

discord.gg/hvRzgf6VYn

LIVE | WEBINAR

A crash course in automated ban management — set it and forget it



Thursday, December 14th | 4 p.m. UTC | 10 a.m. CST

((>)) LIVE | WEBINAR



Daniel Cave
Head of Product Marketing



Adrián Chaves
Python Developer

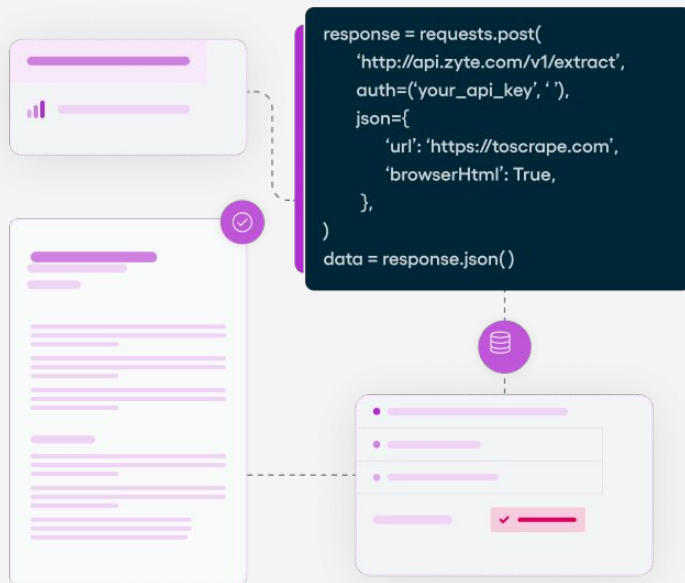
zyte



Beat Bans Smarter

A crash course in automated ban management — set it and forget it

A new **'automation first'** way of handling website bans, and the tech that enables an 'automation first' approach.



Agenda

Strategy

A better scraping strategy to make you faster, more efficient and lets you scale more easily.

Tech

Hands on code session with Adrian to building a MVP scraper that will show a better website unblocking strategy.

Q&A

Open Q&A - post questions in the chat.

Beat Bots Smarter

A crash course in automated ban
management

A new **automation first** way of handling
website bans, and the
an 'automation first' approach.

```
response = requests.post(  
    "http://api.zyte.com/v1/extract/  
    ...  
    browserId=1"
```

"If a fully automated API was as cheap as data centre proxies for simple sites, and can crack the hard sites too, it would make sense to use it first not last."

The background is a solid dark blue. In the top-left and bottom-right corners, there are abstract, hand-drawn style red shapes. These shapes consist of various line segments, dots, and small circles, some of which are connected to form larger, irregular structures. The shapes appear to be decorative elements that frame the central text.

The Traditional Strategies of Developers to Today

Factors that kill growth and experimentation

Time to **create** a scraping system that handles bans, and manually unblocking per website.

Time lost when **fixing configs** and infra when websites anti-bot profiles change long term.

Estimating Costs for any given site across multiple tools is hard.



Anti-bot Evolution

■ Site & ban complexity

Simple

Midrange

Sophisticated

■ Technologies required

DC IPs

Proxy
Rotator

Residential
IPS

Scriptable
Browser

Browser
Farm

■ Typical cost per million

\$XX



\$XXX



\$X,XXX

Using the most expensive tool for all sites makes no sense.

But this means using multiple tools



Problems & friction that generates

Combine
multiple tools,
each with their
own contract

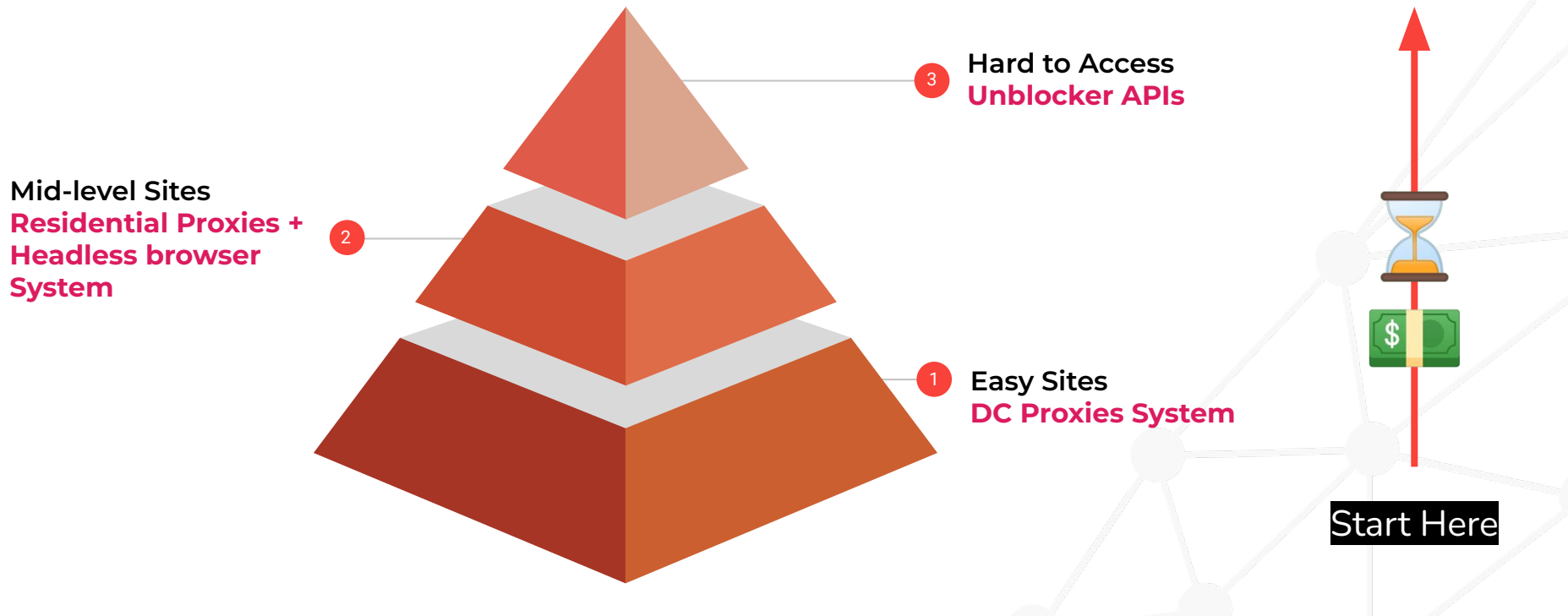
Then the **site**
changes, and it
breaks again

Unhappy customers
Unhappy dev team

Find a config that
works to solve an
opaque ban, monitor
success rates

Loss of **data**,
scrambling
to fix

Handling Anti-bot/Bans



Put simply...

If Data centre and headless doesn't work, I have two options:

- 1) Start experimenting with puppeteer and other tools.
- 2) Add an expensive 'unblocker' to the mix.

"When option 1 is too heavy of heavy lift, I use unblockers for harder sites."



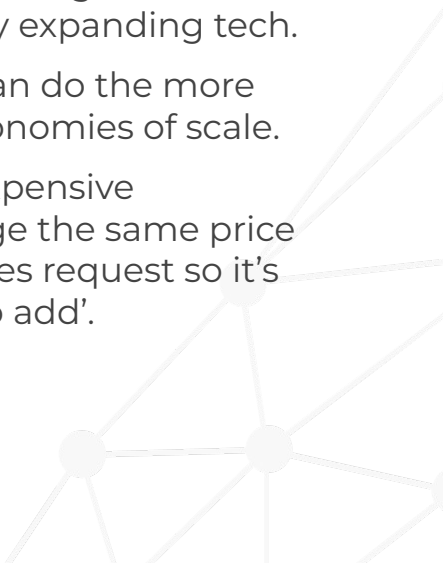


Easy to solve Ban Solutions

- ⚠ Your time is money, and building multiple systems using multiple vendors is a time sink with days of testing trial & error.
- ⚠ Incurs maintenance debt
- ⚠ Fragility: Data outages kill business
- ⚠ Multiple proxy subscriptions (overage risks), Hard to plan tech and time budgets.
- ⚠ Scalability: Any trial and error or manual configuration tasks kill productivity
- ⚠ Always Requires Dev to run
- ⚠ Opportunity cost of not working on core business

Hard to solve Ban Solutions

[All of the same problems Plus]

- ⚠ Accelerating diminishing returns of inhouse development.
 - ⚠ Increasing costs for acquiring less data while maintaining every expanding tech.
 - ⚠ Companies exist that can do the more advanced stuff with economies of scale.
 - ⚠ Many unblockers are expensive sledgehammers... charge the same price for easy sites as hard sites request so it's 'Just another tier/tool to add'.
- 

A better Strategy? Go Automation First



Use one API
for both easy
and hard sites

Both easy and hard
sites **Instantly
Unblocked** for you
using the **minimum
viable tech/cost** to
overcome a ban

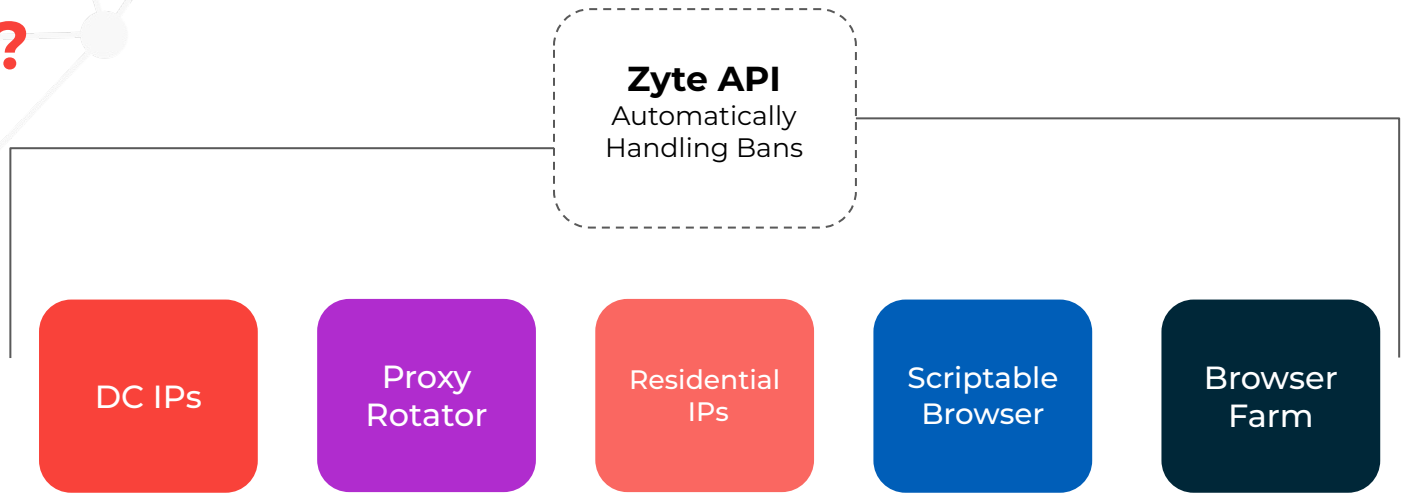
Requests priced per
site based on
difficulty

One tool that unblocks both/all types of site quicker, often cheaper, and definitely easier.

It's priced so you can use it for everything.

Simplifies your planning, budgeting, and development.

How?



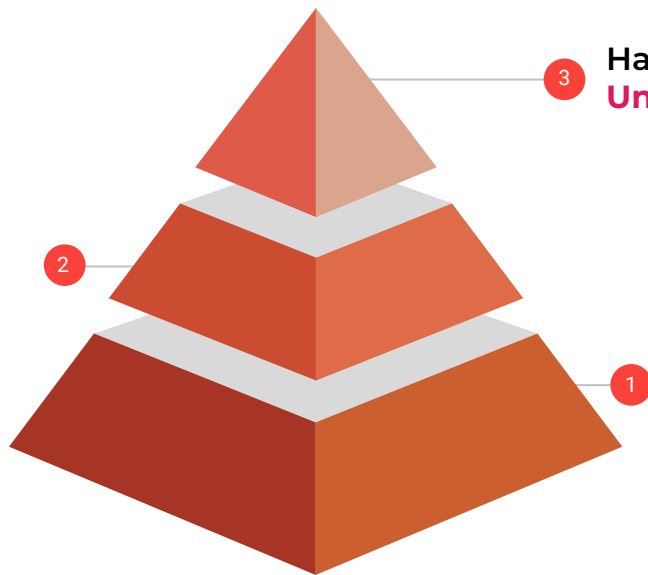
Automatically avoids bans using minimum required infrastructure

Per-site pricing - pay only for what each site needs

So you can use a single, automated tool without a cost trade-off

So you can turn all this effort and mess into...

Mid-level Sites
Residential Proxies
+ Headless
browser System



3 Hard to Access
Unblocker APIs

1 Easy Sites
DC Proxies System

End Here

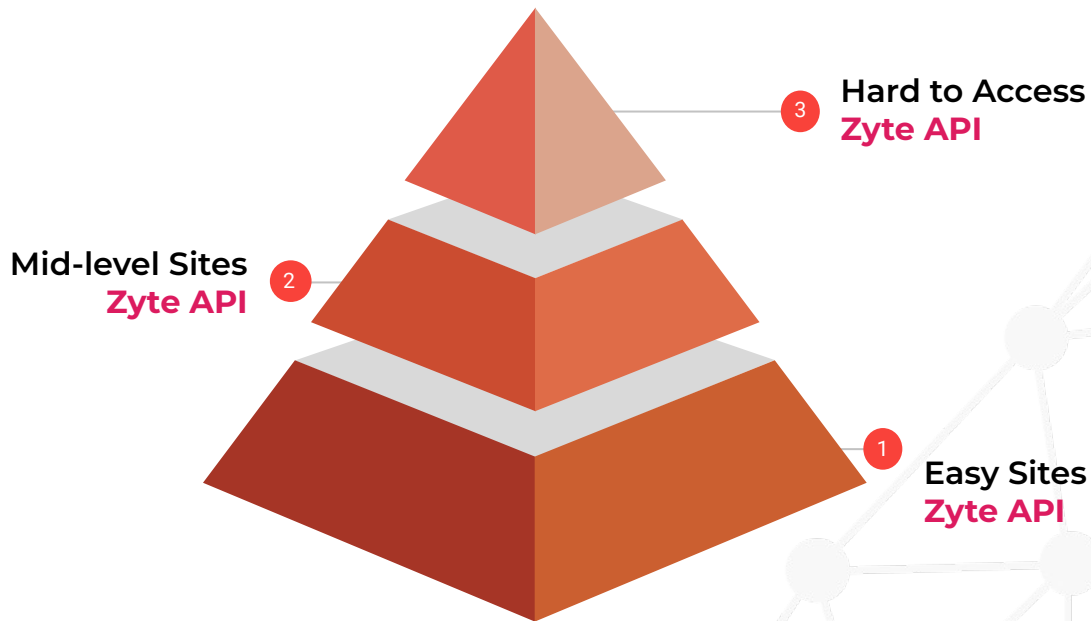
Start Here



Into an 'Automation first'

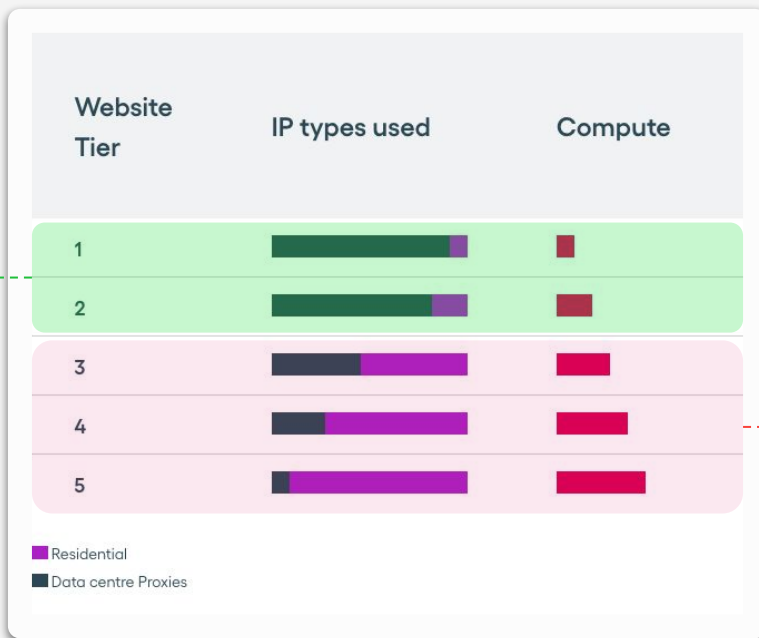
Set all your request to go through Zyte API from day one.

It automatically unblocks all sites for you.. The easy sites stay cheap, and the harder sites are solved using the leanest tech stack possible that heals itself.



Pricing model that makes it all possible.

Simple Bans



Complex Sites

Only pay the tech you need, and don't waste time figuring it out yourself (and monitoring bans) through trial and error.

Run Test Queries to see per-site pricing

Your Domains < Hide outputs Export

Domain	Requests per month	Rendered HTML	Screenshot	Data Extraction	Estimated Price	Actions
toscraper.com	1000000	☐	☐	☐	\$78/month	
toscraper.com	30000	☒	☐	☐	\$15/month	
toscraper.com	10000	☐	☒	☐	\$13/month	
toscraper.com	2000	☐	☐	☒	\$2/month	

Remove all

limit 1,042,000 requests per month

Estimated **\$108/month**

All prices include 61% **volume discount** for a similar monthly commitment

Use the price estimator to help budget and see volume discounts to plan budgets

'Automation first'
method in action.





Adrian Chaves

***Tech Demo: Build a MVP Spider
with Automated set-and-forget
anti-ban***

Developer Experience Team, Scrapy
Developer, Zyte

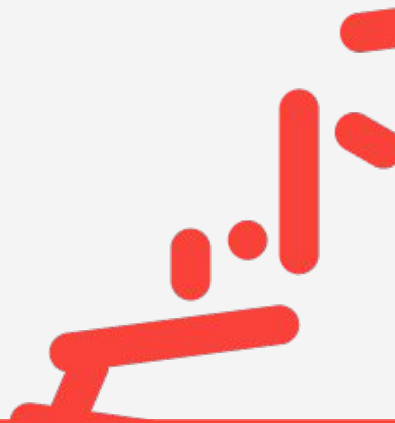
- ✓ Created a spider in minutes
- ✓ Got blocked
- ✓ Set Zyte API as website unblocker
- ✓ Beat bans instantly & cost efficiently
- ✓ Created little/no maintenance
- ✓ Turned on browser rendering
- ✓ Performed browser actions

Mitigating Factors that kill your ability to grow and experiment

~~Time to create a scraping system that doesn't get blocked per website.~~ **Time/cost vastly reduced**

~~Time and Data lost when (re)fixing configs and infra when websites change long term.~~ **Ban maintenance vastly reduced**

~~Uncertain costs for any given site across multiple tools.~~ **Budgets are easy to plan and stable**





Chris Becker

Proxyway Reviewer and tester.

[Read the report](#)

Don't just believe us:
Proxyway

“Zyte had a particularly strong showing, both in terms of success rate and response time.”

Proxyway's test results show that Zyte API has the highest success rate in the market, the fastest response time, and a solution that's competitively priced for handling all levels of bans.

“Zyte's pricing is interesting in general. [Zyte] adjusts request cost dynamically, based on the website's difficulty and optional parameters.”

Hands-on test from The Web Scraping Club



Pierluigi Vinciguerra

Co Founder of Databoutique.com
Author of [The Web Scraping Club](#)

“A stunning 100% success rate for all the websites

I'm really impressed by the Zyte API, when I first heard about it at the Extract Summit I was just curious about it but I thought it would have been difficult for it to tackle all the anti-bot around.

- Integration is straightforward
- 100% success rate
- Option to activate browser rendering makes it more affordable on larger volumes ”

“100/100”



Wrap up

The goal of our 'automation first' strategy is to **boost productivity and scalability with tech** that allows you to delegate effort to an API that automates away bans for you.

- Automate the solving of bans cost efficiently on a per-site basis
- Automate the self-healing and monitoring of bans for you

Zyte API completely short circuits the current industry standard strategies by letting users have one tool that does both types of site quicker, cheaper, more easily, with less admin and less risk of overages, or data outages.



Try Zyte API (free)

<https://www.zyte.com/zyte-api/>

Join us on discord

<https://discord.gg/hvRzgf6VYn>



in next webinar...

“ We need to add EMEA to our

“

If only there was a way I could
automate away build selectors too...
I might be able to scale even faster.

I wonder if Zyte can do that too?....



Data Team Lead

Business L

Q&A