Web 3.0 Web-Search Engine

By Henry (Junho)

Table of Contents

1. Current State of the Web (Search Engine)

Current Web Model

- Centralized DNS
- Centralized web indexing and search (Google, DuckDuckGo, etc)
- Centralized certificate authority system (TLS/SSL)

Today's DNS: archaic stuff

List of Root Servers

HOSTNAME	IP ADDRESSES	OPERATOR
a.root-servers.net	198.41.0.4, 2001:503:ba3e::2:30	Verisign, Inc.
b.root-servers.net	170.247.170.2, 2801:1b8:10::b	University of Southern California, Information Sciences Institute
c.root-servers.net	192.33.4.12, 2001:500:2::c	Cogent Communications
d.root-servers.net	199.7.91.13, 2001:500:2d::d	University of Maryland
e.root-servers.net	192.203.230.10, 2001:500:a8::e	NASA (Ames Research Center)
f.root-servers.net	192.5.5.241, 2001:500:2f::f	Internet Systems Consortium, Inc.
g.root-servers.net	192.112.36.4, 2001:500:12::d0d	US Department of Defense (NIC)
h.root-servers.net	198.97.190.53, 2001:500:1::53	US Army (Research Lab)
i.root-servers.net	192.36.148.17, 2001:7fe::53	Netnod
j.root-servers.net	192.58.128.30, 2001:503:c27::2:30	Verisign, Inc.
k.root-servers.net	193.0.14.129, 2001:7fd::1	RIPE NCC
l.root-servers.net	199.7.83.42, 2001:500:9f::42	ICANN
m.root-servers.net	202.12.27.33, 2001:dc3::35	WIDE Project

https://www.iana.org/domains/root/servers

DNS


DEPARTMENT OF JUSTICE
NATIONAL INTELLECTUAL PROPERTY RIGHTS
IPR CENTER
COORDINATION CENTER
HOMELAND SECURITY INVESTIGATIONS
SPECIAL AGENT
This domain name has been seized by ICE - Homeland Security Investigations, pursuant to a seizure warrant issued by a United States District Court under the authority of 18 U.S.C. §§ 981 and 2323.
Willful copyright infringement is a federal crime that carries penalties for first time offenders of up to five years in federal prison, a $250,000 fine, forfeiture and restitution (17 U.S.C § 506, 18 U.S.C. § 2319). Intentionally and knowingly trafficking in counterfeit goods is a federal crime that carries penalties for first time offenders of up to ten years in federal prison, a $2,000,000 fine, forfeiture and restitution (18 U.S.C. § 2320).

DNS

Domain Name System

- The very last section of the domain is called its **top-level domain (TLD)** name

Top-Level Domain	General Purpose	New TLDs	General Purpose
.com	U.S. Commercial	.biz	Business
.net	Network	.info	Information
.org	Nonprofit organization	.pro	Professional
.edu	U.S. Educational	.museum	Museums
.int	International	.aero	Aerospace industry
.mil	U.S. Military	.coop	Cooperative
.gov	U.S. Government		

Figure 15.10 **Top-level domains, including some relatively new ones**

Certificate Authorities

- Examples:
  - DigiCert
  - GlobalSign
  - Entrust
  - Comodo
  - GoDaddy
  - Let’s Encrypt
  - Trustwave

C=ES,O=ACCV,OU=PKIACCV,CN=ACCVRAIZ1
OU=AC RAIZ FNMT-RCM,O=FNMT-RCM,C=ES
CN=Actalis Authentication Root CA,O=Actalis S.p.A./03358520967,L=Milan,C=IT
CN=AffirmTrust Commercial,O=AffirmTrust,C=US
CN=AffirmTrust Networking,O=AffirmTrust,C=US
CN=AffirmTrust Premium,O=AffirmTrust,C=US
CN=AffirmTrust Premium ECC,O=AffirmTrust,C=US
CN=Amazon Root CA 1,O=Amazon,C=US
CN=Amazon Root CA 2,O=Amazon,C=US
CN=Amazon Root CA 3,O=Amazon,C=US
CN=Amazon Root CA 4,O=Amazon,C=US
C=DE,O=Atos,CN=Atos TrustedRoot 2011
CN=Autoridad de Certificacion Firmaprofesional CIF A62634068,C=ES
CN=Baltimore CyberTrust Root,OU=CyberTrust,O=Baltimore,C=IE
CN=Buypass Class 2 Root CA,O=Buypass AS-983163327,C=NO
CN=Buypass Class 3 Root CA,O=Buypass AS-983163327,C=NO
CN=CA Disig Root R2,O=Disig a.s.,L=Bratislava,C=SK
CN=CFCA EV ROOT,O=China Financial Certification Authority,C=CN
CN=COMODO Certification Authority,O=COMODO CA Limited,L=Salford,ST=Greater Manchester,C=GB
CN=COMODO ECC Certification Authority,O=COMODO CA Limited,L=Salford,ST=Greater Manchester,C=GB
CN=COMODO RSA Certification Authority,O=COMODO CA Limited,L=Salford,ST=Greater Manchester,C=GB
CN=Chambers of Commerce Root,OU=http://www.chambersign.org,O=AC Camerfirma SA CIF A82743287,C=EU
CN=Global Chambersign Root,OU=http://www.chambersign.org,O=AC Camerfirma SA CIF A82743287,C=EU
CN=Certigna,O=Dhimyotis,C=FR
CN=Certigna Root CA,OU=0002 48146308100036,O=Dhimyotis,C=FR
CN=Certum CA,O=Unizeto Sp. z o.o.,C=PL
CN=Certum Trusted Network CA,OU=Certum Certification Authority,O=Unizeto Technologies S.A.,C=PL
CN=Certum Trusted Network CA 2,OU=Certum Certification Authority,O=Unizeto Technologies S.A.,C=PL
CN=Chambers of Commerce Root - 2008,O=AC Camerfirma S.A.,serialNumber=A82743287,L=Madrid (see current address at www.camerfirma.com/address),C=EU
CN=AAA Certificate Services,O=Comodo CA Limited,L=Salford,ST=Greater Manchester,C=GB
CN=Cybertrust Global Root,O=Cybertrust\, Inc
CN=D-TRUST Root CA 3 2013,O=D-Trust GmbH,C=DE

https://www.checktls.com/showcas.html

Google

- Let’s investigate “Bitcoin” search results...

bitcoin

News | Images | Price | Perspectives | Login | USD | Videos | Predictions | Chart | All filters | Tools

About 658,000,000 results (0.47 seconds)

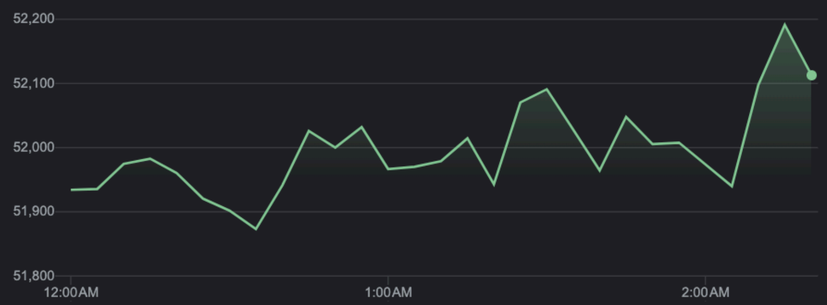
Market Summary > Bitcoin

52,191.40 USD

+257.20 (0.50%) ↑ today

Feb 16, 2:14 AM UTC · Disclaimer

1D | 5D | 1M | 6M | YTD | 1Y | 5Y | Max



1 BTC = 52191.40 USD

Feedback

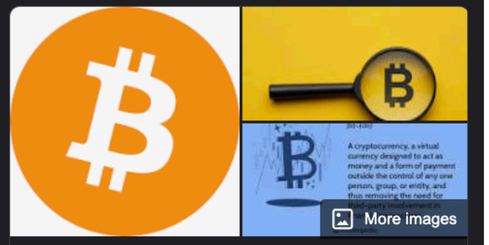
More about Bitcoin →

Bitcoin.org
https://bitcoin.org

Bitcoin - Open source P2P money

Bitcoin is an innovative payment network and a new kind of money. Find all you need to know and get started with **Bitcoin** on **bitcoin**.org.

FAQ · Introduction · Resources · Innovation



Bitcoin

Cryptocurrency

Bitcoin is the first decentralized cryptocurrency. Nodes in the peer-to-peer bitcoin network verify transactions through cryptography and record them in a public distributed ledger, called a blockchain, without central oversight. Wikipedia

Symbols: BTC, ฿, Ƀ

Official users: El Salvador

Block reward: ฿6.25 (as of 2023)

Code: BTC, XBT

Exchange rate: Floating

Hash function: SHA-256 (two rounds)

Implementation(s): Bitcoin Core

People also search for — View 10+ more

Ethereum | Dogecoin | Ripple | Cardano

Bitcoin.com
https://www.bitcoin.com

Buy Bitcoin & cryptocurrency | Wallet, news, education.

The world's gateway to **Bitcoin** & cryptocurrency. Buy, sell, spend, swap, and invest in BTC, ETH, BCH, AVAX, MATIC & hundreds more digital assets.

CoinDesk
https://www.coindesk.com › price › bitcoin

Bitcoin Price | BTC Price Index and Live Chart

Bitcoin USD price, real-time (live) charts, news and videos. Learn about BTC value, **bitcoin** cryptocurrency, crypto trading, and more.

Top stories

Coinbase's quarterly earnings ›

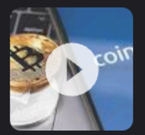
Yahoo Finance

Coinbase upgraded to Neutral while bitcoin rises above $52K

11 hours ago

Investor's Business Daily

Coinbase Spikes After Earnings, Revenue Blast Past Forecasts



4 hours ago

Decrypt

News Explorer — Coinbase Posts Huge $273.4 Million Profit in Q4 2023 as Bitcoi…

4 hours ago

Elizabeth Warren's Satoshi Nakamoto flag

CoinDesk

Satoshi-Era Bitcoin Function 'OP_CAT' Dusted Off as Development Fervor Grows

1 day ago

Bitcoin Magazine

Elizabeth Warren Pivots on Bitcoin, Honors Nakamoto With Flag Over Capitol

11 hours ago

Google

1. bitcoin.org
2. Bitcoin.com
3. CoinDesk
4. Wikipedia
5. @Bitcoin Twitter handle
6. CoinMarketCap
7. Cointelegraph
8. Yahoo
9. The Block
10. CNBC

Google

1. bitcoin.org
2. ~~Bitcoin.com~~
3. ~~CoinDesk~~
4. Wikipedia
5. @Bitcoin Twitter handle
6. ~~CoinMarketCap~~
7. ~~Cointelegraph~~
8. ~~Yahoo~~
9. ~~The Block~~
10. ~~CNBC~~

What about...

- Whitepaper
- Bitcoin subreddit
- Bitcoin Core repo
- Self-custody (ex: hardware wallets)
- Nakamoto Institute
- Gigi's blog
- All the great things my network finds useful or has built

Oh, and...

- I use an ad blocker

Ranking Web Content

- Relevance
- Authoritative
- Page Experience (Ex: page loading speed)
- Backlinks
- User engagement
- Freshness

Costs

- “A dollar saved is a dollar earned”
- Revenue:
  - Google: $305B
  - Verisign: $5.9B
  - GoDaddy: $4.3B
  - Entrust: ~$750M
  - Cogent: ~$560M
  - Digicert: $420M
  - Trustwave: $395.9M
  - Comodo Cybersecurity: $56M
  - GlobalSign: ~$50M

This...

List of Root Servers

HOSTNAME	IP ADDRESSES	OPERATOR
a.root-servers.net	198.41.0.4, 2001:503:ba3e::2:30	Verisign, Inc.
b.root-servers.net	170.247.170.2, 2801:1b8:10::b	University of Southern California, Information Sciences Institute
c.root-servers.net	192.33.4.12, 2001:500:2::c	Cogent Communications
d.root-servers.net	199.7.91.13, 2001:500:2d::d	University of Maryland
e.root-servers.net	192.203.230.10, 2001:500:a8::e	NASA (Ames Research Center)
f.root-servers.net	192.5.5.241, 2001:500:2f::f	Internet Systems Consortium, Inc.
g.root-servers.net	192.112.36.4, 2001:500:12::d0d	US Department of Defense (NIC)
h.root-servers.net	198.97.190.53, 2001:500:1::53	US Army (Research Lab)
i.root-servers.net	192.36.148.17, 2001:7fe::53	Netnod
j.root-servers.net	192.58.128.30, 2001:503:c27::2:30	Verisign, Inc.
k.root-servers.net	193.0.14.129, 2001:7fd::1	RIPE NCC
l.root-servers.net	199.7.83.42, 2001:500:9f::42	ICANN
m.root-servers.net	202.12.27.33, 2001:dc3::35	WIDE Project

Domain Name System

- The very last section of the domain is called its **top-level domain (TLD)** name

Top-Level Domain	General Purpose	New TLDs	General Purpose
.com	U.S. Commercial	.biz	Business
.net	Network	.info	Information
.org	Nonprofit organization	.pro	Professional
.edu	U.S. Educational	.museum	Museums
.int	International	.aero	Aerospace industry
.mil	U.S. Military	.coop	Cooperative
.gov	U.S. Government		

Figure 15.10 **Top-level domains, including some relatively new ones** 15-29

•••

Certificate Authorities

```
C=ES,O=ACCV,OU=PKIACCV,CN=ACCVRAIZ1
OU=AC RAIZ FNMT-RCM,O=FNMT-RCM,C=ES
CN=Actalis Authentication Root CA,O=Actalis S.p.A./03358520967,L=Milan,C=IT
CN=AffirmTrust Commercial,O=AffirmTrust,C=US
CN=AffirmTrust Networking,O=AffirmTrust,C=US
CN=AffirmTrust Premium,O=AffirmTrust,C=US
CN=AffirmTrust Premium ECC,O=AffirmTrust,C=US
CN=Amazon Root CA 1,O=Amazon,C=US
CN=Amazon Root CA 2,O=Amazon,C=US
CN=Amazon Root CA 3,O=Amazon,C=US
CN=Amazon Root CA 4,O=Amazon,C=US
C=DE,O=Atos,CN=Atos TrustedRoot 2011
CN=Autoridad de Certificacion Firmaprofesional CIF A62634068,C=ES
CN=Baltimore CyberTrust Root,OU=CyberTrust,O=Baltimore,C=IE
CN=Buypass Class 2 Root CA,O=Buypass AS-983163327,C=NO
CN=Buypass Class 3 Root CA,O=Buypass AS-983163327,C=NO
CN=CA Disig Root R2,O=Disig a.s.,L=Bratislava,C=SK
CN=CFCA EV ROOT,O=China Financial Certification Authority,C=CN
CN=COMODO Certification Authority,O=COMODO CA Limited,L=Salford,ST=Greater Manchester,C=GB
CN=COMODO ECC Certification Authority,O=COMODO CA Limited,L=Salford,ST=Greater Manchester,C=GB
CN=COMODO RSA Certification Authority,O=COMODO CA Limited,L=Salford,ST=Greater Manchester,C=GB
CN=Chambers of Commerce Root,OU=http://www.chambersign.org,O=AC Camerfirma SA CIF A82743287,C=EU
CN=Global Chambersign Root,OU=http://www.chambersign.org,O=AC Camerfirma SA CIF A82743287,C=EU
CN=Certigna,O=Dhimyotis,C=FR
CN=Certigna Root CA,OU=0002 48146308100036,O=Dhimyotis,C=FR
CN=Certum CA,O=Unizeto Sp. z o.o.,C=PL
CN=Certum Trusted Network CA,OU=Certum Certification Authority,O=Unizeto Technologies S.A.,C=PL
CN=Certum Trusted Network CA 2,OU=Certum Certification Authority,O=Unizeto Technologies S.A.,C=PL
CN=Chambers of Commerce Root - 2008,O=AC Camerfirma S.A.,serialNumber=A82743287,L=Madrid (see current
CN=AAA Certificate Services,O=Comodo CA Limited,L=Salford,ST=Greater Manchester,C=GB
CN=Cybertrust Global Root,O=Cybertrust\, Inc
CN=D-TRUST Root CA 3 2013,O=D-Trust GmbH,C=DE
```

Google Web Ranking Criteria

- Relevance
- Authoritative
- Page Experience (Ex: page loading speed)
- Backlinks
- User engagement
- Freshness

Reminds me of this...

FEDERAL RESERVE press release

For release at 2 p.m. EDT September 21, 2022

The Committee seeks to achieve maximum employment and inflation at the rate of 2 percent over the longer run. In support of these goals, the Committee decided to raise the target range for the federal funds rate to 3 to 3-1/4 percent and anticipates that ongoing increases in the target range will be appropriate. In addition, the

And this...


Google


ENTRUST


COMODO
CYBERSECURITY


GlobalSign®
by GMO


Trustwave®


GoDaddy™


cogent


VERISIGN™


digicert®

Reminds me of this.

Summary: Current Web

- Centralized
- Custodial
- Rent-seekers
- Centrally planned / astroturfed
- Surveillance

2. Envisioning a better web (Search Engine)

A better web, in words...

- No rent seekers
- Non-custodial
- Ranked by users (Reddit-style) and more specifically, web-of-trust
- Decentralized

Let’s work backwards…

- Let’s envision a beautiful web (search engine) of the future

The web search engine: Context

- Lots of web apps
- Some use Nostr, some don’t
- DNS exists; still (partially) used
- Nostr has taken off
- Traditional web search engines obsolete for Nostr users
  - Instead, first step into web search/discovery is through our search engine

The web search engine: Primitives

- Users can generate metadata on websites (ex: liking, commenting, zapping)
- Users/groups can host/point to websites on Nostr, which attaches the website to their Nostr identity
- Users can follow other users

The web search engine: Ranking Criteria

- Through X network hops in their web-of-trust:
  - Engagement (ex: likes, comments, zaps)
  - Websites directly published/hosted on Nostr
- Also (obviously): Content relevance

Re: The Bitcoin example

- Whitepaper
- Bitcoin subreddit
- Bitcoin Core repo
- Self-custody (ex: hardware wallets)
- Nakamoto Institute
- Gigi's blog
- All the great things my network finds useful or has built

Circling back - Benefits

- Better web search
- Remove web hosting / search rent seekers
- Non-custodial domains
- Decentralized web hosting

3. What needs to be built

Basic primitives

- DNS on Nostr
- Web hosting on Nostr
- A website or browser that supports accessing websites via Nostr
- **WoT based web indexing/search system**
- Potentially, distributed server/relay set to serve websites

4. Monetization

Monetization

- Growth first!
- Sell a service:
  - Web hosting service (geo-distributed relay network)
  - Web index / searching service
- Ads...?
- Q: Do zap splits and subscriptions make sense for Nostr / open-source software?
  - **Compute is scarce, open-source code is not**

5. Growth strategy

Growth strategy

- Leave Nostr as an implementation detail
- Users see:
  - Account
  - Web hosting
  - Search engine
- Bootstrap with:
  - Other social graphs
  - Create pointers to Web 2.0
  - Google API

6. Objections

What do you do that no one else does?

- Q: Kind of like a central banker saying “why would anyone do all the work to create a new currency if they can’t create more of it?”
  - A: You can get more of it at the start, get a headstart, and now we have bitcoin
- This idea of trying to create a monopoly on information, it will soon be a very outdated mindset, just like only creating a money if you can have a monopoly on printing it is an outdated fiat mindset
- Focus on providing a service
- Q: “Why would anyone do all the work to create nostr web hosting when they’re not going to have a monopoly?”
  - A: Because you can get a headstart and build out a useful service and earn a profit

Brainstorming

- Why would anyone host their website on your slow dinky app
- Why is replacing DNS, web discovery the crux of moving to a better web? Like why can't we just build Youtube on Nostr, Twitter on Nostr, etc?
- You said, DNS and web discovery are centralized and all that shit but isn't every other internet app like Netflix, Twitch, etc the same?
- What do you offer that makes you different from everyone else?
- Why will you outcompete your competitors?
- What is your 10x improvement?
- Don't LLMs obsolete web search
- How do you moderate content, for example users posting obscene or vulgar stuff
- How do you prevent illegal websites
- Why do you need to host websites on Nostr to build this? Can't you just build the same thing but host websites, including cool Nostr apps, using normal DNS?
- Wouldn't the best way for someone who wants to use your search engine to host their own search engine? That'll scan their web of trust and generate indexes?
- How will you build out the same content relevance functionality as Google? Particularly for complicated search queries
- How do people with large follower lists filter out spammy/low quality stuff from their web of trust?
- How do you compete with Google who has all this extra compute

Fin