

Boye Domain Discovery

June 2021



BOYE & CO



A flashback to warmer days of 2019 :)



Why is Domain Discovery so important?

In a nutshell **Brand Protection**

Why its good to find all your websites?

Brand Image

Information accuracy

Customers experience improvements

Security

Compliance with Regulation

Why is Domain Discovery so important?

Privacy

Security

Accessibility

Quality

Performance

Sustainability

How to we do Domain Discovery?

Vertical domain correlation



scholar.google.com



maps.google.com

Horizontal domain correlation



gmail.com

google.cz

google.com

youtube.com

blogger.com

Where do we get data from?

We use several sub-domain finding tools to guarantee complete coverage

Data is gathered from:

- 1) DNS entries
- 2) SSL Certificates entries
- 3) Search engines results

A quick audit on some top global Universities...

Danish Universities

ku.dk – 8,500

au.dk – 4,800

German Universities

lmu.de – 2,800 (LMU Munich)

tum.de – 29,000 (Tech Uni Munich)

uni-heidelberg.de – 63,500 (Heidelberg)

unibo.it – 10,700

usal.es – 4,200

u-paris.fr – 69

UK Universities

ox.ac.uk – 94,700

cam.ac.uk – 47,700

ucl.ac.uk – 55,300

lse.ac.uk – 19,400

ed.ac.uk – 49,300

US Universities









princeton.edu – 34,100

columbia.edu – 142,500

mit.edu – 93,500

yale.edu – 38,300

stanford.edu – 113,100

Domain		Site Title
 aarhus19.boy-co.com		—
 www.brooklyn21.boy-co.com	200	Boye 21 Brooklyn Conference
 aarhus20.boy-co.com		—
 brooklyn20.boy-co.com	200	Boye 20 Brooklyn Conference - digital leadership conference
 www.boy-co.com	200	Boye & Company - next-stage networking
 2018.boy-co.com		—
 brooklyn19.boy-co.com		—
 brooklyn21.boy-co.com	200	Boye 21 Brooklyn Conference

The Domain Discovery data has serious quality problems!

Automated discovery while producing incredibly valuable data for business also creates tons of hay to hide the needles in...

Example from Harvard.edu domain below is finding 8000 sub-domains starting with "dhcp."

That is 8% of the total of their estate

It is important to be able to find these and then hide them once a human has reviewed the basic rule...

📄 harvard.edu 8110 rows

<input type="checkbox"/>	Site Name & URL	Screenshot	Status Code	Page Title	Site Owner
<input type="checkbox"/>	dhcp-203-232.nmr.mgh.harvard.edu	No Image	-	-	-
<input type="checkbox"/>	dhcp-203-222.nmr.mgh.harvard.edu	No Image	-	-	-
<input type="checkbox"/>	dhcp-203-162.nmr.mgh.harvard.edu	No Image	-	-	-
<input type="checkbox"/>	dhcp-202-243.nmr.mgh.harvard.edu	No Image	-	-	-
<input type="checkbox"/>	dhcp-202-229.nmr.mgh.harvard.edu	No Image	-	-	-

Do Domain Management tools solve the issues of Domain Discovery?

NO!

They don't focus on:

- Sub-domains
- Business ownership
- Ignore rules to automate review
- Version history of change
- Quality measurement reporting

Challenges facing Universities (and so many organisations)

Huge digital estates of thousands of websites

Dozens or hundreds of Content management platforms

Unclear ownership of domains, sites, pages, assets

Staying compliant with regulations and brand standards

Poor customer experiences due to low quality and fragmentation

A quick look at some of the solutions we are building

Site owners and being Social

littleforest.co.uk 15 of 15 rows

Search...

+ Add

Run Status Code





Edit

Remove

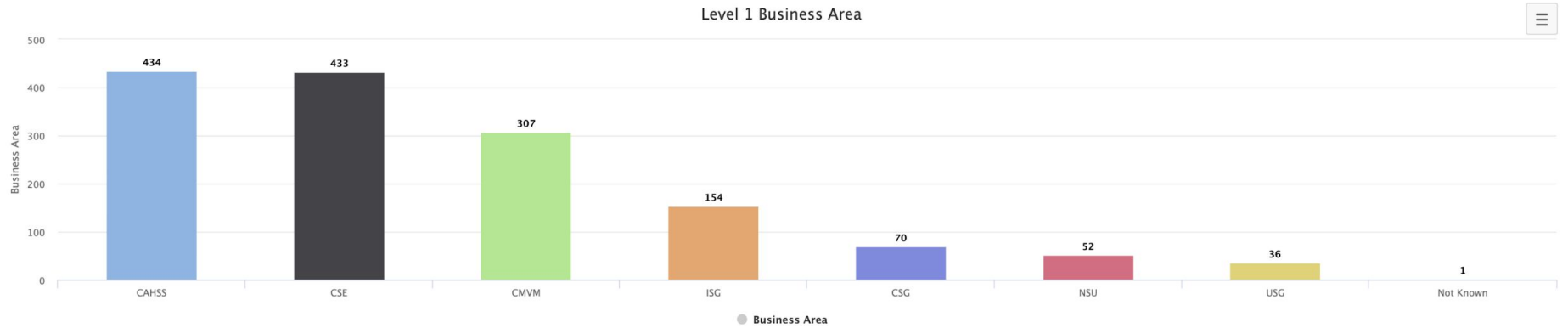
Share

Columns Show 100 entries

Order by Modified Descending Export

<input type="checkbox"/>	Site Name & URL	Status Code	Site Owner	Category	Run Status	Created
<input type="checkbox"/>	https://www2.littleforest.edu/	200	 John Edwards	Urgent to Resolve	Ready	2021-04-06 13:13
<input type="checkbox"/>	https://finance.littleforest.edu/	200	 Jenny Rei	All Clear	Ready	2021-04-06 11:27
<input type="checkbox"/>	https://staging.littleforest.co.uk/	200	 Jenny Rei	Needs IT Dev	Ready	2021-04-06 09:55
<input type="checkbox"/>	http://test.littleforest.co.uk/	200	 Ben Timmons	Needs IT Dev	Ready	2021-04-06 13:14

Visualising domains by Business area



Affiliates and Institute Report



+ Add Site





▶ Run

⚙️ Edit

🗑️ Delete

Export

Show More

	Screenshot	Site	CMS	CMS Version	Accessibility Errors	SSL Certificate Issues	Performance Score	SEO Score	Best Practice Score
<input type="checkbox"/>		ProPublica https://president.harvard.edu/	Wordpress	5.5	3	No	79	54	49
<input type="checkbox"/>		Wyss Institute https://www.health.harvard.edu	Drupal	9	5	Yes	51	53	75
<input type="checkbox"/>		Boston Childrens Hospital https://www.hsph.harvard.edu	Wordpress	5.3	4	No	51	83	81
<input type="checkbox"/>		Joslin Diabetes https://blogs.harvard.edu	Wordpress	5.3	10	No	69	77	61

Content Policies report

Security & Brand

[Report Copyright Infringement](#)

[Report Security Issue](#)

[Trademark Notice](#)

Website

[Accessibility](#)

[Digital Accessibility](#)

[Privacy Statement](#)

Get In Touch

[Contact Harvard](#)

[Maps & Directions](#)

[Jobs](#)

HARVARD
UNIVERSITY



Content Policies report

👁 Policies Report - Exact Match



+ Add Site



▶ Run

⚙ Edit

🗑 Delete

Export

Show More

	Screenshot	Site	Privacy	Accessibility	Terms and Conditions	Data Protection	Modern Slavery	Trademark Notice	Copyright Infringement
<input type="checkbox"/>		president.harvard.edu https://president.harvard.edu/	No	Yes	No	No	Yes	Yes	Yes
<input type="checkbox"/>		health.harvard.edu https://www.health.harvard.edu	No	Yes	Yes	Yes	No	No	Yes

SSL Certificate Checks

SSL Check Report



+ Add Site



▶ Run

⚙ Edit






🗑 Delete

Export





Show More

	Screenshot ▾	Site ▾	Issued by ▾	Valid to ▾	Certificate grade ▾	Issued to ▾	Days to expiry ▾
<input type="checkbox"/>		Harvard University https://www.harvard.edu/ 🔗	Let's Encrypt (US)	2021-05-20	A	None	85 days left
<input type="checkbox"/>		chandra.harvard.edu https://chandra.harvard.edu/ 🔗	Internet2 (US)	2021-06-04	A	President and Fellows of Harvard College	99 days left

Dealing with Redirects and Timeouts

<input type="checkbox"/>	Site Name & URL	Screenshot	Status Code	Page Title
<input type="checkbox"/>	arthur.harvard.edu	No Image	Timeout	-
<input type="checkbox"/>	college.harvard.edu/		200	Harvard College H...
<input type="checkbox"/>	oue.fas.harvard.edu/		200	Office of Undergra...
<input type="checkbox"/>	pfoho.harvard.edu/pfoho		302 200	Pforzheimer House
<input type="checkbox"/>	adamshouse.harvard.edu/adamshouse		302 200	Adams House
<input type="checkbox"/>	fap.fas.harvard.edu/fap		302 200	Harvard First-Year ...

Version history of changes





 Domain	 Sub-domains	 Changed	+ New	 Deleted
TEST harvard.edu	24	24	29	5
harvard.edu	25000	103	25000	0

Version history of changes

🔍 <https://arboretum.harvard.edu/> History

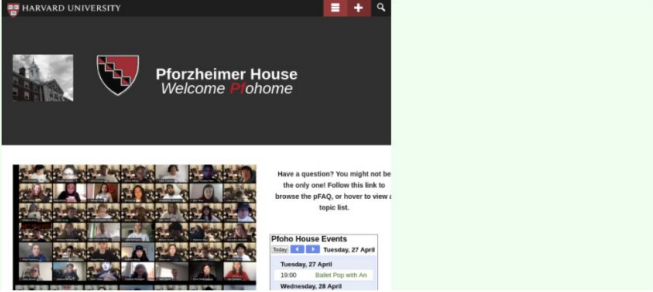
🔍 Search...

Show 10 entries

User Date and Time	Change	Site Name & URL	Restore this version
 Little Forest 27 Apr 2021 15:15:23	Variable https://arboretum.harvard.edu/ Modified	https://arboretum.harvard.edu/	🔄
 Little Forest 22 Apr 2021 14:13:45	Variable https://arboretum.harvard.edu/ Modified	https://arboretum.harvard.edu/	🔄
 Little Forest 22 Apr 2021 13:25:48	Variable https://arboretum.harvard.edu/ Modified	https://arboretum.harvard.edu/	🔄
 Little Forest 19 Apr 2021 22:36:45	Variable https://arboretum.harvard.edu/ Modified	https://arboretum.harvard.edu/	🔄

Version history of changes

🕒 pfoho.harvard.edu/pfoho History ✕

Column	Old Version	Current Version
Screenshot	-	
Status Code	-	302,200
Page Title	-	Pforzheimer House

By Little Forest in 2021-04-27 15:17:38

Cancel Restore

Reporting on Chrome Privacy warnings

Reporting on Chrome Privacy warnings



Your connection is not private

Attackers might be trying to steal your information from [REDACTED] (for example, passwords, messages, or credit cards). NET::ERR_CERT_COMMON_NAME_INVALID

ADVANCED

[Back to safety](#)

Reporting on Chrome warnings



The image shows a Google search interface. The search bar contains the text "Huerta Hof German Shepherds". Below the search bar, there are navigation links for "All", "Maps", "News", "Videos", "Shopping", "More", and "Search tools". The search results show "About 2,980 results (0.36 seconds)". The first result is "Huerta Hof German Shepherds - Facebook" with a URL and a description. The second result is "Huerta Hof German Shepherds" with a URL and a security warning: "This site may harm your computer." A red arrow points to the warning text.

Google

Huerta Hof German Shepherds

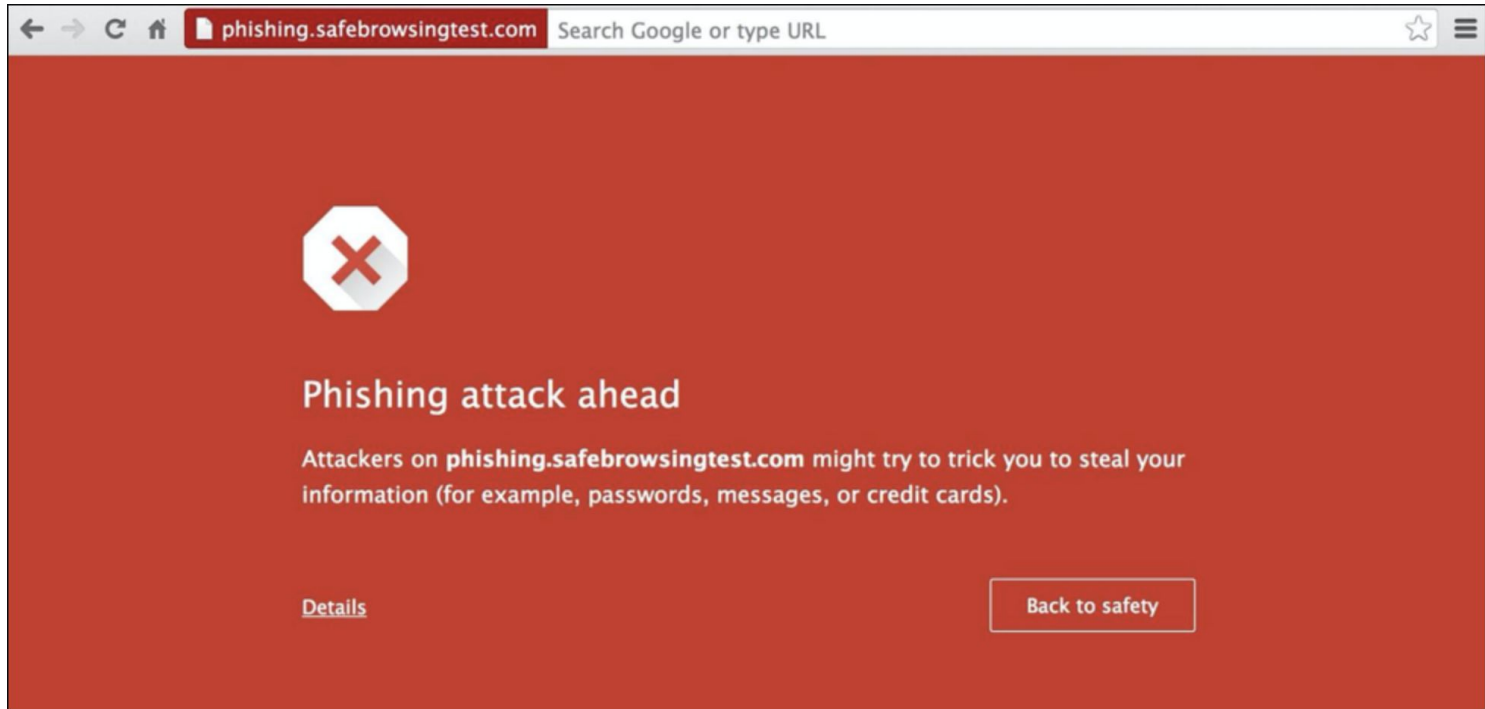
All Maps News Videos Shopping More Search tools

About 2,980 results (0.36 seconds)

Huerta Hof German Shepherds - Facebook
<https://www.facebook.com/HuertaHofGermanShepherds/>
Huerta Hof German Shepherds. 1117 likes · 30 talking about this. Huerta-Hof German Shepherds, where breeding is still considered an art!

Huerta Hof German Shepherds
<teamhuertahof.com/>
This site may harm your computer.

Reporting on Chrome warnings



Reporting on Chrome warnings



The site ahead contains harmful programs

Attackers on [REDACTED] might attempt to trick you into installing programs that harm your browsing experience (for example, by changing your homepage or showing extra ads on sites you visit).

Automatically report details of possible security incidents to Google. [Privacy policy](#)

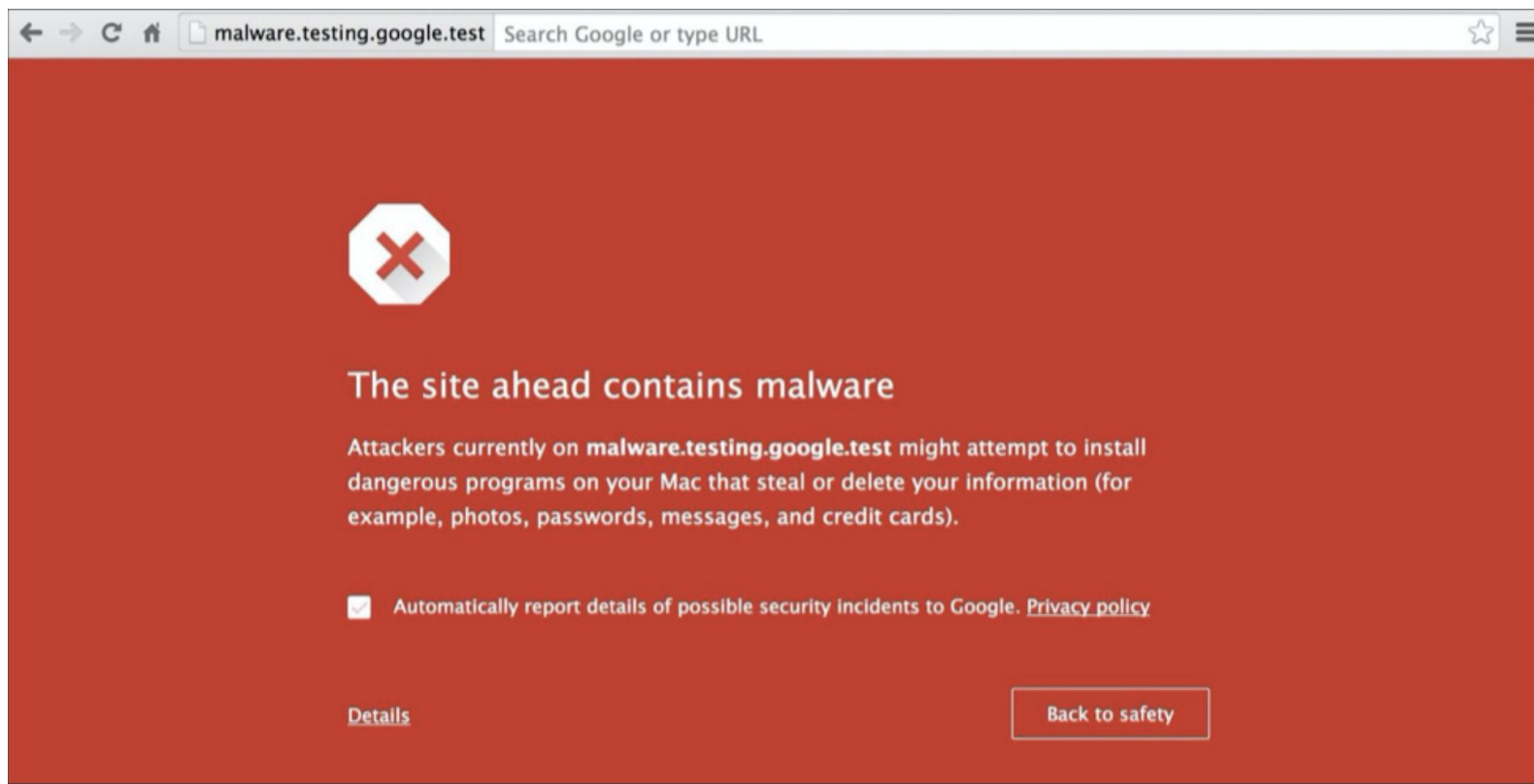
[Hide details](#)

[Back to safety](#)

Google Safe Browsing recently [found harmful programs](#) on www.ownedcore.com.


If you understand the risks to your security, you may [visit this site](#) before the harmful programs have been removed.

Reporting on Chrome warnings



The screenshot shows a Chrome browser window with the address bar displaying `malware.testing.google.test`. The page content is a red warning message. At the top left of the page is a white octagon with a red 'X' inside. Below this is the heading "The site ahead contains malware". The main text reads: "Attackers currently on `malware.testing.google.test` might attempt to install dangerous programs on your Mac that steal or delete your information (for example, photos, passwords, messages, and credit cards)." Below the text is a checkbox that is checked, with the label "Automatically report details of possible security incidents to Google. [Privacy policy](#)". At the bottom left is a link labeled "Details", and at the bottom right is a button labeled "Back to safety".

malware.testing.google.test Search Google or type URL



The site ahead contains malware

Attackers currently on `malware.testing.google.test` might attempt to install dangerous programs on your Mac that steal or delete your information (for example, photos, passwords, messages, and credit cards).

Automatically report details of possible security incidents to Google. [Privacy policy](#)

[Details](#) [Back to safety](#)

Reporting on Chrome Privacy warnings

<input type="checkbox"/>	www.firegrid.eng.ed.ac.uk		Privacy 200	Select an installatio...
--------------------------	--	---	-------------	--------------------------

Important learnings in our journey

1. A central team cannot manage a Web Registry entirely themselves, they must be the owners and then involve many business units to help - for this we create specific “views”
2. The service must create Rules to automatically categorise URLs based on patterns e.g “one.maths.ox.ac.uk” belongs to the “Maths” department
3. We include an “Unassigned” list which all business owners can review to see if any of their sites are in this category

Benefits of your Domain Discovery platform

1. Stay informed and in control of your entire web estate
2. Understand quality levels and Accessibility of all websites
3. Monitor customer experience by eliminating rogue sites
4. Save time via automation and monitoring changes for you
5. Agree ownership of sites within your organisation
6. Improve communication and control across your digital organisation

We offer a **free** Domain Discovery report for Boye & Company members

Can your organisation or a client benefit from this?

Get in touch :)

Thank you



[Edit profile](#)

Gavin Colborne

@gavincolborne

Managing Director of Little Forest [@thelittleforest](#) providing Digital Governance and Accessibility Reporting for large web estates.

 London  littleforest.co.uk  Joined June 2009