Blocking LLM scrapers on Alibaba Cloud from your nginx configuration

There are currently LLM scrapers running off many Alibaba Cloud IPs, that ignore `robots.txt` and pretend to be desktop browsers. They also generate absurd request rates, to the point of being basically a DDoS attack. One way to deal with them is to simply block *all* of Alibaba Cloud.

> This will also block legitimate users of Alibaba Cloud!

Here's how you can block them:

1. Generate a deny entry list at https://www.enjen.net/asn-blocklist/index.php?asn=45102&type=nginx
2. Add the entries to your nginx configuration. It goes directly in the `server { ... }` block.

On NixOS

If you're using Nix or NixOS, you can keep the deny list in a separate file, which makes it easier to maintain and won't clutter up your nginx configuration as much. It would look something like this:

```
services.nginx.virtualHosts.<name>.extraConfig = ''
  ${import ./alibaba-blocklist.nix}
  # other config goes here
''
```

... where you replace `<name>` with the name of your hostname.

Revision #1
Created 2025-02-19 19:07:39 UTC by joepie91
Updated 2025-02-19 19:13:42 UTC by joepie91