February 22, 2010, http://www.ethanzuckerman.com/blog/2010/02/22/internet-freedom-beyond-circumvention/

1- Duplicating your content without mirroring

One simple way to help blocked users access your blog is to duplicate its content on other sites. Though the URL or IP address of your site may be blocked, services that republish your content may be available. Tools like Google Reader, FriendFeed, Google Buzz, Facebook and other RSS readers give visitors more ways to access your site by republishing your content in multiple locations.

In order to use these services, you need to publish an RSS or Atom feed. WordPress automatically creates these feeds for your blog, but if you plan to use these tools to make blocked content accessible to readers, make sure to:

- Include all of your content in your feed, not just basic titles, headlines and excerpts. In WordPress this is controlled by a setting in the Settings > Reading page. Where it says "For each article in a feed, show" ensure that "Full text" is selected.

- Check that the media files used in your blog content are not blocked for your readers. If the images/videos/audio included or linked to from your posts are hosted on your domain, then readers may not be able to view them even if they can read your RSS feed. Instead of hosting your multimedia content on your blocked web site, try to utilize social media sites that are not blocked in your target country. Uploading your content to sites like YouTube, Flickr, blip.tv, or archive.org and linking to it from your blog may help make it available to visitors who would otherwise only be able to see your text.

You should make sure that the services you choose are not banned in your target country. The OpenNet Initiative has a map showing which countries currently block Facebook, Flickr and YouTube.

2- Making your blog secure

Before you start the process of mirroring your blog to evade censorship you should ensure that you are in fact being targeted by a government, rather than experiencing non-political security problems. WordPress, due to its popularity, is targeted by an army of creative and dedicated hacker-spammers who break into WordPress-powered sites so they can add spam links or use other illegal and abusive search engine optimization (SEO) tricks that exploit your site's popularity. One possible effect of these intrusions is that your site may be blocked by filtering software or even Google for including inappropriate content (put on your site by the hackers) or for serving malware/spyware to visitors, another common strategy for hackers who take over WordPress sites.

Ensuring that your site is secure and firmly in your control is a good idea no matter what, but if you think you are blocked then the process of investigating security may turn up new clues that point to a more mundane problem. If you discover that your site has in fact been compromised and you are able to fix it, then you can get your site reinstated in Google or removed from commercial filters by requesting that your site be reviewed (more on this below).

By the same token, if you are worried about governments hoping to censor you, then having a particularly secure site is a good idea either way, as it will protect your site and personal information from politically malicious hacking attempts.

The WordPress Codex offers advice about what do to if you think your site has been hacked and how to harden WordPress to make it more secure. The fundamental goals are to ensure that no users exist on your site that you don't know about and that all the files on your server are the ones you expect. Hackers will upload new files and use them to re-hack you if they are locked out. They will also inject content into your database and hide users where you can't find them.

The easiest way to audit your site is to use some of the many security WordPress plugins that exist for the job. Here is a short list of some of the best ones (you don't need to install them all, but considering what they offer is useful):

Some trusted security WordPress plugins

- ★ WP Security Scan: Scans your WordPress installation for security vulnerabilities and suggests corrective actions.
- ★ WordPress Firewall Plugin: Investigates web requests with simple WordPress-specific heuristics to identify and stop most obvious attacks.
- ★ WordPress File Monitor: Monitors your WordPress installation for added/deleted/changed files. When a change is detected an email alert can be sent to a specified address.
- ★ WordPress Database Backup: Creates backups of your database, including automatic regular ones. This is particularly important because if you are hacked you need a backed up database that you know has not had malicious content added into it.
- ★ Secure WordPress: Helps to secure your WordPress installation: removes error information on login page; adds index.html to plugin directory; removes the wp-version, except in admin area.
- ★ Maximum Security for WordPress: Guards against intrusion; tracks a plethora of events; blocks malicious content that could harm your readers and your search engine ranking; and includes a strong Web application firewall along with a full blown intrusion prevention system.
- ★ Login LockDown WordPress Security: Records the IP address and timestamp of every failed WordPress login attempt. If more than a certain number of attempts are detected within a short period of time from the same IP address, then the login function is disabled for all requests from that range.
- ★ ChapSecureLogin: You can use this plugin to process your password encryption. The encryption process is created by the Chap protocol; this is particularly useful when you can't use SSL or any other kinds of secure protocols.
- ★ Theme Authenticity Checker: TAC searches the source files of every installed theme for signs of malicious code. If any bad code is found.

3- Introduction to Internet filtering techniques

The OpenNet Initiative, a research center based at Harvard University, defines Internet filtering as "the technical approaches to control access to information on the Internet."* Internet filtering is only one part of Internet censorship, which also includes tactics such as threatening online speech and removing offensive web sites from the Internet via take down notices or domain name deregistration. For the purposes of this guide, we will focus on the various technical methods of Internet filtering, which can be divided into four main approaches:

DNS Filtering

The most common way to censor a web site is by blocking access to its domain name for an entire region, for example by disallowing all traffic to yourblog.com. In such cases blocked web sites are often accessible at other domain names or subdomains of such a site: though yourblog.com is blocked, blog.yourblog.com might continue to be accessible.

Uniform Resource Locator (URL) Filtering

Another common method is to block access to specific information (pages or posts) on a web site or a blog by preventing access to particular URLs. This selective blocking only targets specific subdomains or pages without affecting the rest of the web site. For example, a censor might filter the advocacy.globalvoicesonline.org subdomain while leaving the more general globalvoicesonline.org unfiltered, or vice versa.

IP Filtering

Blocking the IP addresses of servers hosting unwanted web sites is the most straightforward form of online censorship. The IP address is the number used to uniquely identify every computer on the Internet, so blocking the IP address of a given machine makes it inaccessible. Although IP filtering is the easiest and cheapest way to ban unwanted content, it can easily lead to over-blocking a large range of web sites. If an IP address associated with a particular web site is blocked, all other web sites that share the same IP address on the web server become blocked as well.

Keyword Filtering

Some countries, such as China and Tunisia, block access to any URL path containing a specific keyword. For example, Tunisia blocks the domain nawaat.org as well as the keyword "nawaat" in all URL paths. This means that @nawaat Twitter account is automatically blocked, as is the Nawaat Facebook account and all Google cache and search result pages that contain "nawaat" in their URL paths.

---

* OpenNet Initiative, "About Filtering," http://opennet.net/about-filtering.

4- Determining if and how your site is blocked

If you can identify what techniques are being being used to block your blog, you can determine the best way to make it available again for blocked users. A good place to start is to check the OpenNet Initiative's country profiles to see which filtering methods are used in the place where you suspect your blog is being blocked.

Assessing whether your site is blocked is an important part of this process. It's possible that setup, security or even connectivity problems, rather than censorship, are causing some readers to find the site inaccessible. To fully determine the nature of any filtering or blocking of your site you will need to communicate with affected individuals and ask them to help you test the situation. Making connections with a group of testers is likely to prove valuable. If you want to crowd-source the testing of whether your site is blocked in different locations you may want to check out Herdict, a site that lets people state their location and whether they can access your site, creating a map of where you are likely blocked.

Censorware

Some countries use filtering software (censorware), such as WebSense or SmartFilter, that block access to sites based on their categorization as "harmful," "gambling," "spam," etc. If you think your site might have been blocked due to being flagged as harmful or spam (see security section above), you can use Google's Safe Browsing diagnostic tool (visit http://www.google.com/safebrowsing/diagnostic?site=[your URL, such as http://globalvoicesonline.org]) or McAfee's SiteAdvisor to check its status. Such blocking can result from an insecure site that has become a source of malware due to being hacked. If you have fixed the security problem or your blog has been flagged in error, you can contact Google to request re-consideration of your site.

IP Blocking

The first step to determining if your IP address is being blocked is to figure out what your IP address is. You can use the IP Lookup: Domain tool to find out which IP address corresponds with your blog's domain name. The next step is to check to see if this IP address is blocked.

It's possible that your site's IP address is blocked not because of your content, but because another site with the same IP on the same shared host was blocked. To see what sites share a particular IP address, you can perform a reverse IP lookup. If you suspect your site is being blocked accidentally because of its shared IP you can contact your hosting service and ask them to change the IP of your site somehow (e.g. by migrating it to another server or cluster). You may need to purchase a dedicated IP address to have it changed, which can be costly. Another solution to IP filtering, although inconvenient, would be to migrate your site to a new host entirely, which would give you a new IP address.

Keyword Filtering

If your blog is a victim of keyword filtering, you will need to purchase another domain name. You should also avoid using that keyword in your blog title, the titles of your blog posts and pages, your tags and your categories, your images and media files. As explained above, keyword filtering targets a specific word, and the ultimate solution is to avoid using this world in all your URL paths.

Keyword filtering is particularly difficult to address, since you will need to change all the URLs within your site to avoid using the blocked word. In the process you will lose incoming links, pingbacks and links coming from popular search engines and aggregators. You may want to try using the Permalink Migration WordPress plugin, which will help you change your URLs without affecting your search engine rankings or breaking pre-existing links to your web site.

5 - Mirroring your WordPress blog

Mirroring is the process of making two or more domain names or subdomains contain the same up-to-date data. In this guide we will address two-way mirroring of a self-hosted WordPress blog. In two-way mirroring data is updated in both directions, keeping the two or more blogs in sync with each other by using the same WordPress installation and database. When you add, change, or delete any kind of content (posts, pages, comments, images, etc.) from the "source" blog at its regular domain or from the "target" blog or blogs at the mirrored domain(s), the same content will be added, changed, or deleted on the other blogs.

Unfortunately mirroring your blog is not a definitive or permanent solution to censorship. In most cases it is only a matter of time before the original censors find out your new mirror and block it as well. If they are closely monitoring your online activities, it will be very easy for them to block all your mirror blogs. Remember that censorship is a cat-and-mouse game and that the mirroring technique explained here may not be the ideal solution for you. It will only help you exploit the breach in the censorship wall by making your content available for a certain period of time, which might be shorter or longer depending on the censors' vigilance.

However, and the mirroring plugins highlighted in this guide allow you to create and manage as many blog mirrors as you want. This will make it easier for you to stay one step ahead of censorship by being prepared to mirror your content as many times as necessary. Even if your blog is not currently being censored, mirroring can be used as a fall-back mechanism for an eventual block.

To mirror your blog, you will need to follow a few steps:

a. Obtain and configure a new domain or subdomain
b. Choose, download and install a mirroring WordPress plugin
c. Configure the plugin
d. Hide your mirror blog from Google
e. Tell your readers about the mirrored site

a. Obtain and configure a new domain or subdomain

To mirror your blog at an alternate URL the first thing you need is the alternate URL. This can be either a subdomain of your existing site (i.e. mirror.yourblog.com) or an entirely new domain (mirrormyblog.com). The choice of which is better for you is a complex one and depends on the exact nature of the blocking/filtering that is making your site inaccessible.

Remember that if you are being blocked at the IP address level, a mirror site hosted on the same server, regardless of its domain, will also be blocked. See the section on IP address filtering above for advice on handling that situation.

Using subdomains (i.e. mirror.yourblog.com)

Sometimes a subdomain of your normal site URL will pass through filters that block your normal site, likely because the filter is only targeting your exact url ("http://yourblog.com" rather than "yourblog.com" or "yourblog"). If this is the case, then mirroring using a subdomain it is the ideal choice. It will be familiar to existing readers, involves less configuration, and in most cases will not incur extra charges with your registrar or host.

Most web hosts allow you to add and configure subdomains from your account's control panel. Read the help documentation for your hosting company or ask them for help if you aren't sure how to do this.

Testing whether a subdomain will make your site accessible to blocked users is fairly easy:

- Register a test subdomain with your host and point it at an empty server directory.
- Upload a simple index.html page with a test message in it.
- Ask a user that you know is blocked from your main site to visit it.

If the test page is not blocked for the someone who cannot otherwise access your blog, then a subdomain will probably work as a mirror location. Otherwise a new domain name may be needed.

Registering a new domain name

A completely new domain is needed if the filtering is done by nuanced domains or via keywords. In such cases subdomains will be blocked along with the main site, and even new domains containing the blocked keywords may be inaccessible. Careful consideration should be given to your choice of a new domain:

- It costs money to register new domain names.
- If you are being keyword blocked, you need a domain that does not contain the blocked keyword.

- Ideally this domain should still be memorable and meaningful to your existing and new readers.

New domains can be registered either from your existing hosting provider or from a dedicated registrar like GoDaddy. If your host offers domain registration, then using their service can be beneficial because it simplifies the process of configuring the domain as they will handle DNS settings for you.*

Once a new domain is registered it can take anywhere from 12 to 72 hours to become active because the DNS settings propagate slowly throughout the internet. While waiting you should prepare for the rest of the tasks below. Once you are able to point your browser to the domain you have created and see the landing page provided by your hosting service, then you are ready to continue.

DNS configuration

The next step is to log in to your web host panel interface (the following examples use DreamHost hosting service) and select Domains > Manage Domains. Your first domain (the one that is blocked) needs to be Fully Hosted. For the domain or subdomain that you want to set as the mirror, click the edit button. In the edit screen you will see several options. For example, DreamHost offers five options:

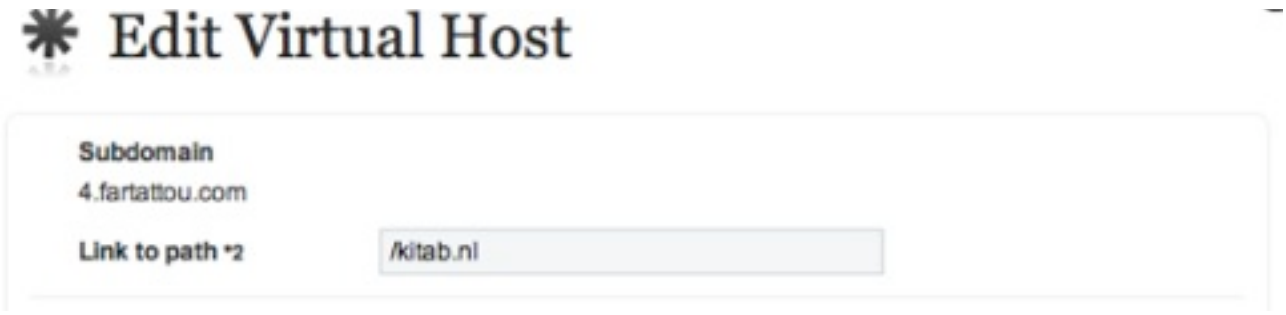
1. Fully Hosted;
2. Redirected;
3. Mirrored;
4. Parked;
5. Cloaked.

Select the Mirrored option, and set the domain to mirror (yourmirrorblog.com) to use your Fully Hosted domain (yourblockedblog.com):

* Many hosts charge exorbitant rates for registering new domains compared with dedicated registrars like GoDaddy (a .com/.net/.org domain should be about 10$ US), so the extra work of pointing a new domain at your host with DNS can be worth it in the long run. You should compare your host's prices with competitors before deciding.

Editing the virtual host

If you are using a web hosting service other than DreamHost, request help from your hosting company on how to configure virtual hosts and alter DNS settings. Normally you will only need to point your new domain or subdomain (example: 4.fartattou.com) to the root directory of your primary blog (example: /kitab.nl):



b. Choosing, downloading & installing a mirroring plugin

Choosing a plugin

Several WordPress plugins exist to help you to with the mirroring process. We recommend using one of the following:

- Domain Mirror Plugin by David McAleavy
- Domain Theme by Stephen Carroll

Both plugins allow a single WordPress installation to display different URL paths, blog titles and different domains. They also allow you to associate different themes with different domains, meaning that you can use a dynamic Web 2.0-style theme for your main blog and a minimalist theme for the mirror blog that speeds up load times and minimizes bandwidth use. Minimalist WordPress themes are recommended for blocked blogs. Fast page load times are crucial for visitors accessing blocked blogs via proxies as well as for visitors living in places with poor connectivity.

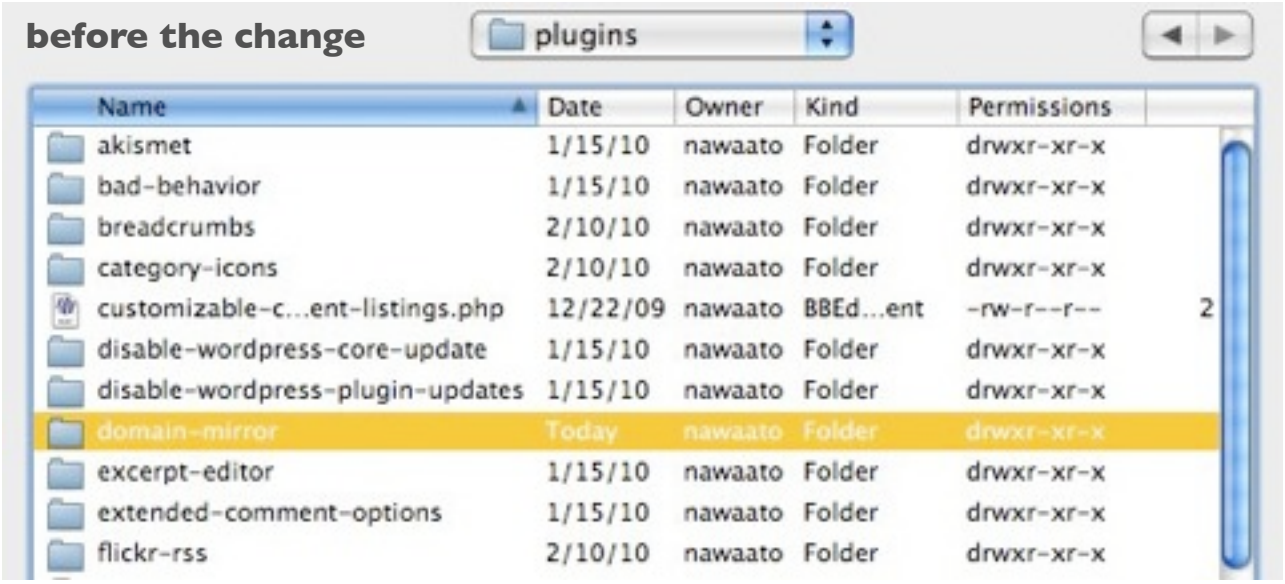
This guide includes instructions for configuring the Domain Mirror plugin. For instructions on using the Domain Theme plugin, please visit Stephen Carroll's plugin site.

Downloading and installing the plugin

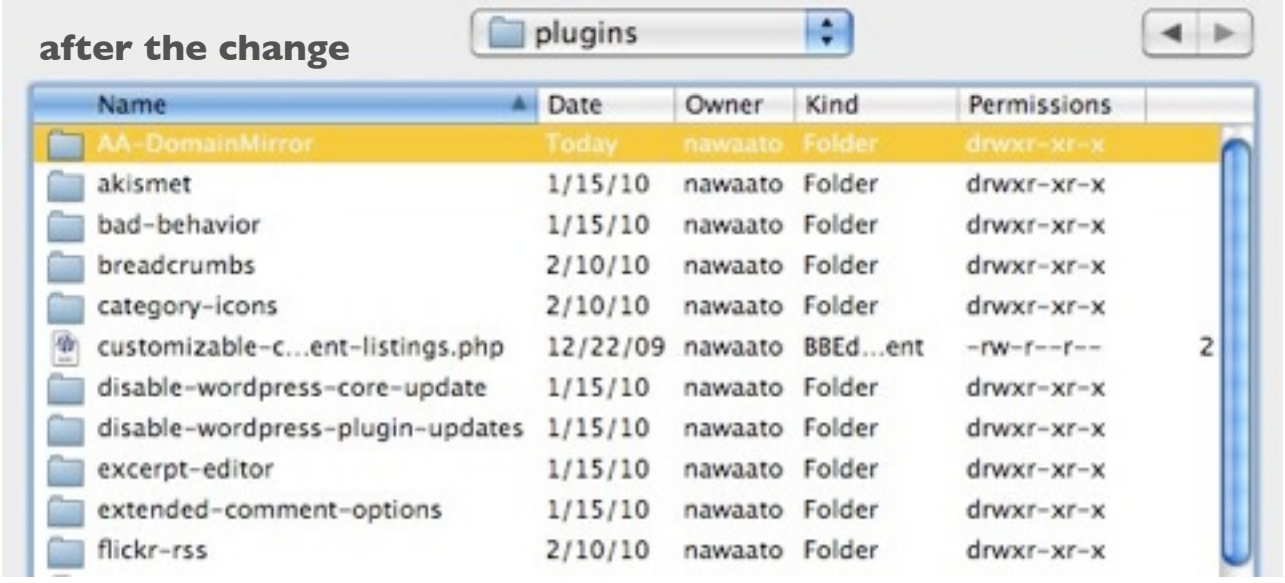
Download the Domain Mirror plugin, unzip it and upload it to the wp-content/plugins/ folder using FTP. You can also install the plugin directly from the WordPress dashboard by going to Plugins > Add New and searching for the plugin name. After finding the plugin you want, click the "Install" link to the right of the search result:

if you choose to install the Domain Mirror plugin from the Wordpress admin interface, you will need to rename the downloaded "domain-mirror" directory to "AA-DomainMirror" in order to force this plugin to load first, which will prevent compatibility problems with other plugins. To do this you will need FTP access to the plugin directory.

before the change



after the change

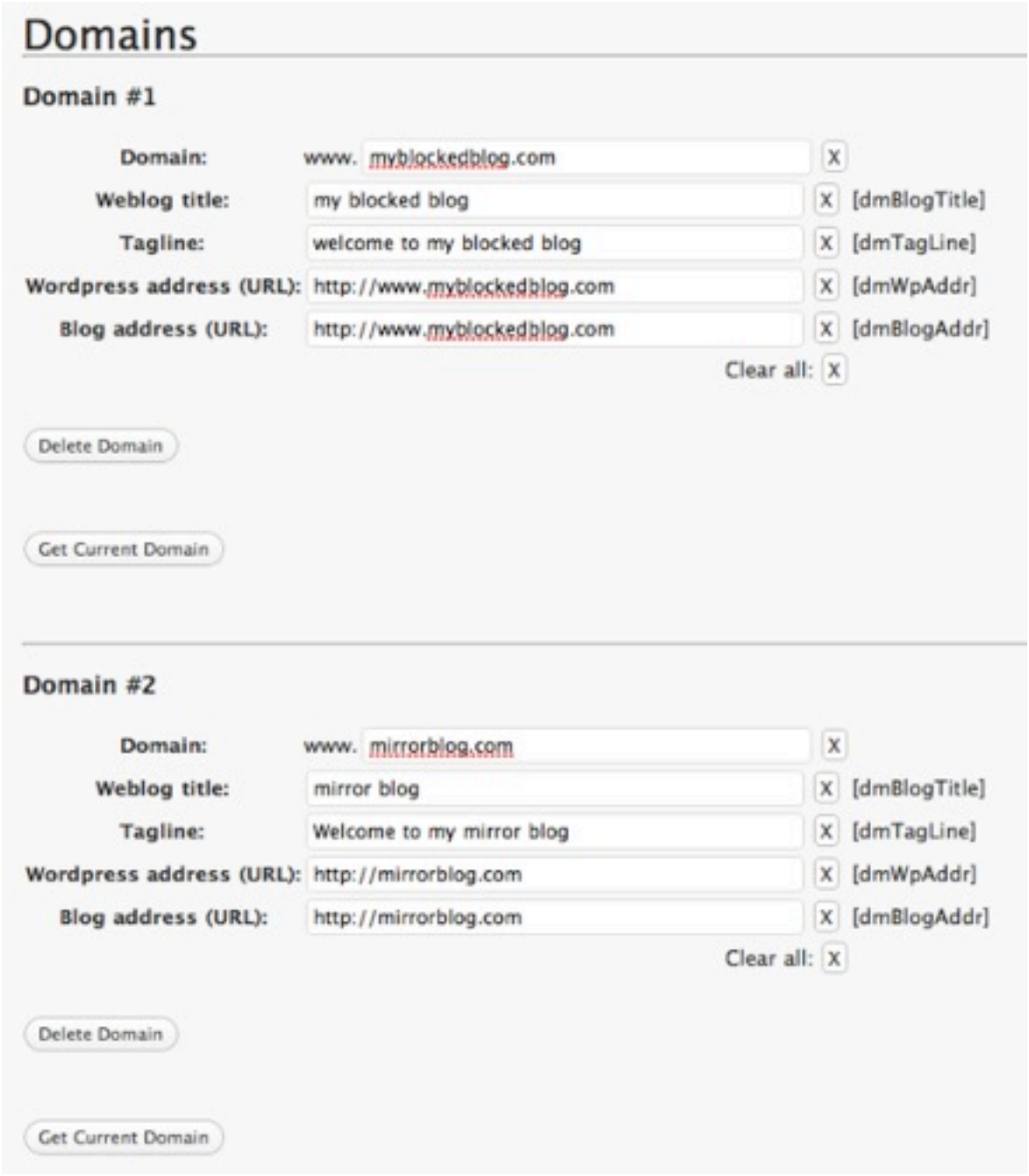


c. Configuring the Domain Mirror Plugin

In your WordPress dashboard, go to your Plugins page and activate the plugin you installed. Then go to Dashboard > Settings > Domain Mirror and fill in the appropriate information for your domain names.

The Domain #1 section should contain the basic information of your primary blog. Click the "Get Current Domain" button to get the values from the database as saved in your WordPress General Settings. The Domain #2 section should contain the details for the mirror domain. After adding this information click "Save Changes."

You can add as many mirror domains as you like by clicking the "Add New Domain" button.



If you have followed the steps above, you now have two copies of your blog. When you visit your primary domain, your blog remains unchanged. When you visit your new mirror site, the blog appears as if configured for that domain. You can see a live example of the Tunisian collective blog using this technique: the primary blog is at nawaat.org and the mirror blog is at twitter.nawaat.org. You will notice that the primary blog (on the right) has a complex theme, while the mirror blog (on the left) uses a minimalist theme to ensure a fast page load times.

d. Hiding your mirror blog from Google

Censors, like regular Internet users, often use Google and other search engines to find online content. If your mirror site is being indexed by Google, censors may be able to find it and block it quickly. You can help prevent this by preventing Google and other search crawlers from indexing your mirror blogs. This may also make it more difficult for new readers to find your blog. For this reason, we recommend spreading the news about your mirror blog using Twitter, an e-mail list, Facebook, Google Buzz and other social media tools.

You can prevent Google from indexing your mirror blog by creating or editing a Robots Exclusion Protocol (also known as robot.txt). This file tells search engines where and where not to look for content on your server and will prevent your mirror domain from being indexed, reducing the risk of censors finding your mirror blog.

If a robot.txt file does not exist in the folder that contains the content for your mirror site, you will need to create one. You can use any text editing software (Notepad or Wordpad for Windows, TextEdit for Mac OS, Vi or Emacs for Linux) to do this. Once you have created the file (or after you have opened the existing file), add the following text:

```
# Disallow Googlebot
User-agent: Googlebot
Disallow: /

User-agent: *
Disallow: /
```

Save the file as robot.txt and upload it to your mirror folder. Be careful not to put the file in the folder that contains your main blog, or you will prevent search engines from indexing your site entirely.

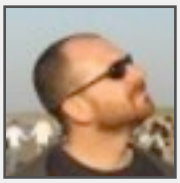
Telling your readers about the mirrored site

After setting up the domain mirror for your blog, your next step is to help readers discover the new link. There are many ways to do this:

- Add a notice to the header of your RSS feed informing your RSS subscribers about the new mirror. Style the notice with CSS so it will be visible. You can use the RSS plugin to add and style the notice.
- If you have an email list, notify your subscribers about the new link.

- If you use TwitterFeed to automatically publish your blog updates to Twitter, make sure to change the link of your blog RSS feed to the RSS feed for the mirror site. This will make sure that all the links published on Twitter will point to the mirror blog and not to the primary blog. See our Cross-Posting for Advocacy guide for tips on how to implement this technique.
- If you have a Facebook account, you can easily import your mirrored blog. You can use the WPBOOK plugin which will add your WordPress blog as a Facebook application. Make sure to use the URL of the mirror blog.

About the authors



Sami Ben Gharbia: Tunisian blogger and activist based in The Netherlands. Sami is Global Voices Advocacy director and behind the Threatened Voices mapping project. The co-founder of nawaat.org (Tunisian collective blog about news and politics), cybversion.org (a blog monitoring online censorship in Tunisia), babtounes (a wordpress twitter client aggregating tweets about Tunisia) and many other online digital activism projects.

Rebekah Heacock: works for Harvard's Berkman Center for Internet and Society as a research assistant for the OpenNet Initiative. She writes about technology, aid & development, East African public transit at Jackfruity. She is also a researcher for Global Voices' Technology for Transparency Network.

Jeremy Clarke: a PHP, HTML, CSS and WordPress hacker from Montreal. He is the developer and designer of Global Voices and participating in open-source software like wordpress His blog is at SimianUprising.com. To learn more about how Global Voices was built and designed see our Design and Tech page. You can watch Jeremy's Intro to WordPress Theming at wordpress.tv.