PRODUCT BROCHURE

SNI Classification Database

IAB content categories, web-filtering categories, and the traffic role and owner domain of each delivered hostname that remains readable in encrypted traffic.

1,880,282,561

hostnames delivered with traffic role and owner

of more than 2.7 billion hostnames held

IAB CONTENT TAXONOMY · WEB-FILTERING CATEGORIES · ROLES AND OWNERS · WEEKLY UPDATES

For the hostnames that stay readable in encrypted traffic: what they are about, how policy should treat them, what they do and who runs them.

1,880,282,561 hostnames delivered with traffic role and owner	**1,725,047,991** of them with an IAB content category	**1,626,210,215** with a web-filtering category
703 / 59 IAB Content Taxonomy 3.0 categories / web-filtering categories	**16** roles: 13 traffic roles, from website and sign-in to advertising and telemetry, and 3 technical classes	**95.8 %** with an identified owner domain

THREE HOSTNAMES FROM THE DATABASE

	api.glb.paypal.com	images.ebay.at	googleads.g.doubleclick.net
IAB category	Personal Finance > Consumer Banking	Shopping	Business and Finance > Business > Marketing and Advertising
Web-filtering category	Financial Services	Marketplaces / Classifieds / Auctions	Advertising/Marketing
Role	API	CDN	Advertising and tracking
Owner domain	paypal.com	ebay.at	google.com

Separate visits from background traffic

Count the pages people open, and set aside images, interfaces, telemetry and advertising calls.

Apply policy per hostname

Block a category for the pages people open while the sign-in, interface and update hosts that applications depend on keep working.

Attribute traffic to its owner domain

Group hostnames by owner domain, including services that run under separate root domains.

Classify unlisted hostnames offline

Engine Offline, included with the database, gives new hostnames a role, an owner and the category of their root domain on your own servers.

02 COVERAGE AND SCALE

One page view, dozens of hostnames, each classified on its own.

Hostnames of bestbuy.com from the database. Only the first is the page a shopper chose to open; the others deliver images and scripts, run the shop's interfaces, sign customers in, measure visits, send mail and push updates.

HOSTNAME	ROLE	IAB CATEGORY	WEB-FILTERING CATEGORY	OWNER
www.bestbuy.com	**Website**	**Technology & Computing > Consumer Electronics**	**Shopping**	**bestbuy.com**
images.bestbuy.com	CDN	Technology & Computing > Consumer Electronics	Shopping	bestbuy.com
assets.bestbuy.com	CDN	Technology & Computing > Consumer Electronics	Shopping	bestbuy.com
api.bestbuy.com	API	Technology & Computing > Consumer Electronics	Shopping	bestbuy.com
my.bestbuy.com	Login and identity	Technology & Computing > Consumer Electronics	Shopping	bestbuy.com
smetrics.bestbuy.com	Advertising and tracking	Technology & Computing > Consumer Electronics	Shopping	bestbuy.com
tag1.bestbuy.com	Advertising and tracking	Technology & Computing > Consumer Electronics	Shopping	bestbuy.com
ingest.bbycontent.bestbuy.com	Telemetry	Technology & Computing > Consumer Electronics	Shopping	bestbuy.com
realtime.bestbuy.com	Push and real-time	Technology & Computing > Consumer Electronics	Shopping	bestbuy.com
mail10.bestbuy.com	Mail	Technology & Computing > Consumer Electronics	Shopping	bestbuy.com
download.digitallibrary.bestbuy.com	Downloads and updates	Technology & Computing > Consumer Electronics	Shopping	bestbuy.com

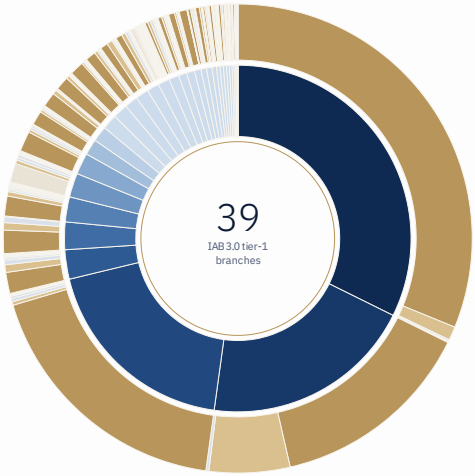
1,880,282,561	1,725,047,991	1,626,210,215	118,380,050
hostnames delivered with traffic role and owner	with an IAB content category	whose root domain has a web-filtering category	distinct root domains in the delivered file

Every hostname carries its root domain, for aggregation to organisation level and as the fallback key for hostnames not yet in the database. The database holds more than 2.7 billion hostnames; those without a classification in the current release are retained for later classification and are not included in the delivered file.

03 IAB CLASSIFICATION

One tiered vocabulary for every destination.

Categories follow the IAB Content Taxonomy, the same categories as our 120 million domain database, so hostname and domain results line up. Each record carries the full path from tier 1 to the deepest applicable tier, so policies and reports can be written broadly or with precision.



Inner ring: tier-1 categories, ordered by share. Outer ring: tier-2 categories within each tier-1 branch; the two largest are highlighted in gold, unshaded segments are records classified at tier 1. Websites and services; ISP customer lines and parked domains excluded. Sample of 1 billion SNIs.

FROM BROAD TO PRECISE

Websites and services; ISP customer lines and parked domains excluded.

Each category path runs from tier 1, the broad subject, through tier 2 and tier 3 to the deepest applicable tier. Policies can act on a whole branch, such as all of Technology & Computing, or on a single leaf.

39

tier-1 branches of IAB Content Taxonomy 3.0, all in use

703

categories in IAB Content Taxonomy 3.0

EXAMPLE CATEGORY, FOUR TIERS

TIER 1	Technology & Computing
TIER 2	Computing
TIER 3	Internet
TIER 4	Cloud Computing

LARGEST TIER-2 BRANCHES

Share of categorised SNIs. Websites and services; ISP customer lines and parked domains excluded.

TIER 1	TIER 2	SNIS IN SAMPLE	SHARE
Technology & Computing	Computing	225,590,109	31.2%
Personal Finance	Home Utilities (Internet Service Providers)	131,941,425	18.2%
Business and Finance	Industries	101,999,459	14.1%
Business and Finance	Business	40,273,098	5.6%
Entertainment	Music	11,655,125	1.6%

04 DEPTH AND CONFIDENCE

Precise where the hostname allows it, explicit about where a category comes from.

Many hostnames are parts of the same website: the page server, the image server, the server for scripts. They share the publisher's category, and their role shows what each one does. Hostnames that run a different service under the same domain receive their own category. The Method attribute states which applies to each row.

Categories are assigned per hostname record and written as the full IAB path. The Multiple Categories attribute carries the assigned category together with its confidence score.

DEPTH OF THE ASSIGNED CATEGORY PATH

Share of categorised SNIs by deepest tier assigned. Websites and services; ISP customer lines and parked domains excluded.

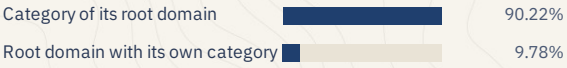


Tier 1	Tier 2	Tier 3	Tier 4
2.8%	**24.4%**	**50.1%**	**23.2%**

Sample of 1 billion SNIs

WHERE THE CATEGORY COMES FROM

Method attribute, share of categorised SNIs.



MULTIPLE CATEGORIES, AS DELIVERED

`api.glb.paypal.com`

Personal Finance > Consumer Banking (1.00)

Each record states its category with a confidence score, so consumers can set their own thresholds for automated action.

97.8% of categorised websites and services classified to tier 2 or deeper

73.3% classified to tier 3 or deeper

703 categories in IAB Content Taxonomy 3.0

91.7% of delivered hostnames carry an IAB category

05 WEB-FILTERING CATEGORIES

From content category to policy decision.

Firewalls, secure web gateways and DNS filters write their policies in web-filtering categories. Each hostname record carries both the IAB category that describes it and the filtering category a policy acts on.

A filtering category alone treats a domain as one block. Combined with the hostname record, a policy can block a category for the pages people open while the sign-in, interface and update hosts that other applications depend on keep working.

Where a site fits more than one category, every applicable filtering category is delivered, so mixed services do not disappear into a single bucket.

86.5%

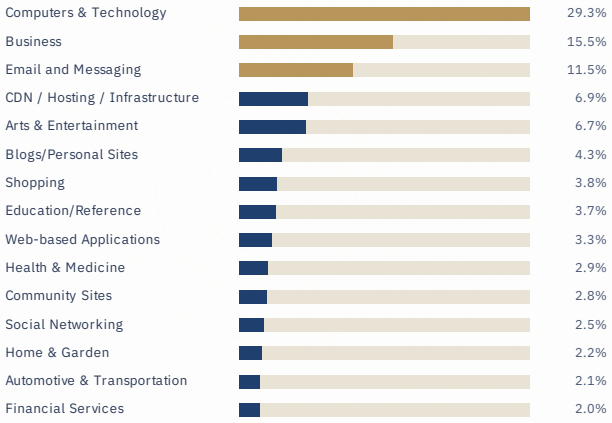
of delivered hostnames carry a web-filtering category

59

web-filtering categories, all present on the sampled SNIs

LEADING WEB-FILTERING CATEGORIES

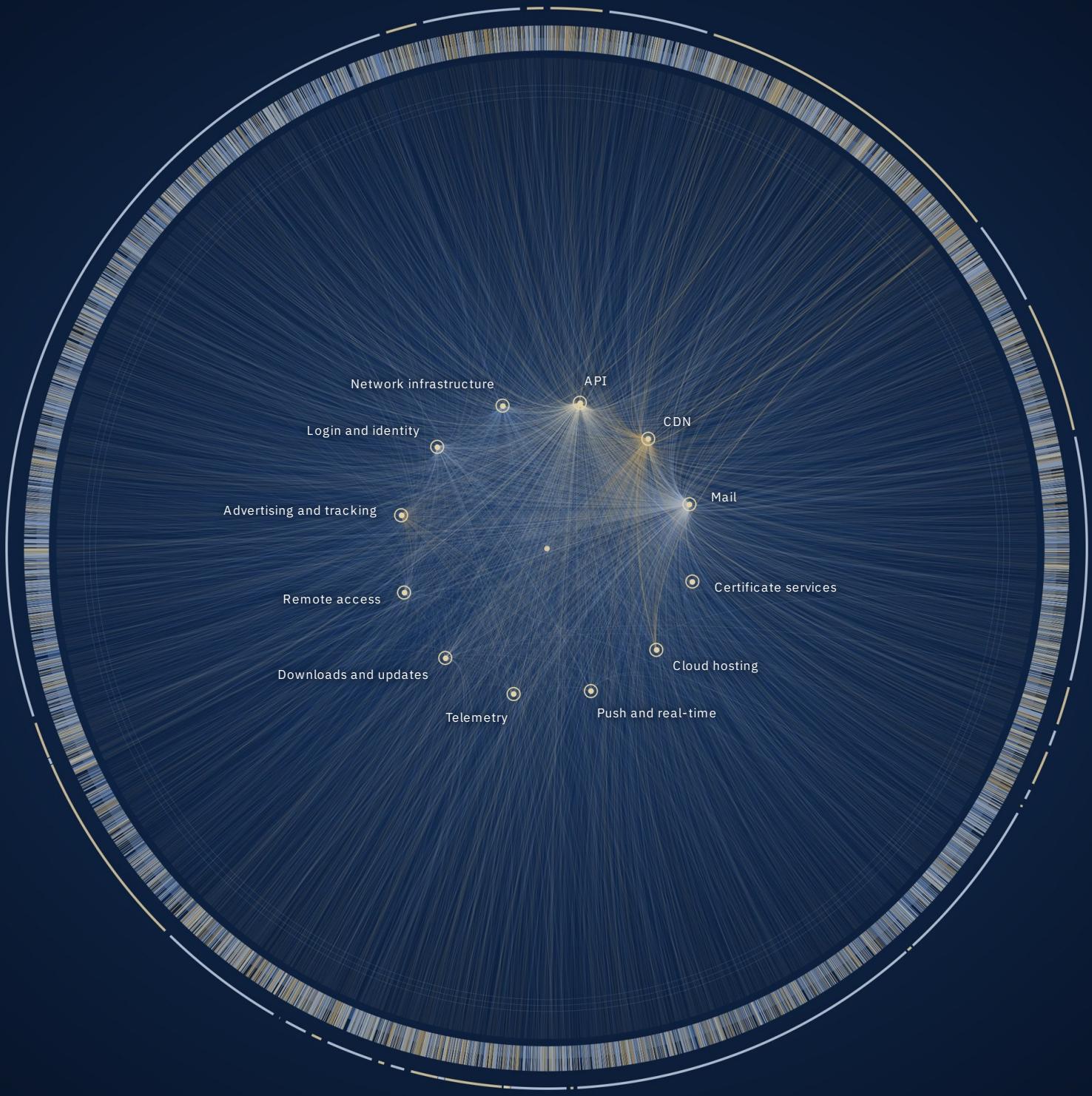
Share of sampled SNIs with a web-filtering category that carry each category; an SNI can carry several.



Sample of 1 billion SNIs

THE SAME HOSTNAMES, TWO VOCABULARIES

SNI	IAB CATEGORY	WEB-FILTERING CATEGORY	ROLE	OWNER
api.glb.paypal.com	Personal Finance > Consumer Banking	Financial Services	API	paypal.com
smtp-in.orange.fr	Personal Finance > Home Utilities > Internet Service Providers	Email and Messaging | Business	Mail	orange.fr
images.ebay.at	Shopping	Marketplaces / Classifieds / Auctions	CDN	ebay.at
accounts.google.co.in	Technology & Computing > Computing > Internet > Search	Search Engines & Platforms	Login and identity	google.co.in
metrics.barclays.co.uk	Personal Finance > Consumer Banking	Online Banking	Telemetry	barclays.co.uk



06 ROLES AND OWNERS

Every hostname has a function. The database names it.

The role separates what people visit from what runs in the background. Counted without it, a single article read becomes six visits; with it, reports keep what people chose to see and policies can treat a sign-in server differently from an advertising server on the same domain.

7,800 SNIs drawn at random from the database: 5,200 endpoints with a non-website role and 2,600 websites, ordered by tier-1 category around the outer ring. Each curve connects an endpoint to its functional role.

06 ROLES AND OWNERS

Function and ownership as additional dimensions.

Many hostnames do not live under the company's own domain: images on a content network, an application on a cloud address. Owner gives the owner domain of the hostname, so traffic can be grouped by the domain of the business that runs the service rather than by the server address.

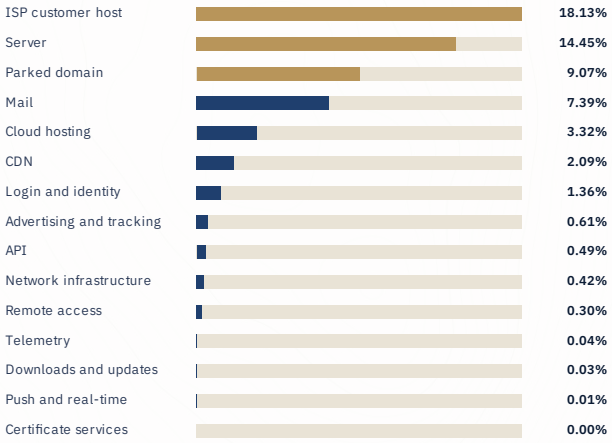
The Non-production flag marks development, test and staging hostnames. Analytics teams usually remove them; security teams often keep them, because test systems reachable from the internet are worth knowing about.

57.7%

of sampled SNIs carry a role other than website

DISTRIBUTION OF NON-WEBSITE ROLES

Share of all sampled SNIs; websites account for the remaining 42.3%.



Sample of 1 billion SNIs

16 ROLES: 13 TRAFFIC ROLES IN THREE GROUPS, AND 3 TECHNICAL CLASSES

The role separates what people visit from what runs in the background. Three technical classes complete the set: ISP subscriber line, server and parked. The SNI Classification Engine uses up to 21 roles when it also runs its online owner checks for new hostnames.

VISITED

- **website** pages people open
- **login** sign-in services
- **email** mail servers and webmail

SUPPORTING

- **cdn** images, video, scripts
- **api** application interfaces
- **push_chat** messaging and push
- **download_update** downloads and updates

BACKGROUND AND INFRASTRUCTURE

- **ads_tracking** ads and analytics
- **telemetry** usage and crash reports
- **remote_access** VPN and remote desktop
- **network** name servers, gateways
- **cloud_hosting** shared storage and hosting
- **certificate** certificate validation

95.8%
with an identified owner domain

42.3%
of sampled SNIs are websites people open

1.1%
flagged as non-production

16
roles: 13 traffic roles and 3 technical classes (ISP subscriber line, server, parked)

07 ANATOMY OF A RECORD

One hostname, classified for analysis and for policy.

An example record from the database: the hostname, its root domain, its IAB classification with confidence, the web-filtering category of its root domain, its traffic role and its owner domain.

ATTRIBUTE	VALUE	WHAT IT HOLDS
SNI	`api.glb.paypal.com`	The hostname exactly as presented in the TLS handshake and in DNS logs; the primary lookup key.
ROOT DOMAIN	`paypal.com`	The registered domain, for aggregation to organisation level.
IAB CATEGORY	Personal Finance > Consumer Banking	Main IAB Content Taxonomy category, written as the full path from tier 1 to the deepest applicable tier.
MULTIPLE CATEGORIES	Personal Finance > Consumer Banking (1.00)	The assigned IAB category with its confidence score between 0 and 1.
WEB-FILTERING CATEGORY	Financial Services	The web-filtering category for policy, delivered in the same record.
ROLE	api	What the server does: website, content delivery, interface, sign-in, mail, advertising, telemetry and more.
OWNER	`paypal.com`	The owner domain of the hostname.

IAB CATEGORY	WEB-FILTERING CATEGORY	ROLE	OWNER DOMAIN
Personal Finance > Consumer Banking	Financial Services	API	`paypal.com`

07 ANATOMY OF A RECORD

Example records from the database.

Each row below is a hostname record, with every attribute shown populated: the IAB category, the web-filtering category of its root domain, the traffic role and the owner domain.

SNI	IAB CATEGORY	WEB-FILTERING CATEGORY	ROLE	OWNER
api.glb.paypal.com	Personal Finance > Consumer Banking	Financial Services	API	paypal.com
smtp-in.orange.fr	Personal Finance > Home Utilities > Internet Service Providers	Email and Messaging | Business	Mail	orange.fr
images.ebay.at	Shopping	Marketplaces / Classifieds / Auctions	CDN	ebay.at
accounts.google.co.in	Technology & Computing > Computing > Internet > Search	Search Engines & Platforms	Login and identity	google.co.in
metrics.barclays.co.uk	Personal Finance > Consumer Banking	Online Banking	Telemetry	barclays.co.uk
googleads.g.doubleclick.net	Business and Finance > Business > Marketing and Advertising	Advertising/Marketing	Advertising and tracking	google.com
remote.lowes.com	Home & Garden > Home Improvement	Home & Garden | Shopping	Remote access	lowes.com

LOOKUP KEY
The SNI, exactly as it appears in traffic and logs

ROOT DOMAIN
For aggregation to organisation level

SCHEMA
The same attributes in every delivery

Platform and service

For hostnames that belong to a known platform, the record also names the service behind them: not only that a hostname belongs to Huawei or Amazon, but that it is Huawei's push service, or Prime Video running on a particular television model. This separates an app's own traffic from the platform services a device uses in the background, such as push notifications, cloud sync, payments and device activation.

HOSTNAME	PLATFORM AND SERVICE	ROLE	OWNER DOMAIN
api.push.hicloud.com	Huawei Push Kit	API	hicloud.com
api21-normal-c-useast1a.tiktokv.com	TikTok app services	API	tiktok.com
api.push.oppomobile.com	OPPO Push	API	oppomobile.com
antv-24-sony-bravia4kgb-505004009.api.amazonvideo.com	Amazon Prime Video app on a Sony Bravia 4K television	API	amazonvideo.com
app-asset.samsungiotcloud.com	Samsung SmartThings IoT cloud	Content delivery	samsungiotcloud.com
p02-remindersws.icloud.com	Apple iCloud services	API	icloud.com
activate.playstation.net	Sony PlayStation Network	API	playstation.net
api-blue.rpay.roku.com	Roku Pay	API	roku.com

08 OPERATIONAL USE

Built for networks where the hostname is the only visible label.

01

Firewalls and web gateways

Write policies by IAB or web-filtering category and by role on encrypted traffic, without breaking it open: block a category, allow its sign-in servers, rate-limit its content delivery.

02

DNS filtering

Every DNS query names a hostname. The database turns those names into categories at the resolver, including the many subdomains a plain domain list misses.

03

Network monitoring and analytics

Report which categories people actually use, with content delivery, tracking and update traffic grouped separately instead of inflating the figures.

04

Security operations

An alert that names only a hostname gains a category, a role and an owner, so an analyst sees at once whether it points to a known service or to something unusual.

05

Internet providers and carriers

Categorise subscriber traffic for usage reporting, network planning and parental controls that block categories without breaking the applications people rely on.

One lookup key, the SNI, connects encrypted traffic to content category, filtering category, function and ownership.

ATTRIBUTES THAT CARRY EACH USE

USE	IAB CATEGORY	FILTERING CATEGORY	ROLE	OWNER
Firewalls and web gateways	•	•	•	
DNS filtering	•	•	•	
Network monitoring and analytics	•		•	•
Security operations	•	•	•	•
Internet providers and carriers	•	•	•	•

09 DELIVERY AND INTEGRATION

Delivered as files. Integrated as a table lookup.

Matching is a plain lookup on an exact key, on the licensee's own servers, with no calls to us, so it keeps pace with the largest log volumes, in batch inside the warehouse or in a stream as logs arrive.

CSV or Parquet

One row per hostname. Parquet loads straight into analytics engines and data warehouses.

Full file or slices

All rows, or slices by top-level domain, country, category or role, delivered by secure download link.

Kept current

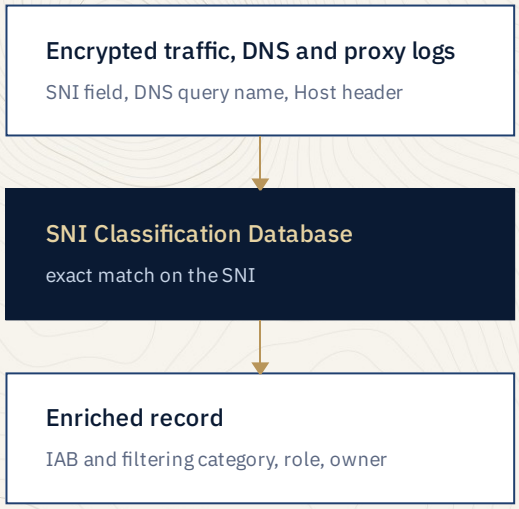
Weekly update files with new and changed hostnames, daily files on request, a full snapshot every quarter.

MATCHING ORDER AND EXAMPLE QUERY

Write the hostname in lower case without a trailing dot or port. Match the exact hostname against the SNI column; otherwise the root domain against our 120 million domain database; anything left is classified by the offline engine included with the database, on your servers.

Example query: one lookup returns the IAB category, web-filtering category, role and owner.

```
SELECT l.*,
  s.category           AS iab_category,
  s.filtering_category,
  s.role,
  s.owner
FROM traffic_log l
LEFT JOIN sni_database s ON s.sni = LOWER(l.hostname);
```



HTTP/3 AND QUIC

QUIC connections carry the server name at the start, so the same lookup applies.

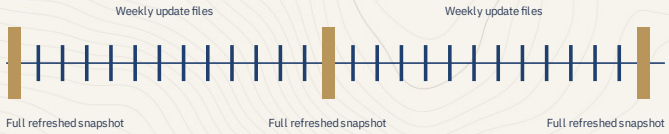
ENCRYPTED CLIENT HELLO

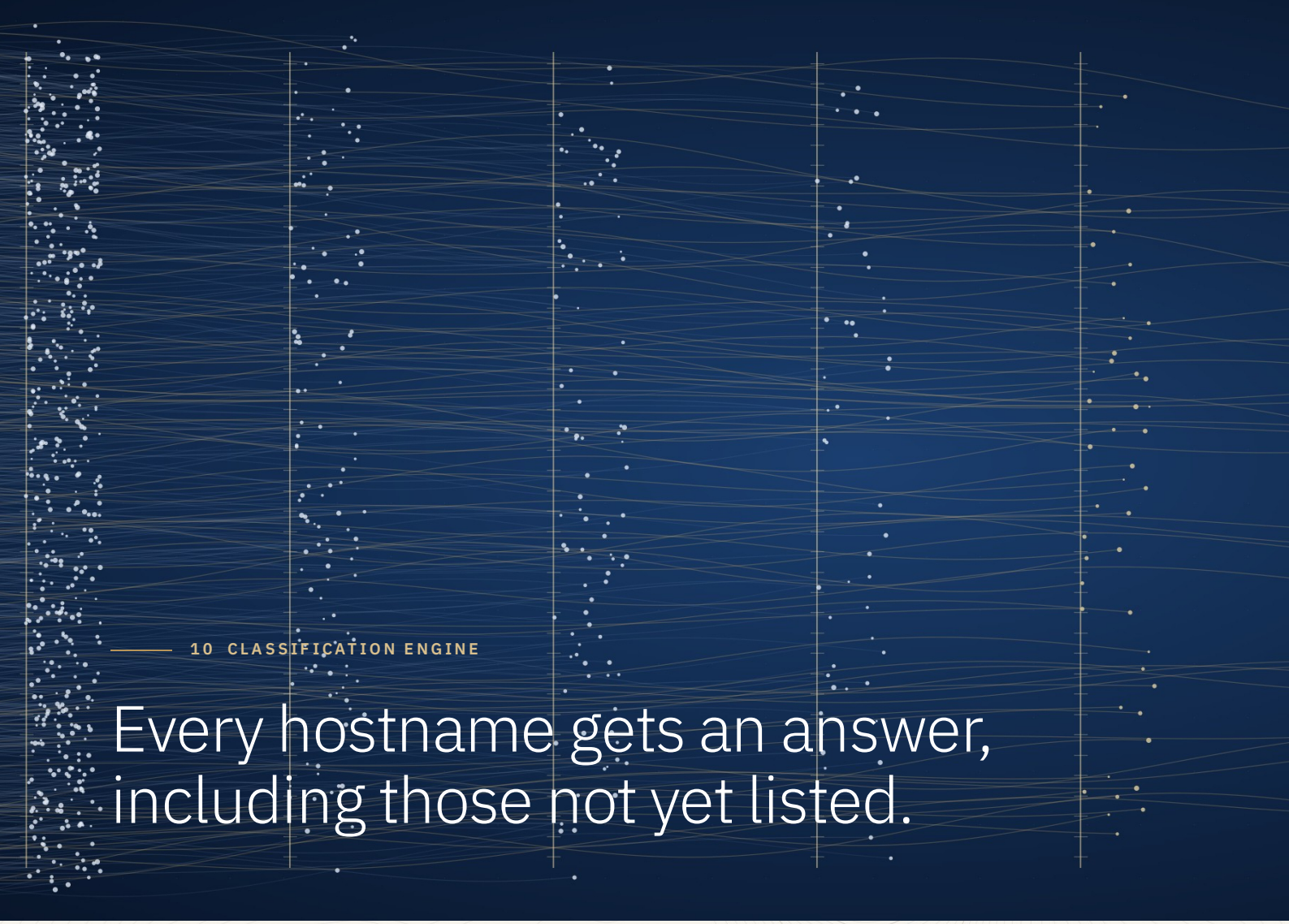
Where the name is hidden, the device still asks for it in DNS. Where the network runs or sees the DNS resolver, as in most company and carrier networks, the DNS query name is matched in the same way.

FRESHNESS

300,000 newly registered domains are reviewed every day.

UPDATE MODEL





10 CLASSIFICATION ENGINE

Every hostname gets an answer, including those not yet listed.

The database is updated every week, but traffic always contains hostnames that appeared after the last update: new campaigns, new subdomains, one-off server names. The SNI Classification Engine classifies them with the same rules that built the database, so every hostname in your data receives a category and a role.

Each engine includes everything of the previous one and adds one step. Engine Offline is included with the database and runs entirely on your servers; the first two engines make no calls to us.

0.93 s

to classify 100,000 SNIs with Engine Offline on a 12-core server, with no network access

Engine Offline

INCLUDED WITH THE DATABASE

A hostname not in the database receives the category of its root domain from the 120 million domain database, a traffic role and an owner. If the root domain is not in that database, role and owner are still assigned and the SNI has no IAB category. No network access at all; suitable for offline or closed environments.

+ Live Owner Checks

FROM YOUR OWN NETWORK

Root domains not in the database are checked through your proxy and resolver: status (live, unresolvable, no address, parked, seized) and the owner where it can be proven. Nothing is sent to us.

+ API Classification

OPTIONAL, UNDER YOUR CONTROL

Registered domains that are still uncategorised can be sent to our classification API, which categorises them from their live website. Only registered domains are sent, never hostnames, users or traffic.

10 CLASSIFICATION ENGINE

Six steps, one answer for every SNI.

THE LOOKUP FLOW

Every SNI is first looked up in the database; only hostnames without an exact match continue, and an answered SNI leaves the flow at once.

1	2	3	4	5	6
Look up the SNI OFFLINE The exact hostname is looked up in the delivered SNI database of 1,880,282,561 hostnames; found rows take category, role and owner directly.	**Root domain** OFFLINE Not found: the registered root domain is looked up in the delivered 120 million domain database and its IAB category applied.	**Traffic role** OFFLINE The rule set assigns the role and owner, also when the root domain is not in the 120 million domain database; the SNI then has no IAB category.	**Live owner checks** YOUR NETWORK Root domain not in the database: status and a proven owner from your own network, through your proxy and resolver.	**Live classification** OPTIONAL API Optional: still uncategorised root domains are categorised by our API from their live website.	**Output** OFFLINE One row per SNI in the database schema, plus a status column.

OPTIMISED ON REAL TRAFFIC

Parked and for-sale domains

Recognised before anything else and reported with their own status.

Owner only with proof

An owner is accepted only when it can be proven, never hallucinated.

Curated reference lists

Parking services, registrars and known companies, refreshed with every weekly update.

Reviewed on real traffic

Checked in rounds on samples of real SNIs and verified on complete hostname lists before each release.

Seized domains

Law-enforcement seizure notices are reported as their own status.

Role before owner

Advertising and tracking hosts are categorised as Marketing and Advertising, whoever operates them.

Built on the full SNI database

The delivered 1,880,282,561 hostnames were classified with exactly these rules.

Built for big batches

Time limits per lookup and per SNI, automatic resume, and a rate limit that respects your network.

Chosen by market leaders across more than 30 sectors.

> Over 300 organisations have used our data and APIs, and more than 50 enterprises have licensed our databases.

REPRESENTATIVE LICENSEES

01 The largest telecommunications company in Europe's second-largest economy

02 A leading global hedge fund, one of the top three globally by industry rankings, managing more than USD 65 billion

03 The largest TV network in Europe's fifth-largest economy

04 One of the largest local news publishers in the United States, whose newspaper websites rank among the most visited local news sites in the country

05 One of the largest advertising technology companies in the United States, whose platform processes more than $200 billion in annualized ad spend

06 A leading global virtualization software corporation, present in 20+ countries

07 A leading US company serving defense and government

Internal use

Analytics, security monitoring and reporting across your teams and systems, with update files for the length of the licence.

In your product

Ship categories and roles inside a firewall, gateway, monitoring tool or data product; the licence covers your customers' use.

Slices

Only the top-level domains, countries, categories or roles you need, extendable later on the same terms.

Licensor

The SNI Classification Database is licensed by Alpha Quantum. Data processing and hosting are located in the European Union.

No personal data

The database lists server names and what those servers do: no users, no IP addresses of people, no browsing histories and no traffic records. Client relationships are treated as confidential.

12 QUESTIONS AND CONTACT

What is an SNI?

The Server Name Indication is the hostname a device sends when it opens an encrypted connection. It is visible to networks even when the rest of the traffic is encrypted.

Can we license only part of it?

Yes. Slices by top-level domain, country, category or role are licensed on their own and can be extended later on the same terms.

How quickly do new hostnames appear?

Weekly update files list added and changed hostnames, daily files on request; between updates, Engine Offline classifies any new hostname as it appears.

Which taxonomy do the categories use?

The IAB Content Taxonomy, the same categories as our 120 million domain database; each record also carries a web-filtering category.

Can we use the data in a product we sell?

Yes. Licences are available for internal use and for products and services you sell to others.

Does it contain personal data?

No. It lists server names and what those servers do: no users, no IP addresses of people, no browsing histories, no traffic records.

SNI Classification Database

To discuss licensing terms, slices, delivery formats or the update model for your organisation, please contact our enterprise data team.

WEB
www.urlcategorizationdatabase.com

EMAIL
info@urlcategorizationdatabase.com

LICENSOR
Alpha Quantum

IAB Tech Lab and the IAB Content Taxonomy are referenced for identification only. All hostnames shown are example records from the database, for illustration; their appearance implies no relationship with the organisations concerned.