SEO Sydney

- News
- SEO Sydney
- Local SEO Sydney
- SEO services Sydney
- search engine optimisation consultants
- More
  local SEO services SydneySEO agencies in SydneySEO service in Sydney SEO services in SydneySEO parramattaSEO consultant SydneySydney SEO consultantSydney SEO consultingkeyword research servicesSEO specialists SydneySEO expert Sydneysearch engine optimisation Sydney local SEO SydneySEO experts SydneySEO packages australiaSEO services expertwhat SEO marketingSEO meaningSEO service SydneySEO agencies SydneySEO agency australiaLocal SEOSEO australiaSEO expertdigital agency SydneySydney SEO consultantlocal SEO specialistsSEO strategy SEO in marketingcontent marketing SydneySEO packagesSEO parramatta SEO Sydney expertSEO Sydney expertsSEO specialistSEO for websiteSEO googleSydney SEO expertsSEO package australiaSEO consultants Sydney expert SEO servicesSEO marketingSEO checkSEO packages SydneySEO keywordsSEO websitelocal SEO australiaSEO consultantSEO package SydneySEO services in SydneySEO companies in australialocal SEO agencyecommerce SEO servicesSEO specialists Sydneybest SEO company in Sydneycontent agency Sydneybest SEO agency SydneySEO agency in SydneySEO company SydneySEO agencies SydneySEO company in SydneySEO company SydneySEO expertsSEO agency Sydneybest SEO SydneySEO agency in SydneySEO services expertSEO agencies in Sydney listing business on googlebest SEO company SydneySEO service Sydney SEO services Sydneysearch engine optimisation Sydneylocal SEO services SEO services providerSydney SEO companySEO company in SydneySEO agency SydneySEO with wordpressSEO consultant SydneySEO expert SydneySydney SEO servicesSEO services company SydneySydney SEO consultingSEO services companySEO servicesSydney SEO expertSEO experts SydneySEO agency australiagoogle listing for businesssearch engine optimisation strategySEO agency
- About Us
- Contact Us

SEO companies in australia

keyword competition

keyword competition

Local SEO specialists"Local SEO specialists focus on optimizing a businesss online presence within a specific region. By targeting local keywords, managing directory listings, and creating location-specific content, these specialists help businesses attract more local customers and improve their community reputation."

Local SEO Sydney"Local SEO services in Sydney focus on optimizing a businesss online presence to attract customers in a specific geographical area. Search Engine Optimisation . By leveraging strategies such as Google My Business optimization, local keyword targeting, and local link building, businesses can dominate local search results, increase foot traffic, and build strong community connections."

Local SEO Sydney"Local SEO in Sydney targets geographically relevant search terms to connect businesses with nearby customers. By optimizing local directories, managing online reviews, and creating location-specific content, these strategies increase visibility and attract more foot traffic to brick-and-mortar stores."

Best SEO Sydney Agency.

keyword density —

- keyword competition
- keyword density
- keyword density analysis
- keyword difficulty analysis
- keyword intent analysis
- keyword intent analysis
- keyword mapping

long-form content keywordsLong-form content keywords support in-depth articles that thoroughly address a topic. These keywords help you capture search traffic from users seeking detailed information and enhance your contents authority.

long-form content optimization"Long-form content optimization involves refining detailed, in-depth articles to improve search visibility and user engagement. By incorporating relevant keywords, structuring content clearly, and adding multimedia elements, businesses can rank higher and provide more value to readers."

long-tail keywords"Long-tail keywords are more specific, less competitive search terms that often have higher conversion rates. By targeting these keywords, businesses can reach a more focused audience, improve rankings, and attract highly qualified traffic."

Best Local SEO Services.

keyword density analysis

long-tail keywords"Long-tail keywords are more specific, less competitive phrases that often yield higher conversion rates. Best SEO Packages Sydney Sydney. These terms attract a more targeted audience, making it easier to rank well and generate quality traffic."

low-competition keywordsLow-competition keywords are easier to rank for because fewer websites target them. Focusing on these terms can help smaller sites gain visibility without needing an extensive backlink profile.

low-competition long-tail keywordsLow-competition long-tail keywords are detailed phrases that are easier to rank for due to limited competition.

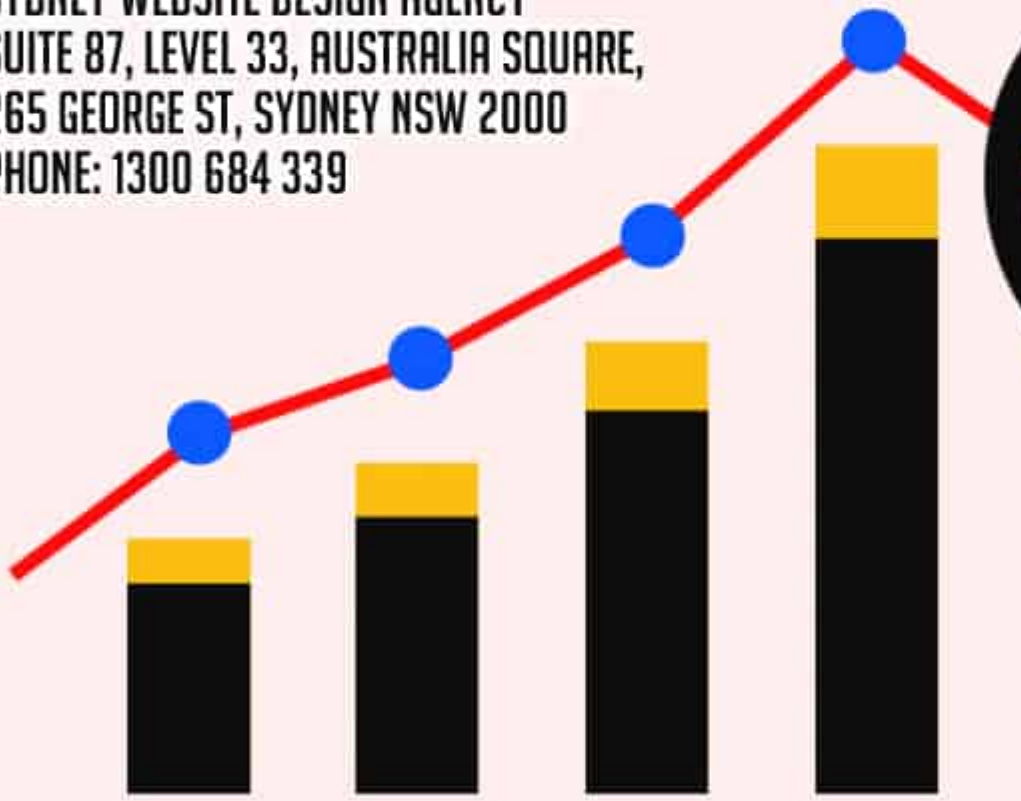
SEO companies in australia - Google algorithm

1. Google site audits
2. Google ranking signals

These keywords help you gain visibility and attract targeted traffic without extensive SEO resources.

HOW SEARCH ENGINE MAR
HELPS BUSINESS GROW OR

SYDNEY WEBSITE DESIGN AGENCY
SUITE 87, LEVEL 33, AUSTRALIA SQUARE,
265 GEORGE ST, SYDNEY NSW 2000
PHONE: 1300 684 339



keyword difficulty analysis

LSI keywords"Latent Semantic Indexing (LSI) keywords are closely related terms that help search engines understand context. By including these keywords, you enhance content relevance and improve search rankings."

market-specific keywordsMarket-specific keywords focus on the unique terms used within a particular industry. Targeting these keywords helps you appeal directly to your niche audience and improve relevancy.

meta description enhancement"Improving meta descriptions makes them more descriptive, engaging, and keyword-rich. SEO Services . A well-crafted meta description helps attract clicks, provides a clear summary of the content, and signals relevance to search engines."

keyword intent analysis

meta description optimization"Meta description optimization involves writing compelling, concise summaries of page content that include relevant keywords. A well-crafted meta description improves click-through rates and helps search engines understand the pages topic."

meta tags"Meta tags provide important information about a web pages content to search engines and users. By crafting compelling meta titles and descriptions, businesses can improve click-through rates and enhance their websites overall visibility in search results."

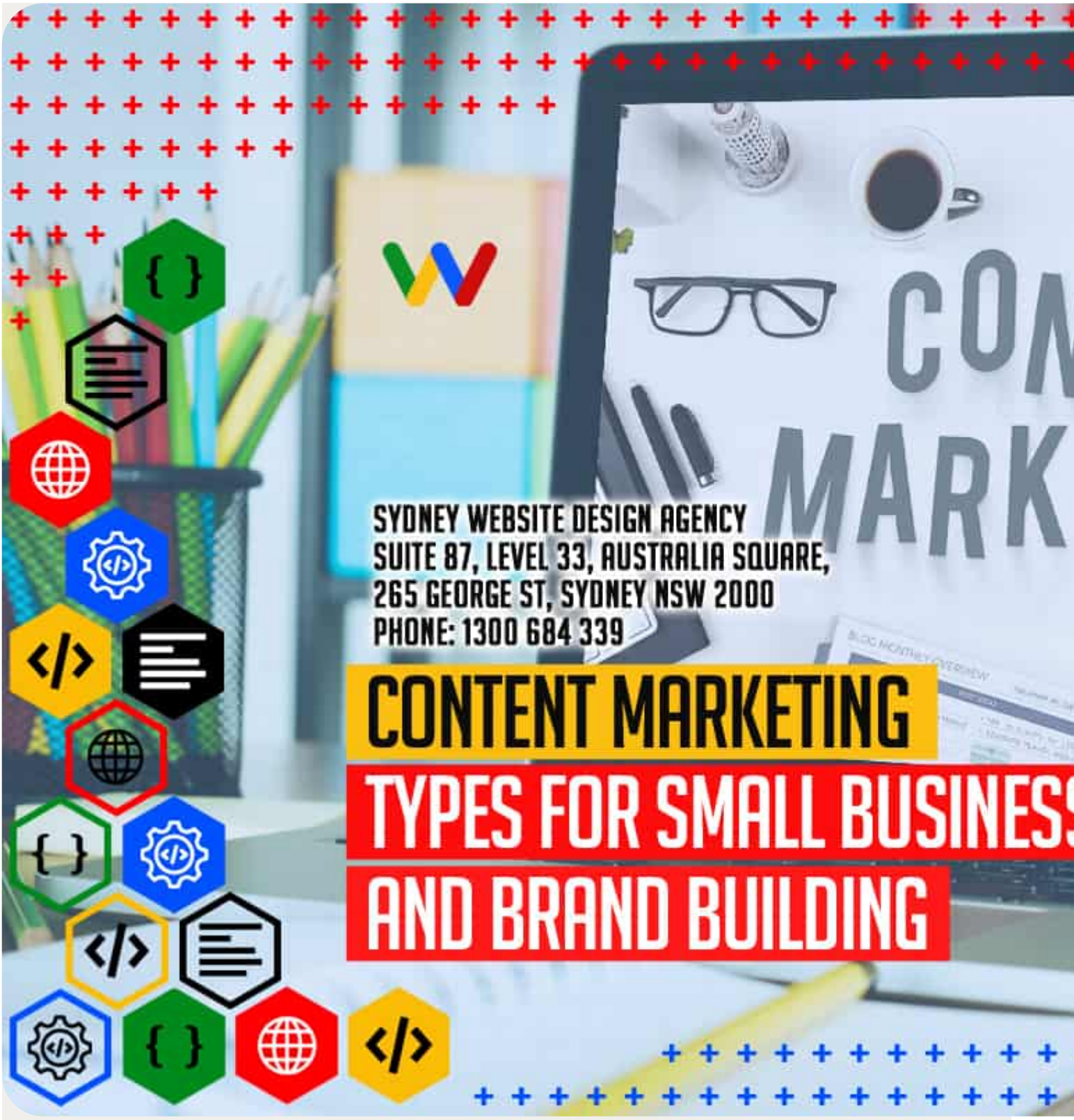
meta tags optimization"Meta tags optimization involves crafting descriptive, keyword-rich title tags and meta descriptions that accurately represent the page content.

SEO companies in australia - Search engine indexing guidelines

- Google Knowledge Panel entries
- Google search visibility

By optimizing these elements, you improve click-through rates and help search engines understand what the page is about, resulting in better rankings."


KEY ADVANTAGE
LOCAL SEO S
28



keyword intent analysis

mobile usability enhancements"Mobile usability enhancements ensure that your website is fully functional and user-friendly on mobile devices. Improvements like responsive design, fast load times, and easy navigation reduce bounce rates and improve rankings in mobile search results."

mobile-first design"Mobile-first design prioritizes the user experience on smartphones and tablets. By designing your site with mobile users in mind, you align with search engines mobile-first indexing guidelines and improve overall rankings."

mobile-first indexing"Mobile-first indexing means that search engines primarily use the mobile version of a websites content for ranking and indexing. By ensuring a responsive design and fast mobile load times, businesses can maintain strong rankings and provide a seamless user experience."

keyword mapping

mobile-friendly content"Mobile-friendly content is designed to be easily accessible and readable on smartphones and tablets. Optimizing for mobile improves user experience, reduces bounce rates, and aligns with search engines mobile-first indexing policies."

mobile-friendly design"Mobile-friendly design ensures that web pages are responsive and easy to navigate on smartphones and tablets. A mobile-optimized site improves user experience, reduces bounce rates, and aligns with search engines preference for mobile-first indexing."

mobile-friendly images"Mobile-friendly images are optimized to display correctly on smartphones and tablets.

SEO companies in australia - Google algorithm

1. Search intent signals
2. Search engine indexing guidelines
3. Google algorithm

By ensuring that images are responsive and load quickly on mobile devices, you improve user satisfaction and help your site perform well in mobile search rankings."

About Web indexing

Image not found or type unknown This article includes a list of general references, but **it lacks sufficient corresponding inline citations**. Please help to improve this article by introducing more precise citations. *(December 2014) (Learn how and when to remove this message)*

Web indexing, or **Internet indexing**, comprises methods for indexing the contents of a website or of the Internet as a whole. Individual websites or intranets may use a back-of-the-book index, while search engines usually use keywords and metadata to provide a more useful vocabulary for Internet or onsite searching. With the increase in the number of periodicals that have articles online, web indexing is also becoming important for periodical websites.[1]

Back-of-the-book-style web indexes may be called "web site A-Z indexes".[2] The implication with "A-Z" is that there is an alphabetical browse view or interface. This interface differs from that of a browse through layers of hierarchical categories (also known as a taxonomy) which are not necessarily alphabetical, but are also found on some web sites. Although an A-Z index could be used to index multiple sites, rather than the multiple pages of a single site, this is unusual.

Metadata web indexing involves assigning keywords, description or phrases to web pages or web sites within a metadata tag (or "meta-tag") field, so that the web page or web site can be retrieved with a list. This method is commonly used by search engine indexing.[3]

See also

[edit]

- Automatic indexing
- Information architecture
- Search engine optimization
- On-page Optimization
- Google Webmaster
- Site map
- Web navigation
- Web search engine
- Information retrieval

Further reading

[edit]

- *Beyond Book Indexing: How to Get Started in Web Indexing, Embedded Indexing, and Other Computer-Based Media*, edited by Marilyn Rowland and Diane Brenner, American Society of Indexers, Info Today, Inc, NJ, 2000, ISBN 1-57387-081-1
- An example of an Internet Index A-Z

- **v**
- **t**
- **e**

Internet search

Types

- Web search engine (List)
- Metasearch engine
- Multimedia search
- Collaborative search engine
- Cross-language search
- Local search
- Vertical search
- Social search
- Image search
- Audio search
- Video search engine
- Enterprise search
- Semantic search
- Natural language search engine
- Voice search

Tools	○ Cross-language information retrieval ○ Search by sound ○ Search engine marketing ○ Search engine optimization ○ Evaluation measures ○ Search oriented architecture ○ Selection-based search ○ Document retrieval ○ Text mining ○ Web crawler ○ Multisearch ○ Federated search ○ Search aggregator ○ Index/Web indexing ○ Focused crawler ○ Spider trap ○ Robots exclusion standard ○ Distributed web crawling ○ Web archiving ○ Website mirroring software ○ Web query ○ Web query classification
Protocols and standards	○ Z39.50 ○ Search/Retrieve Web Service ○ Search/Retrieve via URL ○ OpenSearch ○ Representational State Transfer ○ Wide area information server
See also	○ Search engine ○ Desktop search ○ Online search

References

[edit]

1. ^ *"Web Crawlers:Indexing the Web"*.
2. ^ *Kundu, Malay Kumar; Mohapatra, Durga Prasad; Konar, Amit; Chakraborty, Aruna (2014-05-26). Advanced Computing, Networking and Informatics- Volume 1: Advanced*

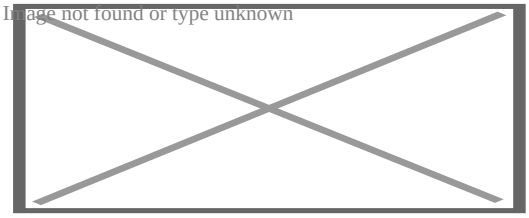
Computing and Informatics Proceedings of the Second International Conference on Advanced Computing, Networking and Informatics (ICACNI-2014). Springer. *ISBN 9783319073538*.
3. ^ *"Indexing the Web | American Society for Indexing"*. *www.asindexing.org. Retrieved 2015-11-25.*

4. What is Website Indexing?

Stub icon Image not found or type unknown This Internet-related article is a stub. You can help Wikipedia by expanding it.

- **v**
- **t**
- **e**

About Semantic Web



A tag cloud (a typical Web 3.0 phenomenon in itself) presenting Web 3.0 themes

- **v**
- **t**
- **e**

Semantics

- **Linguistic**
- **Logical**

Subfields

- Computational
- Lexical (lexis, lexicology)
- Statistical
- Structural

Topics

- Analysis
- Compositionality
- Context
- - Prototype theory
  - Force dynamics
- Semantic feature
- Semantic gap
- Theory of descriptions

Analysis

- Latent
- Computational
- Machine-learning

Applications

- Semantic file system
- Semantic desktop
- Semantic matching
- Semantic parsing
- Semantic similarity
- Semantic query
- - Semantic Web
  - Semantic wiki

Semantics of programming languages

Types

- Action
- Algebraic
- Axiomatic
- Categorical
- Concurrency
- Denotational
- Game
- Operational
- Predicate transformational

Theory

- Abstract interpretation
- Abstract semantic graph

- Language
- Linguistics

The **Semantic Web**, sometimes known as **Web 3.0** (not to be confused with Web3), is an extension of the World Wide Web through standards[1] set by the World Wide Web Consortium (W3C). The goal of the Semantic Web is to make Internet data machine-readable.

To enable the encoding of semantics with the data, technologies such as Resource Description Framework (RDF)[2] and Web Ontology Language (OWL)[3] are used. These technologies are used to formally represent metadata. For example, ontology can describe concepts, relationships between entities, and categories of things. These embedded semantics offer significant advantages such as reasoning over data and operating with heterogeneous data sources.[4] These standards promote common data formats and exchange protocols on the Web, fundamentally the RDF. According to the W3C, "The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries."[5] The Semantic Web is therefore regarded as an integrator across different content and information applications and systems.

History

[edit]

The term was coined by Tim Berners-Lee for a web of data (or **data web**)[6] that can be processed by machines[7]—that is, one in which much of the meaning is machine-readable. While its critics have questioned its feasibility, proponents argue that applications in library and information science, industry, biology and human sciences research have already proven the validity of the original concept.[8]

Berners-Lee originally expressed his vision of the Semantic Web in 1999 as follows:

> I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web – the content, links, and transactions between people and computers. A "Semantic Web", which makes this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines. The "intelligent agents" people have touted for ages will finally materialize.[9]

The 2001 *Scientific American* article by Berners-Lee, Hendler, and Lassila described an expected evolution of the existing Web to a Semantic Web.[10] In 2006, Berners-Lee and colleagues stated that: "This simple idea…remains largely unrealized".[11] In 2013, more than four million Web domains (out of roughly 250 million total) contained Semantic Web markup.[12]

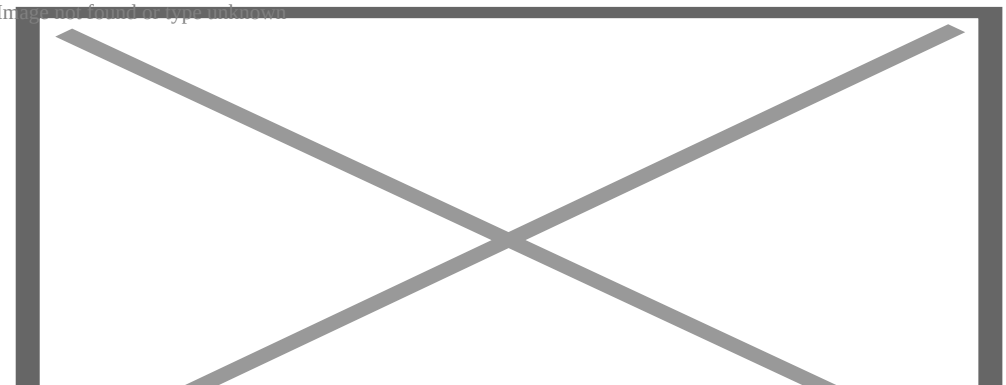
Example

[edit]

In the following example, the text "Paul Schuster was born in Dresden" on a website will be annotated, connecting a person with their place of birth. The following HTML fragment shows how a small graph is being described, in RDFa-syntax using a schema.org vocabulary and a Wikidata ID:

```
<div vocab="https://schema.org/" typeof="Person">
  <span property="name">Paul Schuster</span> was born in
  <span property="birthPlace" typeof="Place" href="https://www.wikidata.org/entity/Q1731">
    <span property="name">Dresden</span>.
  </span>
</div>
```

Image not found or type unknown



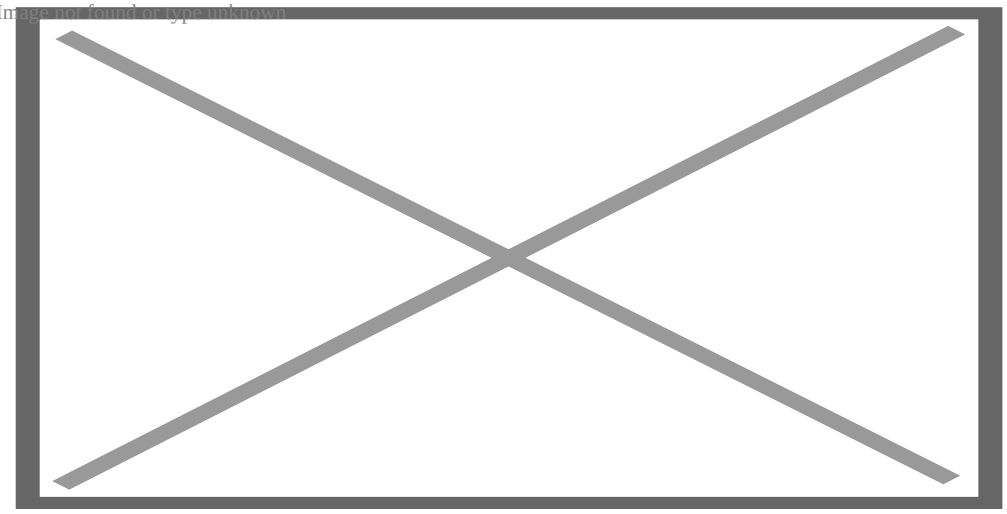
Graph resulting from the RDFa example

The example defines the following five triples (shown in Turtle syntax). Each triple represents one edge in the resulting graph: the first element of the triple (the *subject*) is the name of the node where the edge starts, the second element (the *predicate*) the type of the edge, and the last and third element (the *object*) either the name of the node where the edge ends or a literal value (e.g. a text, a number, etc.).

```
_:a <https://www.w3.org/1999/02/22-rdf-syntax-ns#type> <https://schema.org/Person> .
_:a <https://schema.org/name> "Paul Schuster" .
_:a <https://schema.org/birthPlace> <https://www.wikidata.org/entity/Q1731> .
<https://www.wikidata.org/entity/Q1731> <https://schema.org/itemtype> <https://schema.org/Place>
<https://www.wikidata.org/entity/Q1731> <https://schema.org/name> "Dresden" .
```

The triples result in the graph shown in the given figure.

Image not found or type unknown



Graph resulting from the RDFa example, enriched with further data from the Web

One of the advantages of using Uniform Resource Identifiers (URIs) is that they can be dereferenced using the HTTP protocol. According to the so-called Linked Open Data principles, such a dereferenced URI should result in a document that offers further data about the given URI. In this example, all URIs, both for edges and nodes (e.g. http://schema.org/Person, http://schema.org/birthPlace, http://www.wikidata.org/entity/Q1731) can be dereferenced and will result in further RDF graphs, describing the URI, e.g. that Dresden is a city in Germany, or that a person, in the sense of that URI, can be fictional.

The second graph shows the previous example, but now enriched with a few of the triples from the documents that result from dereferencing https://schema.org/Person (green edge) and https://www.wikidata.org/entity/Q1731 (blue edges).

Additionally to the edges given in the involved documents explicitly, edges can be automatically inferred: the triple

_:a <https://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://schema.org/Person> .

from the original RDFa fragment and the triple

<https://schema.org/Person> <http://www.w3.org/2002/07/owl#equivalentClass> <http://xmlns.com/f

from the document at https://schema.org/Person (green edge in the figure) allow to infer the following triple, given OWL semantics (red dashed line in the second Figure):

_:a <https://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://xmlns.com/foaf/0.1/Person> .

Background

[edit]
Further information: Semantic network § History

The concept of the semantic network model was formed in the early 1960s by researchers such as the cognitive scientist Allan M. Collins, linguist Ross Quillian and psychologist Elizabeth F. Loftus as a form to represent semantically structured knowledge. When applied in the context of the modern internet, it extends the network of hyperlinked human-readable web pages by inserting machine-readable metadata about pages and how they are related to each other. This enables automated agents to access the Web more intelligently and perform more tasks on behalf of users. The term "Semantic Web" was coined by Tim Berners-Lee,[7] the inventor of the World Wide Web and director of the World Wide Web Consortium ("W3C"), which oversees the development of proposed Semantic Web standards. He defines the Semantic Web as "a web of data that can be processed directly and indirectly by machines".

Many of the technologies proposed by the W3C already existed before they were positioned under the W3C umbrella. These are used in various contexts, particularly those dealing with information that encompasses a limited and defined domain, and where sharing data is a common necessity, such as scientific research or data exchange among businesses. In addition, other technologies with similar goals have emerged, such as microformats.

Limitations of HTML

[edit]

Many files on a typical computer can be loosely divided into either human-readable documents, or machine-readable data. Examples of human-readable document files are mail messages, reports, and brochures. Examples of machine-readable data files are calendars, address books, playlists, and spreadsheets, which are presented to a user using an application program that lets the files be viewed, searched, and combined.

Currently, the World Wide Web is based mainly on documents written in Hypertext Markup Language (HTML), a markup convention that is used for coding a body of text interspersed with multimedia objects such as images and interactive forms. Metadata tags provide a method by which computers can categorize the content of web pages. In the examples below, the field names "keywords", "description" and "author" are assigned values such as "computing", and "cheap widgets for sale" and "John Doe".

<meta name="keywords" content="computing, computer studies, computer" />
<meta name="description" content="Cheap widgets for sale" />
<meta name="author" content="John Doe" />

Because of this metadata tagging and categorization, other computer systems that want to access and share this data can easily identify the relevant values.

With HTML and a tool to render it (perhaps web browser software, perhaps another user agent), one can create and present a page that lists items for sale. The HTML of this catalog page can make simple, document-level assertions such as "this document's title is 'Widget Superstore' ", but there is no capability within the HTML itself to assert unambiguously that, for example, item number X586172 is an Acme Gizmo with a retail price of €199, or that it is a consumer product. Rather, HTML can only say that the span of text "X586172" is something that should be positioned near "Acme Gizmo" and "€199", etc. There is no way to say "this is a catalog" or even to establish that "Acme Gizmo" is a kind of title or that "€199" is a price. There is also no way to express that these pieces of information are bound together in describing a discrete item, distinct from other items perhaps listed on the page.

Semantic HTML refers to the traditional HTML practice of markup following intention, rather than specifying layout details directly. For example, the use of denoting "emphasis" rather than <i>, which specifies italics. Layout details are left up to the browser, in combination with Cascading Style Sheets. But this practice falls short of specifying the semantics of objects such as items for sale or prices.

Microformats extend HTML syntax to create machine-readable semantic markup about objects including people, organizations, events and products.[13] Similar initiatives include RDFa, Microdata and Schema.org.

Semantic Web solutions

[edit]

The Semantic Web takes the solution further. It involves publishing in languages specifically designed for data: Resource Description Framework (RDF), Web Ontology Language (OWL), and Extensible Markup Language (XML). HTML describes documents and the links between them. RDF, OWL, and XML, by contrast, can describe arbitrary things such as people, meetings, or airplane parts.

These technologies are combined in order to provide descriptions that supplement or replace the content of Web documents. Thus, content may manifest itself as descriptive data stored in Web-accessible databases,[14] or as markup within documents (particularly, in Extensible HTML (XHTML) interspersed with XML, or, more often, purely in XML, with layout or rendering cues stored separately). The machine-readable descriptions enable content managers to add meaning to the content, i.e., to describe the structure of the knowledge we have about that content. In this way, a machine can process knowledge itself, instead of text, using processes similar to human deductive reasoning and inference, thereby obtaining more meaningful results and helping computers to perform automated information gathering and research.

An example of a tag that would be used in a non-semantic web page:

<item>blog</item>

Encoding similar information in a semantic web page might look like this:

<item rdf:about="https://example.org/semantic-web/">Semantic Web</item>

Tim Berners-Lee calls the resulting network of Linked Data the Giant Global Graph, in contrast to the HTML-based World Wide Web. Berners-Lee posits that if the past was document sharing, the future is data sharing. His answer to the question of "how" provides three points of instruction. One, a URL should point to the data. Two, anyone accessing the URL should get data back. Three, relationships in the data should point to additional URLs with data.

Tags and identifiers

[edit]

Tags, including hierarchical categories and tags that are collaboratively added and maintained (e.g. with folksonomies) can be considered part of, of potential use to or a step towards the

semantic Web vision.[15][16][17]

Unique identifiers, including hierarchical categories and collaboratively added ones, analysis tools and metadata, including tags, can be used to create forms of semantic webs – webs that are to a certain degree semantic.[18] In particular, such has been used for structuring scientific research i.a. by research topics and scientific fields by the projects OpenAlex,[19][20][21] Wikidata and Scholia which are under development and provide APIs, Web-pages, feeds and graphs for various semantic queries.

Web 3.0

[edit]

Tim Berners-Lee has described the Semantic Web as a component of Web 3.0.[22]

> People keep asking what Web 3.0 is. I think maybe when you've got an overlay of scalable vector graphics – everything rippling and folding and looking misty – on Web 2.0 and access to a semantic Web integrated across a huge space of data, you'll have access to an unbelievable data resource …
>
> —Ãƒ Â¢Ã¢â€šÂ¬Ã…Â *Tim Berners-Lee, 2006*

"Semantic Web" is sometimes used as a synonym for "Web 3.0",[23] though the definition of each term varies.

Beyond Web 3.0

[edit]

The next generation of the Web is often termed Web 4.0, but its definition is not clear. According to some sources, it is a Web that involves artificial intelligence,[24] the internet of things, pervasive computing, ubiquitous computing and the Web of Things among other concepts.[25] According to the European Union, Web 4.0 is "the expected fourth generation of the World Wide Web. Using advanced artificial and ambient intelligence, the internet of things, trusted blockchain transactions, virtual worlds and XR capabilities, digital and real objects and environments are fully integrated and communicate with each other, enabling truly intuitive, immersive experiences, seamlessly blending the physical and digital worlds".[26]

Challenges

[edit]

Some of the challenges for the Semantic Web include vastness, vagueness, uncertainty, inconsistency, and deceit. Automated reasoning systems will have to deal with all of these issues in order to deliver on the promise of the Semantic Web.

- Vastness: The World Wide Web contains many billions of pages. The SNOMED CT medical terminology ontology alone contains 370,000 class names, and existing technology has not yet been able to eliminate all semantically duplicated terms. Any automated reasoning system will have to deal with truly huge inputs.
- Vagueness: These are imprecise concepts like "young" or "tall". This arises from the vagueness of user queries, of concepts represented by content providers, of matching query terms to provider terms and of trying to combine different knowledge bases with overlapping but subtly different concepts. Fuzzy logic is the most common technique for dealing with vagueness.
- Uncertainty: These are precise concepts with uncertain values. For example, a patient might present a set of symptoms that correspond to a number of different distinct diagnoses each with a different probability. Probabilistic reasoning techniques are generally employed to address uncertainty.
- Inconsistency: These are logical contradictions that will inevitably arise during the development of large ontologies, and when ontologies from separate sources are combined. Deductive reasoning fails catastrophically when faced with inconsistency, because "anything follows from a contradiction". Defeasible reasoning and paraconsistent reasoning are two techniques that can be employed to deal with inconsistency.
- Deceit: This is when the producer of the information is intentionally misleading the consumer of the information. Cryptography techniques are currently utilized to alleviate this threat. By providing a means to determine the information's integrity, including that which relates to the identity of the entity that produced or published the information, however credibility issues still have to be addressed in cases of potential deceit.

This list of challenges is illustrative rather than exhaustive, and it focuses on the challenges to the "unifying logic" and "proof" layers of the Semantic Web. The World Wide Web Consortium (W3C) Incubator Group for Uncertainty Reasoning for the World Wide Web[27] (URW3-XG) final report lumps these problems together under the single heading of "uncertainty".[28] Many of the techniques mentioned here will require extensions to the Web Ontology Language (OWL) for example to annotate conditional probabilities. This is an area of active research.[29]

Standards

[edit]

Standardization for Semantic Web in the context of Web 3.0 is under the care of W3C.[30]

Components

[edit]

The term "Semantic Web" is often used more specifically to refer to the formats and technologies that enable it.[5] The collection, structuring and recovery of linked data are enabled by technologies that provide a formal description of concepts, terms, and relationships within a given knowledge domain. These technologies are specified as W3C standards and include:

- Resource Description Framework (RDF), a general method for describing information
- RDF Schema (RDFS)
- Simple Knowledge Organization System (SKOS)
- SPARQL, an RDF query language
- Notation3 (N3), designed with human readability in mind
- N-Triples, a format for storing and transmitting data
- Turtle (Terse RDF Triple Language)
- Web Ontology Language (OWL), a family of knowledge representation languages
- Rule Interchange Format (RIF), a framework of web rule language dialects supporting rule interchange on the Web
- JavaScript Object Notation for Linked Data (JSON-LD), a JSON-based method to describe data
- ActivityPub, a generic way for client and server to communicate with each other. This is used by the popular decentralized social network Mastodon.

The Semantic Web Stack illustrates the architecture of the Semantic Web. The functions and relationships of the components can be summarized as follows:[31]

- XML provides an elemental syntax for content structure within documents, yet associates no semantics with the meaning of the content contained within. XML is not at present a necessary component of Semantic Web technologies in most cases, as alternative syntaxes exist, such as Turtle. Turtle is a de facto standard, but has not been through a formal standardization process.
- XML Schema is a language for providing and restricting the structure and content of elements contained within XML documents.
- RDF is a simple language for expressing data models, which refer to objects ("web resources") and their relationships. An RDF-based model can be represented in a variety of syntaxes, e.g., RDF/XML, N3, Turtle, and RDFa. RDF is a fundamental standard of the Semantic Web.[32][33]
- RDF Schema extends RDF and is a vocabulary for describing properties and classes of RDF-based resources, with semantics for generalized-hierarchies of such properties and classes.
- OWL adds more vocabulary for describing properties and classes: among others, relations between classes (e.g. disjointness), cardinality (e.g. "exactly one"), equality,

richer typing of properties, characteristics of properties (e.g. symmetry), and enumerated classes.
- SPARQL is a protocol and query language for semantic web data sources.
- RIF is the W3C Rule Interchange Format. It is an XML language for expressing Web rules that computers can execute. RIF provides multiple versions, called dialects. It includes a RIF Basic Logic Dialect (RIF-BLD) and RIF Production Rules Dialect (RIF PRD).

Current state of standardization

[edit]

Well-established standards:

- RDF - Resource Description Framework
- RDFS - Resource Description Framework Schema
- RIF - Rule Interchange Format
- SPARQL - 'SPARQL Protocol and RDF Query Language'
- Unicode
- URI - Uniform Resource Identifier
- OWL - Web Ontology Language
- XML - Extensible Markup Language

Not yet fully realized:

- Unifying Logic and Proof layers
- SWRL - Semantic Web Rule Language

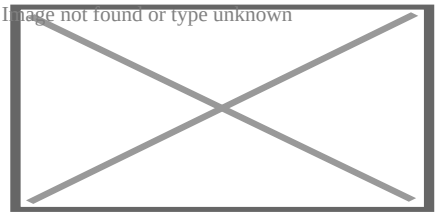
Applications

[edit]

The intent is to enhance the usability and usefulness of the Web and its interconnected resources by creating semantic web services, such as:

- Servers that expose existing data systems using the RDF and SPARQL standards. Many converters to RDF exist from different applications.[34] Relational databases are an important source. The semantic web server attaches to the existing system without affecting its operation.
- Documents "marked up" with semantic information (an extension of the HTML <meta> tags used in today's Web pages to supply information for Web search engines using web crawlers). This could be machine-understandable information about the human-understandable content of the document (such as the creator, title, description, etc.) or it could be purely metadata representing a set of facts (such as resources and services

elsewhere on the site). Note that *anything* that can be identified with a *Uniform Resource Identifier* (URI) can be described, so the semantic web can reason about animals, people, places, ideas, etc. There are four semantic annotation formats that can be used in HTML documents; Microformat, RDFa, Microdata and JSON-LD.[35] Semantic markup is often generated automatically, rather than manually.



Arguments as distinct semantic units with specified relations and version control on Kialo

- Common metadata vocabularies (ontologies) and maps between vocabularies that allow document creators to know how to mark up their documents so that agents can use the information in the supplied metadata (so that *Author* in the sense of 'the Author of the page' will not be confused with *Author* in the sense of a book that is the subject of a book review).
- Automated agents to perform tasks for users of the semantic web using this data.
- Semantic translation. An alternative or complementary approach are improvements to contextual and semantic understanding of texts – these could be aided via Semantic Web methods so that only increasingly small numbers of mistranslations need to be corrected in manual or semi-automated post-editing.
- Web-based services (often with agents of their own) to supply information specifically to agents, for example, a Trust service that an agent could ask if some online store has a history of poor service or spamming.
- Semantic Web ideas are implemented in collaborative structured argument mapping sites where their relations are organized semantically, arguments can be mirrored (linked) to multiple places, reused (copied), rated, and changed as semantic distinct units. Ideas for such, or a more widely adopted "World Wide Argument Web", go back to at least 2007[36] and have been implemented to some degree in Argüman[37] and Kialo. Further steps towards semantic web services may include enabling "Querying", argument search engines,[38] and "summarizing the contentious and agreed-upon points of a discussion".[39]

Such services could be useful to public search engines, or could be used for knowledge management within an organization. Business applications include:

- Facilitating the integration of information from mixed sources[40]
- Dissolving ambiguities in corporate terminology
- Improving information retrieval thereby reducing information overload and increasing the refinement and precision of the data retrieved[41][42][43][44]
- Identifying relevant information with respect to a given domain[45]

- Providing decision making support

In a corporation, there is a closed group of users and the management is able to enforce company guidelines like the adoption of specific ontologies and use of semantic annotation. Compared to the public Semantic Web there are lesser requirements on scalability and the information circulating within a company can be more trusted in general; privacy is less of an issue outside of handling of customer data.

Skeptical reactions

[edit]

Practical feasibility

[edit]

Critics question the basic feasibility of a complete or even partial fulfillment of the Semantic Web, pointing out both difficulties in setting it up and a lack of general-purpose usefulness that prevents the required effort from being invested. In a 2003 paper, Marshall and Shipman point out the cognitive overhead inherent in formalizing knowledge, compared to the authoring of traditional web hypertext:[46]

> While learning the basics of HTML is relatively straightforward, learning a knowledge representation language or tool requires the author to learn about the representation's methods of abstraction and their effect on reasoning. For example, understanding the class-instance relationship, or the superclass-subclass relationship, is more than understanding that one concept is a "type of" another concept. [...] These abstractions are taught to computer scientists generally and knowledge engineers specifically but do not match the similar natural language meaning of being a "type of" something. Effective use of such a formal representation requires the author to become a skilled knowledge engineer in addition to any other skills required by the domain. [...] Once one has learned a formal representation language, it is still often much more effort to express ideas in that representation than in a less formal representation [...]. Indeed, this is a form of programming based on the declaration of semantic data and requires an understanding of how reasoning algorithms will interpret the authored structures.

According to Marshall and Shipman, the tacit and changing nature of much knowledge adds to the knowledge engineering problem, and limits the Semantic Web's applicability to specific domains. A further issue that they point out are domain- or organization-specific ways to express knowledge, which must be solved through community agreement rather than only technical means.[46] As it turns out, specialized communities and organizations for intra-company projects have tended to adopt semantic web technologies greater than peripheral and less-specialized communities.[47] The practical constraints toward adoption have

appeared less challenging where domain and scope is more limited than that of the general public and the World-Wide Web.[47]

Finally, Marshall and Shipman see pragmatic problems in the idea of (Knowledge Navigator-style) intelligent agents working in the largely manually curated Semantic Web:[46]

> In situations in which user needs are known and distributed information resources are well described, this approach can be highly effective; in situations that are not foreseen and that bring together an unanticipated array of information resources, the Google approach is more robust. Furthermore, the Semantic Web relies on inference chains that are more brittle; a missing element of the chain results in a failure to perform the desired action, while the human can supply missing pieces in a more Google-like approach. [...] cost-benefit tradeoffs can work in favor of specially-created Semantic Web metadata directed at weaving together sensible well-structured domain-specific information resources; close attention to user/customer needs will drive these federations if they are to be successful.

Cory Doctorow's critique ("metacrap")[48] is from the perspective of human behavior and personal preferences. For example, people may include spurious metadata into Web pages in an attempt to mislead Semantic Web engines that naively assume the metadata's veracity. This phenomenon was well known with metatags that fooled the Altavista ranking algorithm into elevating the ranking of certain Web pages: the Google indexing engine specifically looks for such attempts at manipulation. Peter Gärdenfors and Timo Honkela point out that logic-based semantic web technologies cover only a fraction of the relevant phenomena related to semantics.[49][50]

Censorship and privacy

[edit]

Enthusiasm about the semantic web could be tempered by concerns regarding censorship and privacy. For instance, text-analyzing techniques can now be easily bypassed by using other words, metaphors for instance, or by using images in place of words. An advanced implementation of the semantic web would make it much easier for governments to control the viewing and creation of online information, as this information would be much easier for an automated content-blocking machine to understand. In addition, the issue has also been raised that, with the use of FOAF files and geolocation meta-data, there would be very little anonymity associated with the authorship of articles on things such as a personal blog. Some of these concerns were addressed in the "Policy Aware Web" project[51] and is an active research and development topic.

Doubling output formats

[edit]

Another criticism of the semantic web is that it would be much more time-consuming to create and publish content because there would need to be two formats for one piece of data: one for human viewing and one for machines. However, many web applications in development are addressing this issue by creating a machine-readable format upon the publishing of data or the request of a machine for such data. The development of microformats has been one reaction to this kind of criticism. Another argument in defense of the feasibility of semantic web is the likely falling price of human intelligence tasks in digital labor markets, such as Amazon's Mechanical Turk.[*citation needed*]

Specifications such as eRDF and RDFa allow arbitrary RDF data to be embedded in HTML pages. The GRDDL (Gleaning Resource Descriptions from Dialects of Language) mechanism allows existing material (including microformats) to be automatically interpreted as RDF, so publishers only need to use a single format, such as HTML.

Research activities on corporate applications

[edit]

The first research group explicitly focusing on the Corporate Semantic Web was the ACACIA team at INRIA-Sophia-Antipolis, founded in 2002. Results of their work include the RDF(S) based Corese[52] search engine, and the application of semantic web technology in the realm of distributed artificial intelligence for knowledge management (e.g. ontologies and multi-agent systems for corporate semantic Web) [53] and E-learning.[54]

Since 2008, the Corporate Semantic Web research group, located at the Free University of Berlin, focuses on building blocks: Corporate Semantic Search, Corporate Semantic Collaboration, and Corporate Ontology Engineering.[55]

Ontology engineering research includes the question of how to involve non-expert users in creating ontologies and semantically annotated content[56] and for extracting explicit knowledge from the interaction of users within enterprises.

Future of applications

[edit]

Tim O'Reilly, who coined the term Web 2.0, proposed a long-term vision of the Semantic Web as a web of data, where sophisticated applications are navigating and manipulating it.[57] The data web transforms the World Wide Web from a distributed file system into a distributed database.[58]

See also

[edit]

- AGRIS
- Business semantics management
- Computational semantics
- Calais (Reuters product)
- DBpedia
- Entity–attribute–value model
- EU Open Data Portal
- History of the World Wide Web
- Hyperdata
- Internet of things
- Linked data
- List of emerging technologies
- Nextbio
- Ontology alignment
- Ontology learning
- RDF and OWL
- Semantic computing
- Semantic Geospatial Web
- Semantic heterogeneity
- Semantic integration
- Semantic matching
- Semantic MediaWiki
- Semantic Sensor Web
- Semantic social network
- Semantic technology
- *Semantic Web*
- Semantically-Interlinked Online Communities
- Smart-M3
- Social Semantic Web
- Web engineering
- Web resource
- Web science

References

[edit]

1. ^ Semantic Web at W3C: https://www.w3.org/standards/semanticweb/
2. ^ *"World Wide Web Consortium (W3C), "RDF/XML Syntax Specification (Revised)", 25 Feb. 2014"*.
3. ^ *"World Wide Web Consortium (W3C), "OWL Web Ontology Language Overview", W3C Recommendation, 10 Feb. 2004"*.

4. ^ Chung, Seung-Hwa (2018). "The MOUSE approach: Mapping Ontologies using UML for System Engineers". Computer Reviews Journal: 8–29. ISSN 2581-6640.
5. ^ **a** **b** "W3C Semantic Web Activity". World Wide Web Consortium (W3C). November 7, 2011. Retrieved November 26, 2011.
6. ^ "Q&A with Tim Berners-Lee, Special Report". Bloomberg. Retrieved 14 April 2018.
7. ^ **a** **b** Berners-Lee, Tim; James Hendler; Ora Lassila (May 17, 2001). "The Semantic Web". Scientific American. Retrieved July 2, 2019.
8. ^ Lee Feigenbaum (May 1, 2007). "The Semantic Web in Action". Scientific American. Retrieved February 24, 2010.
9. ^ Berners-Lee, Tim; Fischetti, Mark (1999). Weaving the Web. HarperSanFrancisco. chapter 12. ISBN 978-0-06-251587-2.
10. ^ Berners-Lee, Tim; Hendler, James; Lassila, Ora (May 17, 2001). "The Semantic Web" (PDF). Scientific American. Vol. 284, no. 5. pp. 34–43. JSTOR 26059207. S2CID 56818714. Archived from the original (PDF) on October 10, 2017. Retrieved March 13, 2008.
11. ^ Nigel Shadbolt; Wendy Hall; Tim Berners-Lee (2006). "The Semantic Web Revisited" (PDF). IEEE Intelligent Systems. Archived from the original (PDF) on March 20, 2013. Retrieved April 13, 2007.
12. ^ Ramanathan V. Guha (2013). "Light at the End of the Tunnel". International Semantic Web Conference 2013 Keynote. Retrieved March 8, 2015.
13. ^ Allsopp, John (March 2007). Microformats: Empowering Your Markup for Web 2.0. Friends of ED. p. 368. ISBN 978-1-59059-814-6.
14. ^ Artem Chebotko and Shiyong Lu, "Querying the Semantic Web: An Efficient Approach Using Relational Databases", LAP Lambert Academic Publishing, ISBN 978-3-8383-0264-5, 2009.
15. ^ "Towards the Semantic Web: Collaborative Tag Suggestions" (PDF).
16. ^ Specia, Lucia; Motta, Enrico (2007). "Integrating Folksonomies with the Semantic Web". The Semantic Web: Research and Applications. Lecture Notes in Computer Science. Vol. 4519. Springer. pp. 624–639. doi:10.1007/978-3-540-72667-8_44. ISBN 978-3-540-72666-1.
17. ^ "Bridging the gap between folksonomies and the semantic web: an experience report" (PDF).
18. ^ Nicholson, Josh M.; Mordaunt, Milo; Lopez, Patrice; Uppala, Ashish; Rosati, Domenic; Rodrigues, Neves P.; Grabitz, Peter; Rife, Sean C. (5 November 2021). "scite: A smart citation index that displays the context of citations and classifies their intent using deep learning". Quantitative Science Studies. **2** (3): 882–898. doi:10.1162/qss_a_00146.
19. ^ Singh Chawla, Dalmeet (24 January 2022). "Massive open index of scholarly papers launches". Nature. doi:10.1038/d41586-022-00138-y. Retrieved 14 February 2022.
20. ^ "OpenAlex: The Promising Alternative to Microsoft Academic Graph". Singapore Management University (SMU). Retrieved 14 February 2022.
21. ^ "OpenAlex Documentation". Retrieved 18 February 2022.
22. ^ Shannon, Victoria (23 May 2006). "A 'more revolutionary' Web". International Herald Tribune. Retrieved 26 June 2006.

23. ^ *"Web 3.0 Explained, Plus the History of Web 1.0 and 2.0". Investopedia. Retrieved 2022-10-21.*
24. ^ https://www.rsisinternational.org/IJRSI/Issue31/75-78.pdf
25. ^ Almeida, F. (2017). Concept and dimensions of web 4.0. International journal of computers and technology, 16(7).
26. ^ *"The Commission wants the EU to lead on 'Web 4.0' — whatever that is". 11 July 2023.*
27. ^ *"W3C Uncertainty Reasoning for the World Wide Web". www.w3.org. Retrieved 2021-05-14.*
28. ^ *"Uncertainty Reasoning for the World Wide Web". W3.org. Retrieved 20 December 2018.*
29. ^ *Lukasiewicz, Thomas; Umberto Straccia (2008). "Managing uncertainty and vagueness in description logics for the Semantic Web" (PDF). Web Semantics: Science, Services and Agents on the World Wide Web. **6** (4): 291–308. doi:10.1016/j.websem.2008.04.001.*
30. ^ *"Semantic Web Standards". W3.org. Retrieved 14 April 2018.*
31. ^ *"OWL Web Ontology Language Overview". World Wide Web Consortium (W3C). February 10, 2004. Retrieved November 26, 2011.*
32. ^ *"Resource Description Framework (RDF)". World Wide Web Consortium.*
33. ^ *Allemang, Dean; Hendler, James; Gandon, Fabien (August 3, 2020). Semantic Web for the Working Ontologist : Effective Modeling for Linked Data, RDFS, and OWL (Third ed.). [New York, NY, USA]: ACM Books; 3rd edition. ISBN 978-1450376143.*
34. ^ *"ConverterToRdf - W3C Wiki". W3.org. Retrieved 20 December 2018.*
35. ^ *Sikos, Leslie F. (2015). Mastering Structured Data on the Semantic Web: From HTML5 Microdata to Linked Open Data. Apress. p. 23. ISBN 978-1-4842-1049-9.*
36. ^ *Kiesel, Johannes; Lang, Kevin; Wachsmuth, Henning; Hornecker, Eva; Stein, Benno (14 March 2020). "Investigating Expectations for Voice-based and Conversational Argument Search on the Web". Proceedings of the 2020 Conference on Human Information Interaction and Retrieval. ACM. pp. 53–62. doi:10.1145/3343413.3377978. ISBN 9781450368926. S2CID 212676751.*
37. ^ *Vetere, Guido (30 June 2018). "L'impossibile necessità delle piattaforme sociali decentralizzate". DigitCult - Scientific Journal on Digital Cultures. **3** (1): 41–50. doi: 10.4399/97888255159096.*
38. ^ *Bikakis, Antonis; Flouris, Giorgos; Patkos, Theodore; Plexousakis, Dimitris (2023). "Sketching the vision of the Web of Debates". Frontiers in Artificial Intelligence. **6**. doi: 10.3389/frai.2023.1124045. ISSN 2624-8212. PMC 10313200. PMID 37396970.*
39. ^ *Schneider, Jodi; Groza, Tudor; Passant, Alexandre. "A Review of Argumentation for the Social Semantic Web" (PDF).* cite journal: Cite journal requires |journal= (help)
40. ^ *Zhang, Chuanrong; Zhao, Tian; Li, Weidong (2015). Geospatial Semantic Web. Springer International Publishing : Imprint: Springer. ISBN 978-3-319-17801-1.*
41. ^ Omar Alonso and Hugo Zaragoza. 2008. Exploiting semantic annotations in information retrieval: ESAIR '08. SIGIR Forum 42, 1 (June 2008), 55–58. doi: 10.1145/1394251.1394262
42. ^ Jaap Kamps, Jussi Karlgren, and Ralf Schenkel. 2011. Report on the third workshop on exploiting semantic annotations in information retrieval (ESAIR). SIGIR Forum 45, 1 (May 2011), 33–41. doi:10.1145/1988852.1988858

43. ^ Jaap Kamps, Jussi Karlgren, Peter Mika, and Vanessa Murdock. 2012. Fifth workshop on exploiting semantic annotations in information retrieval: ESAIR '12). In Proceedings of the 21st ACM international conference on information and knowledge management (CIKM '12). ACM, New York, NY, USA, 2772–2773. doi:10.1145/2396761.2398761
44. ^ Omar Alonso, Jaap Kamps, and Jussi Karlgren. 2015. Report on the Seventh Workshop on Exploiting Semantic Annotations in Information Retrieval (ESAIR '14). SIGIR Forum 49, 1 (June 2015), 27–34. doi:10.1145/2795403.2795412
45. ^ *Kuriakose, John (September 2009). "Understanding and Adopting Semantic Web Technology". Cutter IT Journal. **22** (9). CUTTER INFORMATION CORP.: 10–18.*
46. ^ ***a b c** Marshall, Catherine C.; Shipman, Frank M. (2003). Which semantic web? (PDF). Proc. ACM Conf. on Hypertext and Hypermedia. pp. 57–66. Archived from the original (PDF) on 2015-09-23. Retrieved 2015-04-17.*
47. ^ ***a b** Ivan Herman (2007). State of the Semantic Web (PDF). Semantic Days 2007. Retrieved July 26, 2007.*
48. ^ *Doctorow, Cory. "Metacrap: Putting the torch to seven straw-men of the meta-utopia". www.well.com/. Retrieved 11 September 2023.*
49. ^ *Gärdenfors, Peter (2004). How to make the Semantic Web more semantic. IOS Press. pp. 17–34.* cite book: |work= ignored (help)
50. ^ *Honkela, Timo; Könönen, Ville; Lindh-Knuutila, Tiina; Paukkeri, Mari-Sanna (2008). "Simulating processes of concept formation and communication". Journal of Economic Methodology. **15** (3): 245–259. doi:10.1080/13501780802321350. S2CID 16994027.*
51. ^ *"Policy Aware Web Project". Policyawareweb.org. Retrieved 2013-06-14.*
52. ^ *Corby, Olivier; Dieng-Kuntz, Rose; Zucker, Catherine Faron; Gandon, Fabien (2006). "Searching the Semantic Web: Approximate Query Processing based on Ontologies". IEEE Intelligent Systems. **21**: 20–27. doi:10.1109/MIS.2006.16. S2CID 11488848.*
53. ^ *Gandon, Fabien (7 November 2002). Distributed Artificial Intelligence And Knowledge Management: Ontologies And Multi-Agent Systems For A Corporate Semantic Web (phdthesis). Université Nice Sophia Antipolis.*
54. ^ *Buffa, Michel; Dehors, Sylvain; Faron-Zucker, Catherine; Sander, Peter (2005). "Towards a Corporate Semantic Web Approach in Designing Learning Systems: Review of the Trial Solutioins Project" (PDF). International Workshop on Applications of Semantic Web Technologies for E-Learning. Amsterdam, Holland. pp. 73–76.*
55. ^ *"Corporate Semantic Web - Home". Corporate-semantic-web.de. Retrieved 14 April 2018.*
56. ^ *Hinze, Annika; Heese, Ralf; Luczak-Rösch, Markus; Paschke, Adrian (2012). "Semantic Enrichment by Non-Experts: Usability of Manual Annotation Tools" (PDF). ISWC'12 - Proceedings of the 11th international conference on The Semantic Web. Boston, USA. pp. 165–181.*
57. ^ *Mathieson, S. A. (6 April 2006). "Spread the word, and join it up". The Guardian. Retrieved 14 April 2018.*
58. ^ *Spivack, Nova (18 September 2007). "The Semantic Web, Collective Intelligence and Hyperdata". novaspivack.typepad.com/nova_spivacks_weblog [This Blog has Moved to NovaSpivack.com]. Retrieved 14 April 2018.*

Further reading

[edit]

- *Liyang Yu (December 14, 2014). A Developer's Guide to the Semantic Web, 2nd ed. Springer. ISBN 978-3-662-43796-4.*
- Aaron Swartz's A Programmable Web: An unfinished Work donated by Morgan & Claypool Publishers after Aaron Swartz's death in January 2013.
- *Grigoris Antoniou, Frank van Harmelen (March 31, 2008). A Semantic Web Primer, 2nd Edition. The MIT Press. ISBN 978-0-262-01242-3.*
- *Allemang, Dean; Hendler, James; Gandon, Fabien (August 3, 2020). Semantic Web for the Working Ontologist : Effective Modeling for Linked Data, RDFS, and OWL (Third ed.). [New York, NY, USA]: ACM Books; 3rd edition. ISBN 978-1450376143.*
- *Pascal Hitzler; Markus Krötzsch; Sebastian Rudolph (August 25, 2009). Foundations of Semantic Web Technologies. CRCPress. ISBN 978-1-4200-9050-5.*
- *Thomas B. Passin (March 1, 2004). Explorer's Guide to the Semantic Web. Manning Publications. ISBN 978-1-932394-20-7.*
- *Jeffrey T. Pollock (March 23, 2009). Semantic Web For Dummies. For Dummies. ISBN 978-0-470-39679-7.*
- *Hitzler, Pascal (February 2021). "A Review of the Semantic Web Field". Communications of the ACM. **64** (2): 76–83. doi:10.1145/3397512.*
- *Unni, Deepak (March 2023). "FAIRification of health-related data using semantic web technologies in the Swiss Personalized Health Network". Scientific Data. **10** (1): 127. Bibcode:2023NatSD..10..127T. doi:10.1038/s41597-023-02028-y. PMC 10006404. PMID 36899064.*

External links

[edit]

Semantic Web at Wikipedia's sister projects

- Image not found or type unknown Media from Commons
- Image not found or type unknown Textbooks from Wikibooks
- Image not found or type unknown Data from Wikidata

Image not found or type unknown

Scholia has a *topic* profile for ***Semantic Web***.

- Official website

- **v**
- **t**
- **e**

Semantic Web

Background	○ Databases ○ Hypertext ○ Internet ○ Ontologies ○ Semantics ○ Semantic networks ○ World Wide Web
Sub-topics	○ Dataspaces ○ Hyperdata ○ Linked data ○ Rule-based systems
Applications	○ Semantic analytics ○ Semantic broker ○ Semantic computing ○ Semantic mapper ○ Semantic matching ○ Semantic publishing ○ Semantic reasoner ○ Semantic search ○ Semantic service-oriented architecture ○ Semantic wiki ○ Solid

Related topics

- Collective intelligence
- Description logic
- Folksonomy
- Geotagging
- Information architecture
- iXBRL
- Knowledge extraction
- Knowledge management
- Knowledge representation and reasoning
- Library 2.0
- Digital library
- Digital humanities
- Metadata
- References
- Topic map
- Web 2.0
- Web engineering
- Web Science Trust

Standards		

	Syntax and supporting technologies	○ HTTP
		○ IRI
		  ○ URI
		○ RDF
		  ○ triples
		  ○ RDF/XML
		  ○ JSON-LD
		  ○ Turtle
		  ○ TriG
		  ○ Notation3
		  ○ N-Triples
		  ○ TriX (no W3C standard)
		○ RRID
		○ SPARQL
		○ XML
		○ Semantic HTML

	Schemas, ontologies and rules	○ Common Logic ○ OWL ○ RDFS ○ Rule Interchange Format ○ Semantic Web Rule Language ○ ALPS ○ SHACL
	Semantic annotation	○ eRDF ○ GRDDL ○ Microdata ○ Microformats ○ RDFa ○ SAWSDL ○ Facebook Platform
	Common vocabularies	○ DOAP ○ Dublin Core ○ FOAF ○ Schema.org ○ SIOC ○ SKOS
	Microformat vocabularies	○ hAtom ○ hCalendar ○ hCard ○ hProduct

- **v**
- **t**
- **e**

Emerging technologies

Fields **Information and communications**

- Ambient intelligence
  - Internet of things
- Artificial intelligence
  - Applications of artificial intelligence
  - Machine translation
  - Machine vision
  - Mobile translation
  - Progress in artificial intelligence
  - Semantic Web
  - Speech recognition
- Atomtronics
- Carbon nanotube field-effect transistor
- Cybermethodology
- Extended reality
- Fourth-generation optical discs
  - 3D optical data storage
  - Holographic data storage
- GPGPU
- Memory
  - CBRAM
  - ECRAM
  - FRAM
  - Millipede
  - MRAM
  - NRAM
  - PRAM
  - Racetrack memory
  - RRAM
  - SONOS
  - UltraRAM
- Optical computing
- RFID
  - Chipless RFID
- Software-defined radio
- Three-dimensional integrated circuit

Topics

- Automation
- Collingridge dilemma
- Differential technological development
- Disruptive innovation
- Ephemeralization
- Ethics
  - Bioethics
  - Cyberethics
  - Neuroethics
  - Robot ethics
- Exploratory engineering
- Proactionary principle
- Technological change
  - Technological unemployment
- Technological convergence
- Technological evolution
- Technological paradigm
- Technology forecasting
  - Accelerating change
  - Future-oriented technology analysis
  - Horizon scanning
  - Moore's law
  - Technological singularity
  - Technology scouting
- Technology in science fiction
- Technology readiness level
- Technology roadmap
- Transhumanism

- Image not found or type unknown **List**

- **v**
- **t**
- **e**

Digital humanities

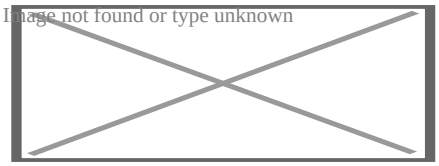
- Computational archaeology
- Computational philosophy
- Computational theory of mind
- Computers and writing
- Cultural analytics
- Cybertext
- Digital classics
- Digital history
- Digital library
- Digital Medievalist
- Digital ontology
- Digital physics
- Digital religion
- Digital rhetoric
- Digital scholarship
- Digital theology
- Digitization
- E-research
- Electronic literature
- Humanistic informatics
- New media
- Philosophy of computer science
- Semantic Web
- Systems theory
- Text Encoding Initiative
- Transliteracy

Authority control databases Image not found or type unknown Edit this at Wikidata

International	FAST
National	Germany United States France BnF data Japan Czech Republic Spain Latvia Israel

About Domain name

This article is about domain names in the Internet. For other uses, see Domain (disambiguation).



An annotated example of a domain name

In the Internet, a **domain name** is a string that identifies a realm of administrative autonomy, authority or control. Domain names are often used to identify services provided through the Internet, such as websites, email services and more. Domain names are used in various networking contexts and for application-specific naming and addressing purposes. In general, a domain name identifies a network domain or an Internet Protocol (IP) resource, such as a personal computer used to access the Internet, or a server computer.

Domain names are formed by the rules and procedures of the Domain Name System (DNS). Any name registered in the DNS is a domain name. Domain names are organized in subordinate levels (*subdomains*) of the DNS root domain, which is nameless. The first-level set of domain names are the *top-level domains* (TLDs), including the *generic top-level domains* (gTLDs), such as the prominent domains com, info, net, edu, and org, and the *country code top-level domains* (ccTLDs). Below these top-level domains in the DNS hierarchy are the second-level and third-level domain names that are typically open for reservation by end-users who wish to connect local area networks to the Internet, create other publicly accessible Internet resources or run websites, such as "wikipedia.org". The registration of a second- or third-level domain name is usually administered by a domain name registrar who sell its services to the public.

A *fully qualified domain name* (FQDN) is a domain name that is completely specified with all labels in the hierarchy of the DNS, having no parts omitted. Traditionally a FQDN ends in a dot (**.**) to denote the top of the DNS tree.[1] Labels in the Domain Name System are case-insensitive, and may therefore be written in any desired capitalization method, but most commonly domain names are written in lowercase in technical contexts.[2] A *hostname* is a domain name that has at least one associated IP address.

Purpose

[edit]

Domain names serve to identify Internet resources, such as computers, networks, and services, with a text-based label that is easier to memorize than the numerical addresses used in the Internet protocols. A domain name may represent entire collections of such resources or individual instances. Individual Internet host computers use domain names as host identifiers, also called hostnames. The term *hostname* is also used for the leaf labels in the domain name system, usually without further subordinate domain name space. Hostnames appear as a component in Uniform Resource Locators (URLs) for Internet resources such as websites (e.g., en.wikipedia.org).

Domain names are also used as simple identification labels to indicate ownership or control of a resource. Such examples are the realm identifiers used in the Session Initiation Protocol (SIP), the Domain Keys used to verify DNS domains in e-mail systems, and in many other Uniform Resource Identifiers (URIs).

An important function of domain names is to provide easily recognizable and memorizable names to numerically addressed Internet resources. This abstraction allows any resource to be moved to a different physical location in the address topology of the network, globally or locally in an intranet. Such a move usually requires changing the IP address of a resource and the corresponding translation of this IP address to and from its domain name.

Domain names are used to establish a unique identity. Organizations can choose a domain name that corresponds to their name, helping Internet users to reach them easily.

A generic domain is a name that defines a general category, rather than a specific or personal instance, for example, the name of an industry, rather than a company name. Some examples of generic names are *books.com*, *music.com*, and *travel.info*. Companies have created brands based on generic names, and such generic domain names may be valuable.[3]

Domain names are often simply referred to as *domains* and domain name registrants are frequently referred to as *domain owners*, although domain name registration with a registrar does not confer any legal ownership of the domain name, only an exclusive right of use for a particular duration of time. The use of domain names in commerce may subject them to trademark law.

History

[edit]
Main article: List of the oldest currently registered Internet domain names

The practice of using a simple memorable abstraction of a host's numerical address on a computer network dates back to the ARPANET era, before the advent of today's commercial Internet. In the early network, each computer on the network retrieved the hosts file (*host.txt*) from a computer at SRI (now SRI International),[4][5] which mapped computer hostnames to numerical addresses. The rapid growth of the network made it impossible to maintain a

centrally organized hostname registry and in 1983 the Domain Name System was introduced on the ARPANET and published by the Internet Engineering Task Force as RFC 882 and RFC 883.

The following table shows the first five .com domains with the dates of their registration:[6]

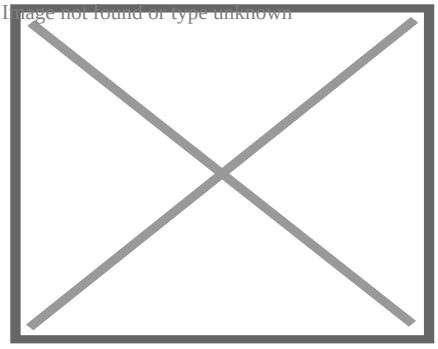
Domain name	Registration date
symbolics.com	15 March 1985
bbn.com	24 April 1985
think.com	24 May 1985
mcc.com	11 July 1985
dec.com	30 September 1985

and the first five .edu domains:[7]

Domain name	Registration date
berkeley.edu	24 April 1985
cmu.edu	24 April 1985
purdue.edu	24 April 1985
rice.edu	24 April 1985
ucla.edu	24 April 1985

Domain name space

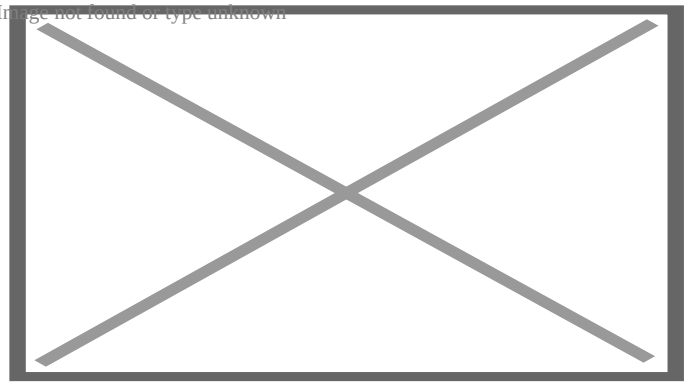
[edit]



The hierarchical domain name system, organized into zones, each served by domain name servers

Today, the Internet Corporation for Assigned Names and Numbers (ICANN) manages the top-level development and architecture of the Internet domain name space. It authorizes domain name registrars, through which domain names may be registered and reassigned.

Image not found or type unknown



The hierarchy of labels in a fully qualified domain name

The domain name space consists of a tree of domain names. Each node in the tree holds information associated with the domain name. The tree sub-divides into *zones* beginning at the DNS root zone.

Domain name syntax

[edit]

A domain name consists of one or more parts, technically called *labels*, that are conventionally concatenated, and delimited by dots, such as *example.com*.

- The right-most label conveys the top-level domain; for example, the domain name *www.example.com* belongs to the top-level domain *com*.
- The hierarchy of domains descends from the right to the left label in the name; each label to the left specifies a subdivision, or subdomain of the domain to the right. For example: the label *example* specifies a node *example.com* as a subdomain of the *com* domain, and *www* is a label to create *www.example.com*, a subdomain of *example.com*. Each label may contain from 1 to 63 octets. The empty label is reserved for the root node and when fully qualified is expressed as the empty label terminated by a dot. The full domain name may not exceed a total length of 253 ASCII characters in its textual representation.[8]
- A hostname is a domain name that has at least one associated IP address. For example, the domain names *www.example.com* and *example.com* are also hostnames, whereas the *com* domain is not. However, other top-level domains, particularly country code top-level domains, may indeed have an IP address, and if so, they are also hostnames.
- Hostnames impose restrictions on the characters allowed in the corresponding domain name. A valid hostname is also a valid domain name, but a valid domain name may not necessarily be valid as a hostname.

Top-level domains

[edit]

When the Domain Name System was devised in the 1980s, the domain name space was divided into two main groups of domains.[9] The country code top-level domains (ccTLD) were primarily based on the two-character territory codes of ISO-3166 country abbreviations. In addition, a group of seven generic top-level domains (gTLD) was implemented which represented a set of categories of names and multi-organizations.[10] These were the domains gov, edu, com, mil, org, net, and int. These two types of top-level domains (TLDs) are the highest level of domain names of the Internet. Top-level domains form the DNS root zone of the hierarchical Domain Name System. Every domain name ends with a top-level domain label.

During the growth of the Internet, it became desirable to create additional generic top-level domains. As of October 2009, 21 generic top-level domains and 250 two-letter country-code top-level domains existed.[11] In addition, the ARPA domain serves technical purposes in the infrastructure of the Domain Name System.

During the 32nd International Public ICANN Meeting in Paris in 2008,[12] ICANN started a new process of TLD naming policy to take a "significant step forward on the introduction of new generic top-level domains." This program envisions the availability of many new or already proposed domains, as well as a new application and implementation process.[13] Observers believed that the new rules could result in hundreds of new top-level domains to be registered.[14] In 2012, the program commenced, and received 1930 applications.[15] By 2016, the milestone of 1000 live gTLD was reached.

The Internet Assigned Numbers Authority (IANA) maintains an annotated list of top-level domains in the DNS root zone database.[16]

For special purposes, such as network testing, documentation, and other applications, IANA also reserves a set of special-use domain names.[17] This list contains domain names such as example, local, localhost, and test. Other top-level domain names containing trade marks are registered for corporate use. Cases include brands such as BMW, Google, and Canon.[18]

Second-level and lower level domains

[edit]

Below the top-level domains in the domain name hierarchy are the second-level domain (SLD) names. These are the names directly to the left of .com, .net, and the other top-level domains. As an example, in the domain *example.co.uk*, *co* is the second-level domain.

Next are third-level domains, which are written immediately to the left of a second-level domain. There can be fourth- and fifth-level domains, and so on, with virtually no limitation. Each label is separated by a full stop (dot). An example of an operational domain name with four levels of domain labels is *sos.state.oh.us*. 'sos' is said to be a sub-domain of 'state.oh.us', and 'state' a sub-domain of 'oh.us', etc. In general, subdomains are domains subordinate to

their parent domain. An example of very deep levels of subdomain ordering are the IPv6 reverse resolution DNS zones, e.g., 1.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.ip6.arpa, which is the reverse DNS resolution domain name for the IP address of a loopback interface, or the localhost name.

Second-level (or lower-level, depending on the established parent hierarchy) domain names are often created based on the name of a company (e.g., *bbc*.co.uk), product or service (e.g. *hotmail*.com). Below these levels, the next domain name component has been used to designate a particular host server. Therefore, *ftp.example.com* might be an FTP server, *www.example.com* would be a World Wide Web server, and *mail.example.com* could be an email server, each intended to perform only the implied function. Modern technology allows multiple physical servers with either different (cf. load balancing) or even identical addresses (cf. anycast) to serve a single hostname or domain name, or multiple domain names to be served by a single computer. The latter is very popular in Web hosting service centers, where service providers host the websites of many organizations on just a few servers.

The hierarchical DNS labels or components of domain names are separated in a fully qualified name by the full stop (dot, .).

Internationalized domain names

[edit]
Main article: Internationalized domain name

The character set allowed in the Domain Name System is based on ASCII and does not allow the representation of names and words of many languages in their native scripts or alphabets. ICANN approved the Internationalized domain name (IDNA) system, which maps Unicode strings used in application user interfaces into the valid DNS character set by an encoding called Punycode. For example, københavn.eu is mapped to xn--kbenhavn-54a.eu. Many registries have adopted IDNA.

Domain name registration

[edit]

History

[edit]

The first commercial Internet domain name, in the TLD *com*, was registered on 15 March 1985 in the name symbolics.com by Symbolics Inc., a computer systems firm in Cambridge, Massachusetts.

By 1992, fewer than 15,000 *com* domains had been registered.

In the first quarter of 2015, 294 million domain names had been registered.[19] A large fraction of them are in the *com* TLD, which as of December 21, 2014, had 115.6 million domain names,[20] including 11.9 million online business and e-commerce sites, 4.3 million entertainment sites, 3.1 million finance related sites, and 1.8 million sports sites.[21] As of July 15, 2012, the *com* TLD had more registrations than all of the ccTLDs combined.[22]

As of December 31, 2023, 359.8 million domain names had been registered.[23]

Administration

[edit]

The right to use a domain name is delegated by domain name registrars, which are accredited by the Internet Corporation for Assigned Names and Numbers (ICANN), the organization charged with overseeing the name and number systems of the Internet. In addition to ICANN, each top-level domain (TLD) is maintained and serviced technically by an administrative organization operating a registry. A registry is responsible for maintaining the database of names registered within the TLD it administers. The registry receives registration information from each domain name registrar authorized to assign names in the corresponding TLD and publishes the information using a special service, the WHOIS protocol.

Registries and registrars usually charge an annual fee for the service of delegating a domain name to a user and providing a default set of name servers. Often, this transaction is termed a sale or lease of the domain name, and the registrant may sometimes be called an "owner", but no such legal relationship is actually associated with the transaction, only the exclusive right to use the domain name. More correctly, authorized users are known as "registrants" or as "domain holders".

ICANN publishes the complete list of TLD registries and domain name registrars. Registrant information associated with domain names is maintained in an online database accessible with the WHOIS protocol. For most of the 250 country code top-level domains (ccTLDs), the domain registries maintain the WHOIS (Registrant, name servers, expiration dates, etc.) information.

Some domain name registries, often called *network information centers* (NIC), also function as registrars to end-users. The major generic top-level domain registries, such as for the *com*, *net*, *org*, *info* domains and others, use a registry-registrar model consisting of hundreds of domain name registrars (see lists at ICANN[24] or VeriSign).[25] In this method of management, the registry only manages the domain name database and the relationship with the registrars. The *registrants* (users of a domain name) are customers of the registrar, in some cases through additional layers of resellers.

There are also a few other alternative DNS root providers that try to compete or complement ICANN's role of domain name administration, however, most of them failed to receive wide recognition, and thus domain names offered by those alternative roots cannot be used universally on most other internet-connecting machines without additional dedicated configurations.

Technical requirements and process

[edit]

In the process of registering a domain name and maintaining authority over the new name space created, registrars use several key pieces of information connected with a domain:

- *Administrative contact*. A registrant usually designates an administrative contact to manage the domain name. The administrative contact usually has the highest level of control over a domain. Management functions delegated to the administrative contacts may include management of all business information, such as name of record, postal address, and contact information of the official registrant of the domain and the obligation to conform to the requirements of the domain registry in order to retain the right to use a domain name. Furthermore, the administrative contact installs additional contact information for technical and billing functions.
- *Technical contact*. The technical contact manages the name servers of a domain name. The functions of a technical contact include assuring conformance of the configurations of the domain name with the requirements of the domain registry, maintaining the domain zone records, and providing continuous functionality of the name servers (that leads to the accessibility of the domain name).
- *Billing contact*. The party responsible for receiving billing invoices from the domain name registrar and paying applicable fees.
- *Name servers*. Most registrars provide two or more name servers as part of the registration service. However, a registrant may specify its own authoritative name servers to host a domain's resource records. The registrar's policies govern the number of servers and the type of server information required. Some providers require a hostname and the corresponding IP address or just the hostname, which must be resolvable either in the new domain, or exist elsewhere. Based on traditional requirements (RFC 1034), typically a minimum of two servers is required.

A domain name consists of one or more labels, each of which is formed from the set of ASCII letters, digits, and hyphens (a–z, A–Z, 0–9, -), but not starting or ending with a hyphen. The labels are case-insensitive; for example, 'label' is equivalent to 'Label' or 'LABEL'. In the textual representation of a domain name, the labels are separated by a full stop (period).

Business models

[edit]

Domain names are often seen in analogy to real estate in that domain names are foundations on which a website can be built, and the highest *quality* domain names, like sought-after real estate, tend to carry significant value, usually due to their online brand-building potential, use in advertising, search engine optimization, and many other criteria.

A few companies have offered low-cost, below-cost or even free domain registration with a variety of models adopted to recoup the costs to the provider. These usually require that domains be hosted on their website within a framework or portal that includes advertising wrapped around the domain holder's content, revenue from which allows the provider to recoup the costs. Domain registrations were free of charge when the DNS was new. A domain holder may provide an infinite number of subdomains in their domain. For example, the owner of *example.org* could provide subdomains such as *foo.example.org* and *foo.bar.example.org* to interested parties.

Many desirable domain names are already assigned and users must search for other acceptable names, using Web-based search features, or WHOIS and dig operating system tools. Many registrars have implemented **domain name suggestion** tools which search domain name databases and suggest available alternative domain names related to keywords provided by the user.

Resale of domain names

[edit]
Main article: List of most expensive domain names

The business of resale of registered domain names is known as the domain aftermarket. Various factors influence the perceived value or market value of a domain name. Most of the high-prize domain sales are carried out privately.[26] Also, it is called confidential domain acquiring or anonymous domain acquiring.[27]

Domain name confusion

[edit]

Intercapping is often used to emphasize the meaning of a domain name, because DNS names are not case-sensitive. Some names may be misinterpreted in certain uses of capitalization. For example: *Who Represents*, a database of artists and agents, chose *whorepresents.com*,[28] which can be misread. In such situations, the proper meaning may be clarified by placement of hyphens when registering a domain name. For instance, Experts Exchange, a programmers' discussion site, used *expertsexchange.com*, but changed its domain name to *experts-exchange.com*.[29]

Uses in website hosting

[edit]

The domain name is a component of a uniform resource locator (URL) used to access websites, for example:

- URL: http://www.example.net/index.html
- Top-level domain: net
- Second-level domain: example
- Hostname: www

A domain name may point to multiple IP addresses to provide server redundancy for the services offered, a feature that is used to manage the traffic of large, popular websites.

Web hosting services, on the other hand, run servers that are typically assigned only one or a few addresses while serving websites for many domains, a technique referred to as virtual web hosting. Such IP address overloading requires that each request identifies the domain name being referenced, for instance by using the HTTP request header field *Host:*, or Server Name Indication.

Abuse and regulation

[edit]

Critics often claim abuse of administrative power over domain names. Particularly noteworthy was the VeriSign Site Finder system which redirected all unregistered .com and .net domains to a VeriSign webpage. For example, at a public meeting with VeriSign to air technical concerns about Site Finder,[30] numerous people, active in the IETF and other technical bodies, explained how they were surprised by VeriSign's changing the fundamental behavior of a major component of Internet infrastructure, not having obtained the customary consensus. Site Finder, at first, assumed every Internet query was for a website, and it monetized queries for incorrect domain names, taking the user to VeriSign's search site. Other applications, such as many implementations of email, treat a lack of response to a domain name query as an indication that the domain does not exist, and that the message can be treated as undeliverable. The original VeriSign implementation broke this assumption for mail, because it would always resolve an erroneous domain name to that of Site Finder. While VeriSign later changed Site Finder's behaviour with regard to email, there was still widespread protest about VeriSign's action being more in its financial interest than in the interest of the Internet infrastructure component for which VeriSign was the steward.

Despite widespread criticism, VeriSign only reluctantly removed it after the Internet Corporation for Assigned Names and Numbers (ICANN) threatened to revoke its contract to administer the root name servers. ICANN published the extensive set of letters exchanged,

committee reports, and ICANN decisions.[31]

There is also significant disquiet regarding the United States Government's political influence over ICANN. This was a significant issue in the attempt to create a .xxx top-level domain and sparked greater interest in alternative DNS roots that would be beyond the control of any single country.[32]

Additionally, there are numerous accusations of domain name front running, whereby registrars, when given whois queries, automatically register the domain name for themselves. Network Solutions has been accused of this.[33]

Truth in Domain Names Act

[edit]

In the United States, the Truth in Domain Names Act of 2003, in combination with the PROTECT Act of 2003, forbids the use of a misleading domain name with the intention of attracting Internet users into visiting Internet pornography sites.

The Truth in Domain Names Act follows the more general Anticybersquatting Consumer Protection Act passed in 1999 aimed at preventing typosquatting and deceptive use of names and trademarks in domain names.

Seizures

[edit]

- Seizure notices
  absolutepoker.com
- Image not found or type unknown

absolutepoker.com
channelsurfing.net

- Image not found or type unknown

channelsurfing.net
libertyreserve.com

Image not found or type unknown

libertyreserve.com

In the early 21st century, the US Department of Justice (DOJ) pursued the seizure of domain names, based on the legal theory that domain names constitute property used to engage in criminal activity, and thus are subject to forfeiture. For example, in the seizure of the domain name of a gambling website, the DOJ referenced 18 U.S.C. § 981 and 18 U.S.C. § 1955(d).[34][1] In 2013 the US government seized Liberty Reserve, citing 18 U.S.C. § 982(a)(1).[35]

The U.S. Congress passed the Combating Online Infringement and Counterfeits Act in 2010. Consumer Electronics Association vice president Michael Petricone was worried that seizure was a *blunt instrument* that could harm legitimate businesses.[36][37] After a joint operation on February 15, 2011, the DOJ and the Department of Homeland Security claimed to have seized ten domains of websites involved in advertising and distributing child pornography, but also mistakenly seized the domain name of a large DNS provider, temporarily replacing 84,000 websites with seizure notices.[38]

In the United Kingdom, the Police Intellectual Property Crime Unit (PIPCU) has been attempting to seize domain names from registrars without court orders.[39]

Suspensions

[edit]

PIPCU and other UK law enforcement organisations make domain suspension requests to Nominet which they process on the basis of breach of terms and conditions. Around 16,000 domains are suspended annually, and about 80% of the requests originate from PIPCU.[40]

Property rights

[edit]

Because of the economic value it represents, the European Court of Human Rights has ruled that the exclusive right to a domain name is protected as property under article 1 of Protocol 1 to the European Convention on Human Rights.[41]

IDN variants

[edit]

ICANN Business Constituency (BC) has spent decades trying to make IDN variants work at the second level, and in the last several years at the top level. Domain name variants are domain names recognized in different character encodings, like a single domain presented in traditional Chinese and simplified Chinese. It is an Internationalization and localization

problem. Under Domain Name Variants, the different encodings of the domain name (in simplified and traditional Chinese) would resolve to the same host.[42][43]

According to John Levine, an expert on Internet related topics, "Unfortunately, variants don't work. The problem isn't putting them in the DNS, it's that once they're in the DNS, they don't work anywhere else."[42]

Fictitious domain name

[edit]

A *fictitious domain name* is a domain name used in a work of fiction or popular culture to refer to a domain that does not actually exist, often with invalid or unofficial top-level domains such as ".web", a usage exactly analogous to the dummy 555 telephone number prefix used in film and other media. The canonical fictitious domain name is "example.com", specifically set aside by IANA in RFC 2606 for such use, along with the *.example* TLD.

Domain names used in works of fiction have often been registered in the DNS, either by their creators or by cybersquatters attempting to profit from it. This phenomenon prompted NBC to purchase the domain name Hornymanatee.com after talk-show host Conan O'Brien spoke the name while ad-libbing on his show. O'Brien subsequently created a website based on the concept and used it as a running gag on the show.[44] Companies whose works have used fictitious domain names have also employed firms such as MarkMonitor to park fictional domain names in order to prevent misuse by third parties.[45]

Misspelled domain names

[edit]

This section **does not cite any sources**. Please help improve this section by adding citations to reliable sources. Unsourced material may be challenged and removed. *(December 2022) (Learn how and when to remove this message)*

Misspelled domain names, also known as typosquatting or URL hijacking, are domain names that are intentionally or unintentionally misspelled versions of popular or well-known domain names. The goal of misspelled domain names is to capitalize on internet users who accidentally type in a misspelled domain name, and are then redirected to a different website.

Misspelled domain names are often used for malicious purposes, such as phishing scams or distributing malware. In some cases, the owners of misspelled domain names may also attempt to sell the domain names to the owners of the legitimate domain names, or to individuals or organizations who are interested in capitalizing on the traffic generated by internet users who accidentally type in the misspelled domain names.

To avoid being caught by a misspelled domain name, internet users should be careful to type in domain names correctly, and should avoid clicking on links that appear suspicious or unfamiliar. Additionally, individuals and organizations who own popular or well-known domain names should consider registering common misspellings of their domain names in order to prevent others from using them for malicious purposes.

Domain name spoofing

[edit]

The term **Domain name spoofing** (or simply though less accurately, **Domain spoofing**) is used generically to describe one or more of a class of phishing attacks that depend on falsifying or misrepresenting an internet domain name.[46][47] These are designed to persuade unsuspecting users into visiting a web site other than that intended, or opening an email that is not in reality from the address shown (or apparently shown).[48] Although website and email spoofing attacks are more widely known, any service that relies on domain name resolution may be compromised.

Types

[edit]

There are a number of better-known types of domain spoofing:

- Typosquatting, also called "URL hijacking", a "sting site", or a "fake URL", is a form of cybersquatting, and possibly brandjacking which relies on mistakes such as typos made by Internet users when inputting a website address into a web browser or composing an email address. Should a user accidentally enter an incorrect domain name, they may be led to any URL (including an alternative website owned by a cybersquatter).[49]

  The typosquatter's URL will usually be one of five kinds, all *similar to* the victim site address:
  - A common misspelling, or foreign language spelling, of the intended site
  - A misspelling based on a typographical error
  - A plural of a singular domain name
  - A different top-level domain: (i.e. .com instead of .org)
  - An abuse of the Country Code Top-Level Domain (ccTLD) (.cm, .co, or .om instead of .com)

- IDN homograph attack. This type of attack depends on registering a domain name that is similar to the 'target' domain, differing from it only because its spelling includes one or more characters that come from a different alphabet but look the same to the naked eye. For example, the Cyrillic, Latin, and Greek alphabets each have their own letter A, each of which has its own binary code point. Turkish has a dotless letter i (Ăƒâ€žÃ‚±) that

may not be perceived as different from the ASCII letter i. Most web browsers warn of 'mixed alphabet' domain names,[50][51][52][53] Other services, such as email applications, may not provide the same protection. Reputable top level domain and country code domain registrars will not accept applications to register a deceptive name but this policy cannot be presumed to be infallible.
- DNS spoofing – Cyberattack using corrupt DNS data
- Website spoofing – Creating a website, as a hoax, with the intention of misleading readers
- Email spoofing – Creating email spam or phishing messages with a forged sender identity or address

Risk mitigation

[edit]

- Domain Name System Security Extensions – Suite of IETF specifications
- Sender Policy Framework – Simple email-validation system designed to detect email spoofing
- DMARC – System to prevent email fraud ("Domain-based Message Authentication, Reporting and Conformance")
- DomainKeys Identified Mail – Email authentication method designed to detect email spoofing
- Public key certificate – Electronic document used to prove the ownership of a public key (SSL certificate)

Legitimate technologies that may be subverted

[edit]

- URL redirection – Technique for making a Web page available under more than one URL address
- Domain fronting – Technique for Internet censorship circumvention

See also

[edit]

- Domain hack
- Domain hijacking
- Domain name registrar
- Domain name speculation
- Domain name warehousing
- Domain registration

- Domain tasting
- Geodomain
- List of Internet top-level domains
- Reverse domain hijacking
- Reverse domain name notation

References

[edit]

1. ^ *Stevens, W. Richard (1994). TCP/IP Illustrated, Volume 1: The Protocols. Vol. 1 (1 ed.). Addison-Wesley. ISBN 9780201633467.*
2. ^ *Arends, R.; Austein, R.; Larson, M.; Massey, D.; Rose, S. (2005). RFC 4034 – Resource Records for the DNS Security Extensions (Technical report). IEFT. doi: 10.17487/RFC4034. Archived from the original on 2018-09-20. Retrieved 2015-07-05.*
3. ^ *Low, Jerry. "Why are generic domains so expensive?". TheRealJerryLow.com. Archived from the original on 20 March 2019. Retrieved 27 September 2018.*
4. ^ RFC 3467, *Role of the Domain Name System (DNS)*, J.C. Klensin, J. Klensin (February 2003)
5. ^ *Cricket Liu, Paul Albitz (2006). DNS and BIND (5th ed.). O'Reilly. p. 3. Archived from the original on 2011-09-05. Retrieved 2011-10-22.*
6. ^ *"The first ever 20 domain names registered". ComputerWeekly.com. Archived from the original on 2020-08-08. Retrieved 2020-07-30.*
7. ^ *Rooksby, Jacob H. (2015). "Defining Domain: Higher Education's Battles for Cyberspace". Brooklyn Law Review. **80** (3): 857–942. Archived from the original on 2018-11-07. Retrieved 2015-10-27.* at p. 869
8. ^ *Mockapetris, P. (November 1987). "Domain names - Implementation and specification (RFC 1035)". IETF Datatracker. Retrieved January 21, 2024.*
9. ^ *"Introduction to Top-Level Domains (gTLDs)". Internet Corporation for Assigned Names and Numbers (ICANN). Archived from the original on 2009-06-15. Retrieved 2009-06-26.*
10. ^ RFC 920, *Domain Requirements*, J. Postel, J. Reynolds, The Internet Society (October 1984)
11. ^ "New gTLD Program" Archived 2011-11-25 at the Wayback Machine, ICANN, October 2009
12. ^ *"32nd International Public ICANN Meeting". ICANN. 2008-06-22. Archived from the original on 2009-03-08. Retrieved 2009-06-26.*
13. ^ *"New gTLS Program". ICANN. Archived from the original on 2011-09-10. Retrieved 2009-06-15.*
14. ^ ICANN Board Approves Sweeping Overhaul of Top-level Domains Archived 2009-06-26 at the Wayback Machine, CircleID, 26 June 2008.
15. ^ *"About the Program - ICANN New gTLDs". ICANN. Archived from the original on 2016-11-03. Retrieved 2016-11-09.*
16. ^ *"Root Zone Database". IANA. Archived from the original on 2019-05-04. Retrieved 2020-11-01.*

17. ^ *Cheshire, S.; Krochmal, M. (February 2013). "RFC6761 - Special-Use Domain Names". Internet Engineering Task Force. doi:10.17487/RFC6761. Archived from the original on 13 November 2020. Retrieved 3 May 2015.*
18. ^ *"Executive Summary - dot brand observatory". observatory.domains. Archived from the original on 2016-11-10. Retrieved 2016-11-09.*
19. ^ Internet Grows to 294 Million Domain Names in the First Quarter of 2015 Archived 2017-12-20 at the Wayback Machine, Jun 30, 2015.
20. ^ *"Thirty years of .COM domains - and the numbers are up". Geekzone. Mar 13, 2015. Archived from the original on April 7, 2016. Retrieved Mar 25, 2016.*
21. ^ Evangelista, Benny. 2010. "25 years of .com names." San Francisco Chronicle. March 15, p. 1
22. ^ *"Domain domination: The com TLD larger than all ccTLDs combined". Royal.pingdom.com. Archived from the original on 2012-07-23. Retrieved 2012-07-25.*
23. ^ *"DNIB Quarterly Report Q4 2023". Domain Name Industry Brief (DNIB). Retrieved 16 February 2024.*
24. ^ *"ICANN-Accredited Registrars". ICANN. Archived from the original on 2019-05-19. Retrieved 2012-09-13.*
25. ^ *"Choose A Top Domain Registrar Of Your Choice Using Our Search Tool". Verisign. Archived from the original on 2015-09-04. Retrieved 2015-08-10.*
26. ^ *Arif, Sengoren (1 October 2024). "Confidentially domain acquiring".*
27. ^ *"Anonymous Domain Ownership". Conference: 2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC). 1 October 2024.*
28. ^ *Courtney, Curzi (14 October 2014). "WhoRepresents helps brands connect with celebrity influencers". DM News. Archived from the original on 8 July 2019. Retrieved 8 July 2019.*
29. ^ *Ki, Mae Heussner (2 June 2010). "'Slurls': Most Outrageous Website URLs". ABC News. Archived from the original on 31 May 2019. Retrieved 8 July 2019.*
30. ^ *McCullagh, Declan (2003-10-03). "VeriSign fends off critics at ICANN confab". CNET News. Archived from the original on January 4, 2013. Retrieved 2007-09-22.*
31. ^ *"Verisign's Wildcard Service Deployment". ICANN. Archived from the original on 2008-12-02. Retrieved 2007-09-22.*
32. ^ *Mueller, M (March 2004). Ruling the Root. MIT Press. ISBN 0-262-63298-5.*
33. ^ Slashdot.org Archived 2010-02-17 at the Wayback Machine, NSI Registers Every Domain Checked
34. ^ *FBI / DOJ (15 April 2011). "Warning". Archived from the original on 2011-04-14. Retrieved 2011-04-15.*
35. ^ *Dia, Miaz (4 February 2010). "website laten maken". Kmowebdiensten. Archived from the original on December 20, 2016. Retrieved 8 December 2016.*
36. ^ *Gabriel, Jeffrey (18 June 2020). "Past Congressional Attempts to Combat Online Copyright Infringement". Saw. Archived from the original on 2020-06-20. Retrieved 2020-06-19.*
37. ^ *Jerome, Sarah (6 April 2011). "Tech industry wary of domain name seizures". The Hill. Archived from the original on 2011-04-10. Retrieved 2011-04-15.*

38. ^ *"U.S. Government Shuts Down 84,000 Websites, 'By Mistake'". Archived from the original on 2018-12-25. Retrieved 2012-12-16.*
39. ^ *Jeftovic, Mark (8 October 2013). "Whatever Happened to "Due Process" ?". Archived from the original on 5 December 2014. Retrieved 27 November 2014.*
40. ^ Tackling online criminal activity Archived 2017-12-16 at the Wayback Machine, 1 November 2016 – 31 October 2017, Nominet
41. ^ ECHR 18 September 2007, no. 25379/04, 21688/05, 21722/05, 21770/05, *Paeffgen v Germany*.
42. ^ ***a b*** *Levine, John R. (April 21, 2019). "Domain Name Variants Still Won't Work". Archived from the original on July 29, 2020. Retrieved May 23, 2020.*
43. ^ *"Comment on ICANN Recommendations for Managing IDN Variant Top-Level Domains" (PDF). ICANN. April 21, 2019. Archived (PDF) from the original on 2022-10-09. Retrieved May 23, 2020.*
44. ^ "So This Manatee Walks Into the Internet Archived 2017-01-23 at the Wayback Machine", *The New York Times*, December 12, 2006. Retrieved April 12, 2008.
45. ^ *Allemann, Andrew (2019-11-05). "Part of MarkMonitor sold to OpSec Security". Domain Name Wire | Domain Name News. Retrieved 2024-11-26.*
46. ^ *"Canadian banks hit by two-year domain name spoofing scam". Finextra. 9 January 2020. Archived from the original on 6 November 2021. Retrieved 27 August 2021.*
47. ^ *"Domain spoofing". Barracuda Networks. Archived from the original on 2021-11-04. Retrieved 2021-08-27.*
48. ^ *Tara Seals (August 6, 2019). "Mass Spoofing Campaign Abuses Walmart Brand". threatpost. Archived from the original on November 6, 2021. Retrieved August 27, 2021.*
49. ^ *"Example Screenshots of Strider URL Tracer With Typo-Patrol". Microsoft Research. Archived from the original on 21 December 2008.*
50. ^ *"Internationalized Domain Names (IDN) in Google Chrome". chromium.googlesource.com. Archived from the original on 2020-11-01. Retrieved 2020-08-26.*
51. ^ *"Upcoming update with IDN homograph phishing fix - Blog". Opera Security. 2017-04-21. Archived from the original on 2020-08-08. Retrieved 2020-08-26.*
52. ^ *"About Safari International Domain Name support". Archived from the original on 2014-06-17. Retrieved 2017-04-29.*
53. ^ *"IDN Display Algorithm". Mozilla. Archived from the original on 2016-01-31. Retrieved 2016-01-31.*

External links

[edit]

Image not found or type unknown

Look up ***homograph*** in Wiktionary, the free dictionary.

Image not found or type unknown

Wikimedia Commons has media related to ***Domain name space***.

- (domain bias in web search) a research by Microsoft
- Top Level Domain Bias in Search Engine Indexing and Rankings
- Icann New gTLD Program Factsheet - October 2009 (PDF)
- IANA Two letter Country Code TLD
- ICANN - Internet Corporation for Assigned Names and Numbers
- Internic.net, public information regarding Internet domain name registration services
- Internet Domain Names: Background and Policy Issues Congressional Research Service
- RFC 1034, Domain Names — Concepts and Facilities, an Internet Protocol Standard
- RFC 1035, Domain Names — Implementation and Specification, an Internet Protocol Standard
- UDRP, Uniform Domain-Name Dispute-Resolution Policy
- Special use domain names

- **v**
- **t**
- **e**

Website management

Concepts

Web hosting

- Clustered
- Peer-to-peer
- Self-hosting
- Virtual

Web analytics

- Click analytics
- Mobile web analytics
- Web tracking
  - Click tracking

- Overselling
- Web document
- Web content
- Web content lifecycle
- Web server
- Web cache
- Webmaster
- Website governance

Web hosting control panels (comparison)

- AlternC
- cPanel
- DirectAdmin
- Domain Technologie Control
- Froxlor
- i-MSCP
- InterWorx
- ISPConfig
- Ispmanager
- Kloxo
- Plesk
- Usermin
- Webmin

Top-level domain registries

- AFNIC
- auDA
- DNS Belgium
- CentralNic
- CIRA
- CNNIC
- CZ.NIC
- DENIC
- EURid
- Freenom
- GoDaddy
- Google Domains
- Identity Digital
- IPM
- JPRS
- KISA
- NIC México
- Nominet
- PIR
- Tucows
- Verisign

Domain name managers and registrars

- Bluehost
- Domainz
- DreamHost
- Dynadot
- Enom
- Epik
- Gandi
- GlowHost
- GMO Internet
- GoDaddy
- Google Domains
- Hover
- Infomaniak
- Jimdo
- Name.com
- Namecheap
- Hostinger
- NameSilo
- NearlyFreeSpeech
- Network Solutions
- OVH
- Register.com
- Squarespace
- Tucows
- UK2
- Webcentral
- Web.com
- Wix.com

Web content management system

- Document management system
- Wiki software
- Blog software

Authority control databases: National Image not found or type unknown **Edit this at Wikidata**

- Germany
- United States
- France
- BnF data
- Japan
- Israel

Check our other pages :

- SEO company Sydney
- SEO expert Sydney
- local SEO specialists
- SEO agency
- SEO with wordpress

Frequently Asked Questions

How do I find the best SEO company in Sydney?

To find the best SEO company in Sydney, look for a provider with a proven track record of success, transparent reporting, and a clear understanding of your business•s goals. Check reviews, case studies, and client testimonials to ensure you are choosing a reputable partner.

What is the difference between local SEO and general SEO?

General SEO focuses on improving a website's visibility on a broader scale, often targeting national or international audiences. Local SEO, on the other hand, zeroes in on geographic areas, helping businesses attract nearby customers through local keywords, directory listings, and Google My Business optimization.

What should I expect from SEO agencies in Sydney?

SEO agencies in Sydney typically offer comprehensive services such as keyword research, technical audits, on-page and off-page optimization, content creation, and performance tracking. Their goal is to increase your site's search engine rankings and drive more targeted traffic to your website.

Why is keyword research important for SEO?

Keyword research helps identify the terms and phrases that potential customers are using to search for products or services. By targeting these keywords in your content, you can improve your visibility in search engine results, attract more qualified leads, and drive higher conversion rates.

What sets SEO specialists in Sydney apart?

SEO specialists in Sydney often have deep expertise in the local market. They understand the competitive landscape, know which keywords resonate with Sydney-based audiences, and are skilled at optimizing websites to rank well in local search results.

What is SEO?

SEO, or search engine optimisation, is the practice of improving a website's visibility on search engines like Google. It involves optimizing various elements of a site such as keywords, content, meta tags, and technical structure to help it rank higher in search results.

SEO companies in australia

SEO Sydney

Phone : 1300 684 339

City : Sydney

State : NSW

Zip : 2000

Google Business Profile

Google Business Website

Company Website : https://sydney.website/seo-sydney/

USEFUL LINKS

SEO Website

SEO Services Sydney

Local SEO Sydney

SEO Ranking

SEO optimisation

LATEST BLOGPOSTS

SEO community

SEO Buzz

WordPress SEO

SEO Audit

Sitemap

Privacy Policy

About Us

SEO Castle Hill | SEO Fairfield | SEO Hornsby | SEO Liverpool | SEO North Sydney | SEO Norwest | SEO Parramatta | SEO Penrith | SEO Strathfield | SEO Wetherill Park

Follow us