

Introduction

We find ourselves at a critical moment in the evolution of the semantic web. AI is becoming widely accessible and integrated into many of our day-to-day applications. Now, in its third year, this chapter provides a unique opportunity to analyze the past year's trends and examine the rapid developments occurring over time. Looking at the previous editions, we can offer a comprehensive view of where structured data stands today and where it's headed.

The Expanding Landscape of Structured Data

Over the past 18 months, there have been significant changes in the structured data landscape. In 2023, Google deprecated rich results for FAQs and HowTos from its search engine results pages (SERP). In November 2024 Google will also remove the Sitelinks Search Box from search results starting. However, in parallel, there has been a new wave of innovation and expansion in using structured data from both Google and Bing.

Key Developments in 2023-2024:

1. **New Structured Data Types:** Google introduced several new types, including Vehicle listings, Course info, Vacation Rentals, and 3D Models for products. Also on the eCommerce space Google has integrated loyalty programs into its structured data offerings, particularly through the Merchant Center and schema.org.
2. **Enhanced Existing Types:** Improvements to organization data, product variants, and the introduction of discount-rich results.
3. **Structured Data Carousels:** The beta launch of structured data carousels, combining ItemList with other types, opens new content presentation possibilities on Google's SERP.
4. **GS1 Integrations:** There has been increased support for GS1 standards such as the GS1 Digital Link, which aims to bridge the gap between physical and digital product information. This technology enables manufacturers and retailers to connect physical products to their digital identities through QR codes. When scanned, these codes provide access to comprehensive product information, enhancing transparency and customer engagement. Also the gs1:CertificationDetails property has been officially adopted by Google as schema:Certification, demonstrating how industry-specific extensions can successfully influence and integrate with schema.org standards.

Beyond Traditional Implementation

As the structured data ecosystem matures, we're witnessing a diversification in implementation strategies:

1. **Schema.org as Markup:** The traditional method of embedding structured data directly into webpages remains a cornerstone of modern SEO practices.
2. **Schema.org as a Data Standard:** Beyond markup, Schema.org is increasingly used to standardize data sent via APIs or feeds, enabling more dynamic and scalable

implementations. For instance, Google's Data Commons initiative leverages a superset of Schema.org entities to represent data from hundreds of organizations around the world in its knowledge graph.

3. **Digital Product Passports (DPPs)**: The rise of DPPs, especially in Europe, is transforming how product information is structured and shared. This development is expected to accelerate the adoption of eCommerce platforms in the coming years.

Structured Data in the Age of AI and Machine Learning

The rise of generative AI and advanced machine learning has further underscored the importance of structured data:

- **Fact Validation**: Structured data provides a reliable source for AI systems to verify information and combat misinformation.
- **Enhanced Search Understanding**: It enables a more nuanced interpretation of content by search engines and AI-powered systems.
- **Training Data**: Well-structured data is high-quality training material for machine learning models.

What This Chapter Provides

This chapter offers a data-driven analysis of structured data trends in 2023-2024, highlighting key developments and best practices:

1. **Evolution of the Landscape**:
 - Key shifts in structured data, especially with the rise of **AI-powered search** like **Google AI Overview** and **Bing Chat**.
 - Changes in **Google** and **Bing** structured data policies, and their impact on SEO.
2. **Prevalence and Growth**:
 - Trends in popular formats like **JSON-LD**, **Microdata**, and **RDFa**.
 - Adoption rates by schema types such as **Product**, **Organization**, and **Article**.
3. **Implementation and Best Practices**:
 - Best practices for structured data, including **JSON-LD** usage.
 - Common mistakes and how to avoid them.
4. **Rich Results & SERP Features**:
 - Effects of deprecated features like **FAQ** and **HowTo**.
 - Introduction of **carousels** and **Product Knowledge Panels**.
5. **AI-Powered Search**:
 - The role of structured data in **AI-driven search** and **voice assistants**.
 - Trends in **AI-powered content discovery**.
6. **E-commerce Innovations**:
 - Growth of **Digital Product Passports** and **GS1 Digital Links**.
 - Structured data's role in **e-commerce** and new rich result types.
7. **Knowledge Graphs & Graph RAG**:

- The rising importance of **knowledge graphs** and **Graph RAG** for enhancing AI outputs.
8. **Quality & Data Integrity:**
 - Best practices for maintaining high-quality structured data.
 9. **Emerging Schemas & Use Cases:**
 - Innovations in schema types and their application in **SEO** and **e-commerce**.
 10. **Future Outlook:**
 - The evolving role of structured data in **AI**, **semantic SEO**, and **content discovery**.

This chapter provides a comprehensive view of structured data's impact on SEO, AI, and e-commerce, with actionable insights for developers and marketers.

Key Concepts

As structured data evolves in complexity, exploring and explaining key concepts is crucial before diving into a deeper analysis. This section outlines fundamental ideas and recent developments in the field.

Linked Data and the Semantic Web

Linked data remains a cornerstone of structured data. We create an interconnected web of information by adding structured data to web pages and providing URI links to referenced entities. This contributes to the semantic web, where data is linked through the Resource Description Framework (RDF), enabling machines to treat web pages as databases.

The concept of semantic triples (subject-predicate-object) continues to be fundamental in expressing relationships between entities. While [SPARQL](#) remains helpful for querying RDF data, the focus has shifted towards more accessible ways of leveraging this linked data structure, such as [GraphQL](#).

Open Data and the 5 Stars Model

Tim Berners-Lee's 5 stars of the open data model remain relevant. It emphasizes the importance of web-available, structured, non-proprietary, URI-identified, and interlinked data. Structured data plays a crucial role in achieving higher levels of this model, contributing to a more open and interconnected web ecosystem.

AI-Powered Search, Voice Assistants, and Digital Assistants

The landscape of search and digital assistance has dramatically evolved with the integration of AI, LLMs, and advanced natural language processing. This convergence has blurred the lines between traditional search engines, voice-activated systems, and AI-powered digital assistants.

Semantic Search Engines and AI-Powered Search

Semantic search has progressed beyond traditional keyword matching to include sophisticated AI-powered experiences. These systems leverage structured data to provide more accurate, contextual, and often conversational search results. Key developments include:

- **Google AI Overview:** A feature that provides comprehensive (sometimes misleading) AI-generated summaries on complex topics.
- **Microsoft Bing Chat:** Integrates chat-based AI interactions directly into Bing search results.
- **Meta AI:** Meta's AI assistant is integrated across platforms like Facebook, Messenger, Instagram, and WhatsApp.
- **SearchGPT (and ChatGPT):** OpenAI's AI search engine that integrates search results into conversational responses.
- **Perplexity.ai:** An AI-powered search engine that provides detailed, sourced answers to queries.
- **You.com:** Offers AI-summarized search results and a chat interface for more interactive searching.

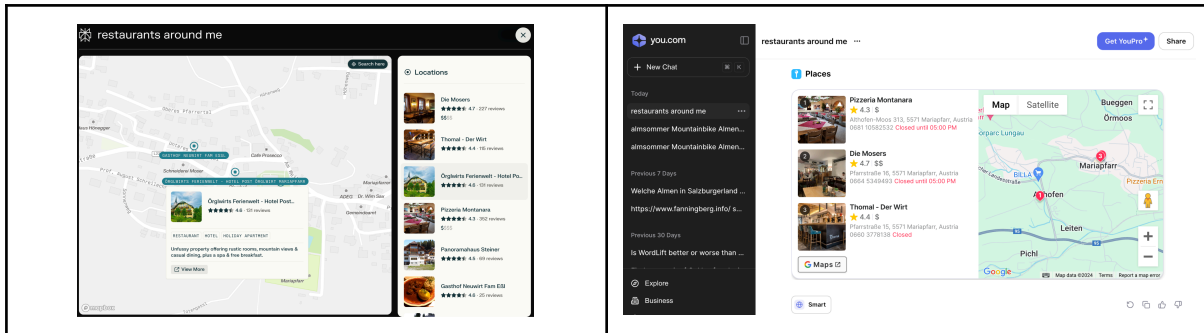
These platforms demonstrate an enhanced ability to understand user intent and context, significantly improving search accuracy and user experience. They often combine traditional web indexing with real-time information retrieval and natural language generation.

The Role of Structured Data

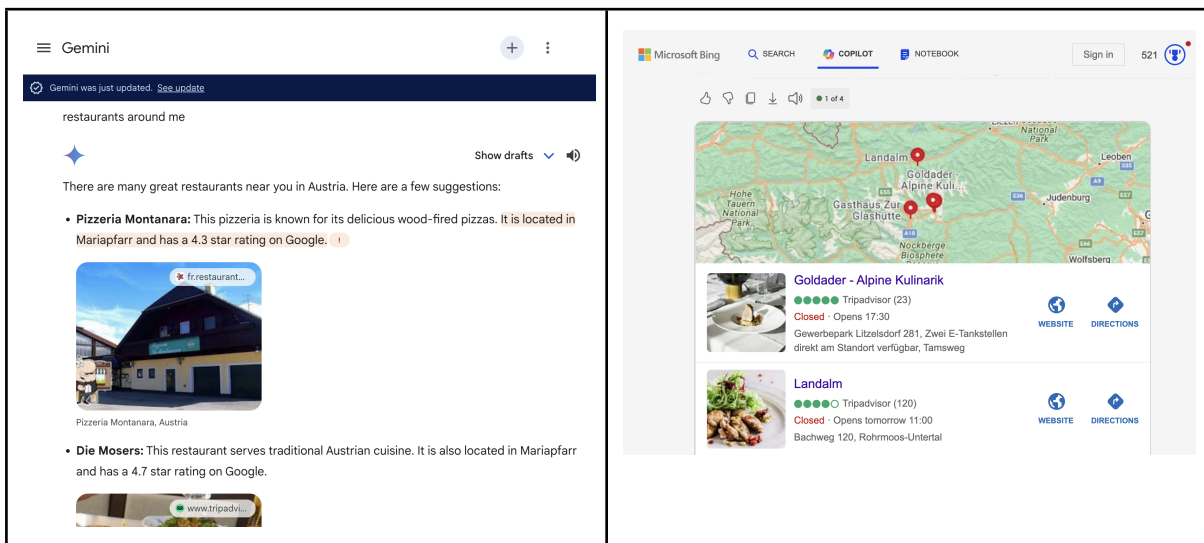
Structured data plays a crucial role in these AI-powered systems by:

1. **Enhancing Entity Recognition:** Helping systems accurately identify and disambiguate entities mentioned in queries.
2. **Providing Context:** Offering additional information about entities and their relationships, improving response accuracy.
3. **Facilitating Knowledge Graph Integration:** Allowing these systems to tap into vast, interconnected information databases.
4. **Enabling Rich Responses:** Supporting the generation of detailed, multi-faceted answers that often include visual elements or interactive features.
5. **Improving Voice Query Interpretation:** Assisting in understanding the intent behind spoken queries, which can be more ambiguous than text-based searches.

While it is still challenging to assess the impact of structured data on Generative AI and AI search engines, in some cases, such as geo-referencing queries, we can observe the early emergence of entities in the user experience of Perplexity.ai and You.com.



This is way more consistent when interacting with Bing Copilot or Gemini by Google.



Empirically, AI-powered search systems, as seen above, are sourcing data from a variety of established knowledge bases and authoritative platforms:

- Map services: Google Maps and Bing Maps serve as crucial data sources for location-based information.
- Authoritative websites: Platforms rich in structured data markup, such as TripAdvisor, contribute significantly to the knowledge base of AI search systems.
- Vertical-specific databases: Industry-specific databases and platforms provide specialized information for AI-powered search in various sectors.

The Shift to AI-Powered Search and Its Implications

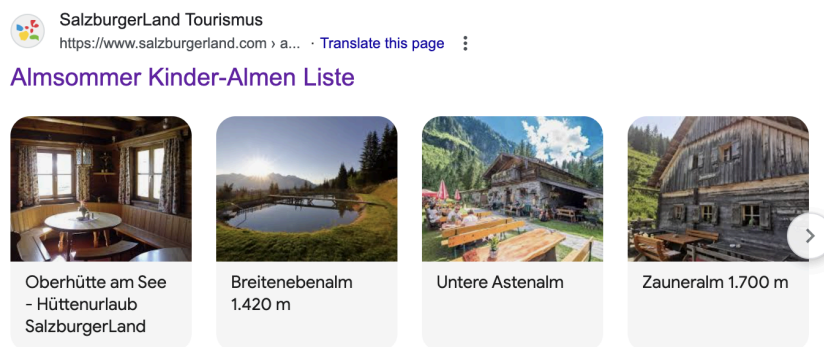
This transition from traditional search to AI-powered search demands a broader, more nuanced approach to optimization:

1. **Multi-platform visibility:** SEO strategies must now account for visibility across a diverse array of AI surfaces and platforms, including:
 - Traditional search engines (Google, Bing)
 - AI chatbots (ChatGPT, Google's Gemini, Perplexity, Anthropic's Claude)
 - Integrated assistants (Microsoft Copilot, potential Apple-ChatGPT integration)
 - Ecosystem-specific tools (Google Workspace, Microsoft 365)
 - Browser and device-level integrations

2. **Beyond conventional optimization:** Success in this landscape goes beyond optimizing for specific features like Google's AI Overview. It requires a holistic approach to making content discoverable and comprehensible across all emerging search interfaces.
3. **Leveraging structured data strategically:** The key to improved visibility lies not just in publishing structured data using schema markup but in facilitating access to structured information about entities that matter to your business or content. This involves:
 - Ensuring clear, structured information is available and easily interpretable by various AI systems.
 - Ensuring that the metadata used to describe the webpage for bots is consistent with the content presented to human readers.
 - Directly feeding accurate information to relevant platforms and marketplaces (e.g., Google Merchant, Amazon) for products and services.

Rich Results and Knowledge Panels

Rich results and knowledge panels, powered by structured data, are essential features of search engine results pages (SERPs). These enhanced displays offer users immediate and relevant information, significantly boosting click-through rates and user engagement. As rich results become more diverse and sophisticated, they present new opportunities for content visibility. A recent example from Google is the introduction of a structured data carousel for listicle pages related to local businesses (including subtypes like restaurants, hotels, vacation rentals), products, and events.



An example of the new beta carousel rich result.

Another notable example, while not directly influenced by structured data, is the new **Google Merchant Knowledge Panel**, which extends the functionality of the **Product Knowledge Graph Panel**. Structured data acts as a signal that contributes to entity disambiguation, helping search engines accurately identify businesses and their attributes, which can lead to the appearance of these panels. This feature helps businesses, both small and large, build trust with users by displaying key information about the merchant directly on Google's search results page.

Knowledge Graphs and Graph RAG

Knowledge graphs have become increasingly central to structured data applications. They provide a powerful way to represent and query complex relationships between entities. The emergence of [Graph RAG](#) (Retrieval-Augmented Generation) represents a significant advancement, combining knowledge graphs with large language models to enhance AI-generated responses with factual, structured information.

Difference Between Labeled Property Graphs and RDF Graphs

Labeled Property Graphs (LPGs) and Resource Description Framework (RDF) graphs are two distinct approaches to organizing and representing data. LPGs, commonly used in databases like Neo4j, structure data with nodes and relationships, each carrying labels and properties. This allows for a flexible and intuitive way to model complex data relationships. On the other hand, RDF graphs, which are foundational to the Semantic Web, use a triple-based structure (subject-predicate-object) to represent data. RDF emphasizes interoperability and standardization, making it ideal for linking data across different systems and domains. While LPGs offer ease of use and performance for certain applications, RDF provides a robust framework for semantic data integration and reasoning.

The importance of structured data in creating knowledge graphs cannot be overstated. **Structured data enables the precise definition of entities and their relationships**, which is crucial for the development of accurate and reliable knowledge graphs. By leveraging structured data, organizations can build comprehensive knowledge graphs that enhance data discoverability, interoperability, and the overall quality of AI-generated insights.

Data Commons

[Data Commons](#) is an open-source and open-data initiative by Google that organizes public datasets from various global sources, such as the United Nations and national census bureaus, to make them universally accessible. The platform provides over 250 billion data points and 2.5 trillion triples, encompassing a wide range of statistical variables. Schema.org is utilized to encode structured data in Data Commons, creating a unified knowledge graph that standardizes and normalizes diverse datasets, enabling easier access and exploration through a common framework. This structured approach helps to integrate vast amounts of data into a coherent, searchable system.

Digital Product Passports and GS1 Digital Link

In the e-commerce and supply chain sectors, [Digital Product Passports](#) (DPPs) and the [GS1 Digital Link standard](#) are revolutionizing how product information is shared and accessed. These technologies leverage structured data to create comprehensive, easily accessible digital representations of physical products, enhancing traceability, sustainability efforts, and consumer information access.

AI, Machine Learning, and Structured Data

The synergy between structured data and AI/ML has deepened. Structured data is crucial in training machine learning models, providing consistent, machine-readable labels. It's particularly important in areas such as:

- **Large Language Models (LLMs):** Fine-tuning with structured data for improved performance in specific domains.
- **Explainable AI:** Using knowledge graphs to trace and explain AI decision-making processes.
- **Multimodal AI:** Linking different data types (text, images, video) in AI systems.

Semantic SEO and Data Quality

Semantic SEO has evolved from focusing purely on keyword matching to leveraging structured data for deeper contextual understanding, which is now essential for search engines to provide more accurate and relevant results. It allows websites to build thematic depth and meaning into their content by using entities and structured metadata, enhancing the ability of search engines like Google and Bing to grasp the intent behind queries rather than just focusing on keyword frequency.

By implementing semantic SEO, businesses can create content clusters based on topics, not just individual keywords, making their content more discoverable and contextually relevant across various search platforms, including voice search assistants. This approach significantly boosts search engine rankings and user engagement, as structured data allows search engines to understand the content at a more granular level, making it easier to match user intent.

Data quality plays a key role here as well. High-quality structured data ensures consistency and accuracy, which is crucial not only for search engines but also in combating misinformation. It helps maintain trustworthiness across the web, especially as structured data is increasingly used in AI-powered systems like knowledge graphs for fact validation and enhancing large language model (LLM) training.

For example, organizations such as EssilorLuxottica, Shiseido, and others are using semantic technologies like knowledge graphs to link content and provide users with more detailed, contextually relevant results. This practice also aids in AI-powered content discovery and makes content easier to retrieve through Generative Search like Perplexity or You.com.

Investing in semantic SEO and maintaining high-quality structured data not only enhances search visibility but also lays a foundation for future-proofing content for AI-driven discovery.

A Year In Review

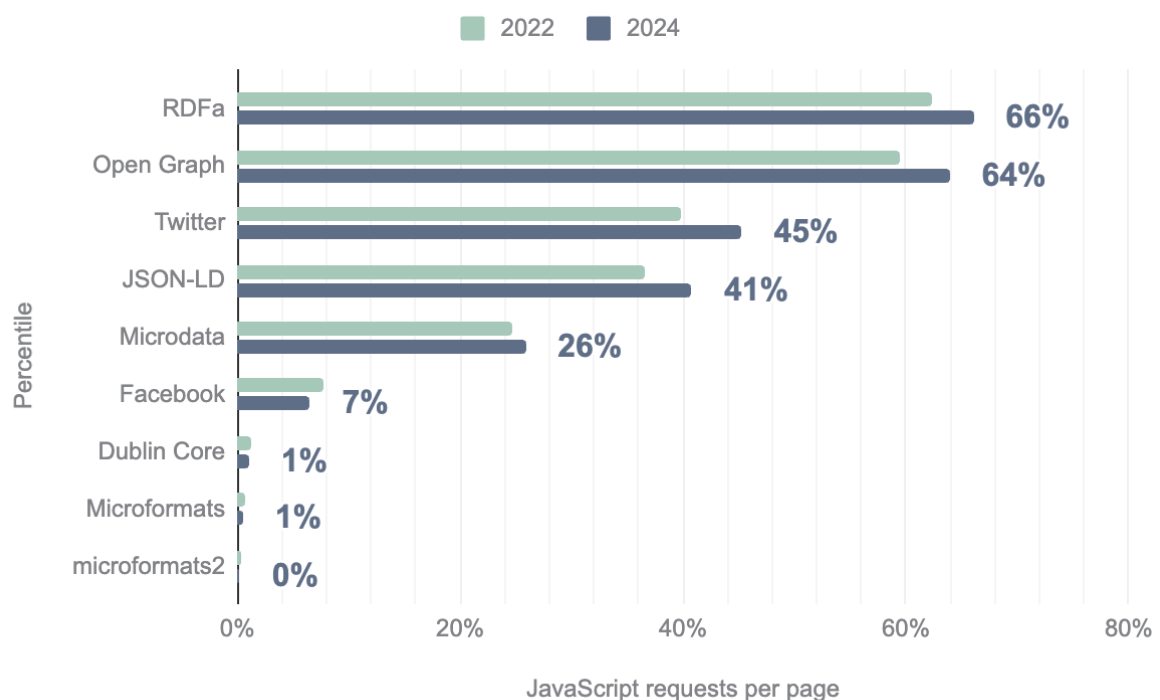
Structured Data

The landscape of structured data implementation continues to evolve, with RDFa and Open Graph maintaining their dominant positions, now reaching 66% and 64% of pages respectively. X (Twitter) meta tags have shown significant growth, appearing on 45% of pages, while JSON-LD usage has expanded to 41% of pages, reflecting its growing adoption for structured data implementation.

Microdata appears on 26% of pages, showing steady usage despite being a less popular format nowadays. The lower tier of implementations includes Facebook meta tags at 7%, while Dublin Core and Microformats maintain minimal presence at around 1% of pages. Notably, microformats2 shows the lowest adoption at less than 1%, despite its previous growth spikes in earlier years.

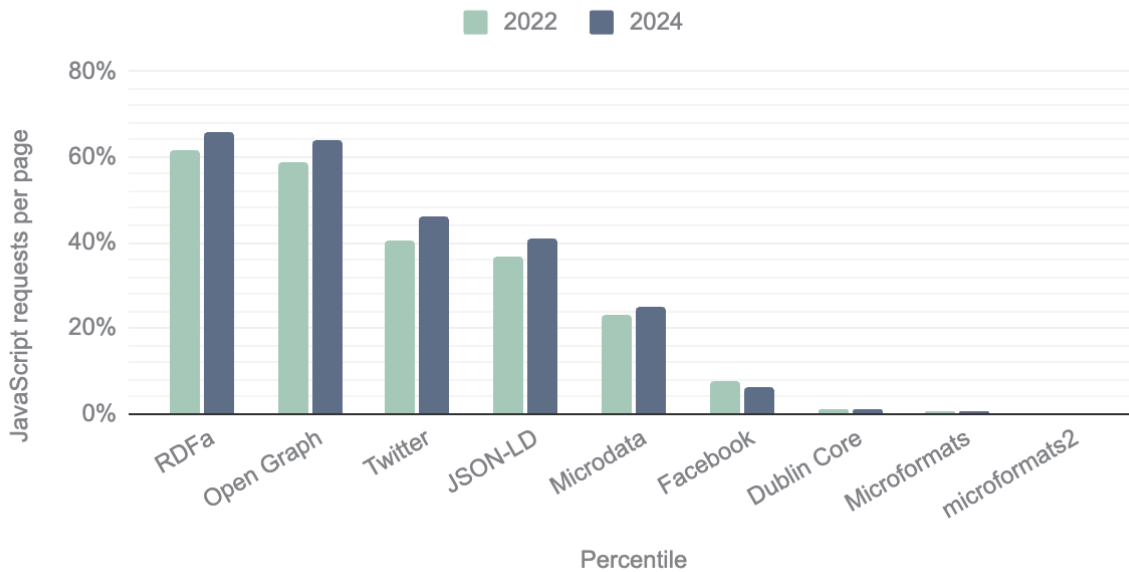
Structured data usage by year

Web Almanac 2024: Structured Data (mobile)



Structured data usage by year

Web Almanac 2024: Structured Data (desktop)



The data shows interesting platform differences between desktop and mobile implementations, though these variations have become less pronounced compared to previous years, suggesting a more standardized approach to structured data implementation across device types. This evolution aligns with the industry's movement toward more consistent cross-platform structured data strategies, particularly as search engines and AI systems increasingly rely on this data for understanding and presenting content.

Structured Data Usage Trends (2022-2024): The data shows notable changes across various structured data formats over the two-year period. RDFa has seen an increase from around 60% to 65% of desktop pages, reinforcing its position as the most widely adopted format most probably driven by templates on legacy CMSs. Open Graph maintains strong adoption with a slight increase, remaining the second most prevalent format at approximately 62%. X meta tags have shown substantial growth, climbing from about 35% to 45% of pages. JSON-LD continues its upward trajectory, increasing from roughly 34% to 41% of pages, reflecting its growing popularity among developers and SEO professionals.

Among the less widely used formats, Microdata shows modest growth to about 25% adoption, while Facebook meta tags have experienced a decline, dropping to around 7% from their previous levels. Dublin Core, Microformats, and microformats2 remain at minimal usage levels, with little change over the period. This pattern suggests a consolidation around the major formats (RDFa, Open Graph, Twitter, and JSON-LD) while smaller formats either stagnate or decline in usage.

These trends reflect the industry's growing maturity in structured data implementation, with clear preferences emerging for certain formats based on their utility, ease of implementation, and support from major platforms.

Comparison of JSON-LD, Microdata, and RDFa usage

The three main structured data formats show distinctly different adoption patterns:

JSON-LD	RDFa	Microdata
<ul style="list-style-type: none">● Present on 41% of pages (up from 34% in 2022)● Growing fastest among the three formats● Preferred by Google and gaining wider developer adoption● Most commonly used for:<ul style="list-style-type: none">○ Organization data○ Local business information○ Product listings○ Articles and creative works	<ul style="list-style-type: none">● Highest adoption at 66% of pages● Most prevalent on legacy CMS platforms● Common implementations:<ul style="list-style-type: none">○ Navigation elements (breadcrumbs)○ Basic page structure○ Image and document metadata○ List items	<ul style="list-style-type: none">● Present on 26% of pages● Showing steady but slower growth● Primarily used for:<ul style="list-style-type: none">○ Webpage structure (8.34% of pages)○ Site navigation (6.42%)○ Headers and footers (5.97% and 5.33%)○ Organization information (4.87%)

Let's analyze now more in detail each type.

RDFa

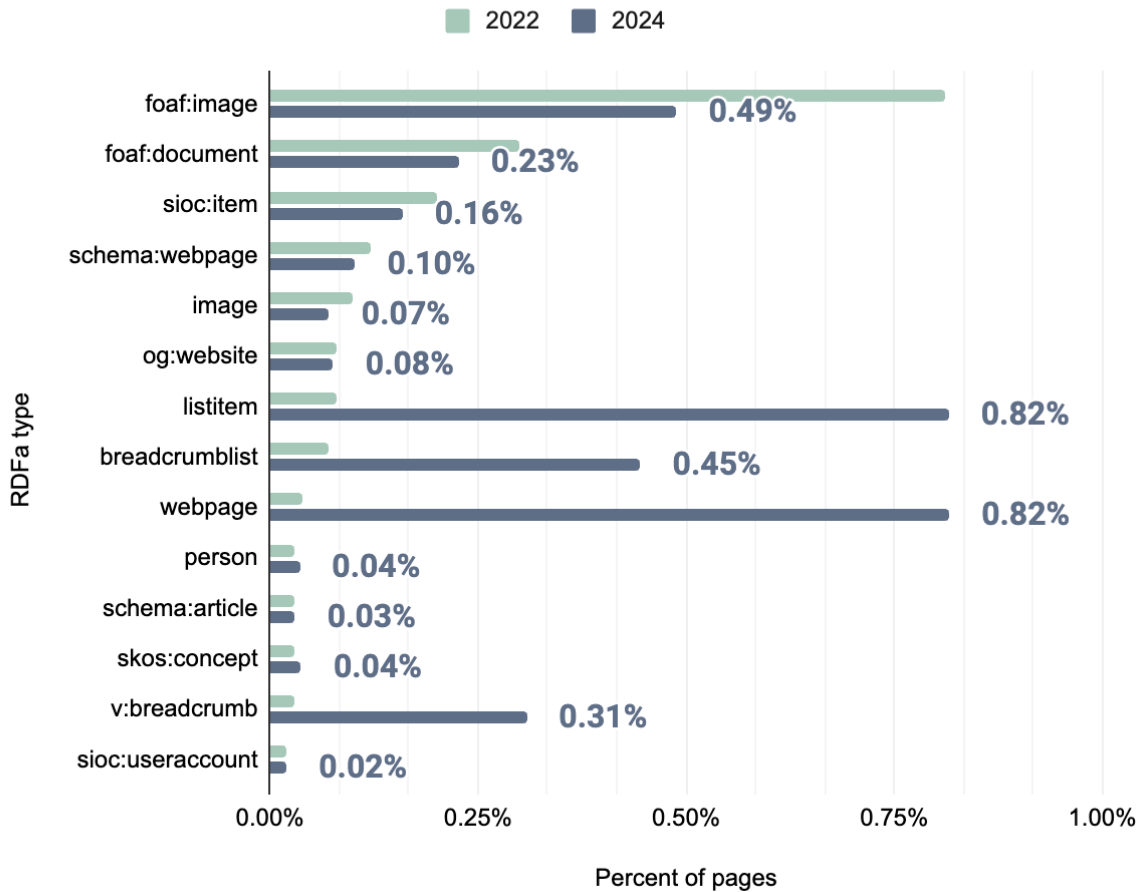
RDFa continues to play a significant role in structured data, particularly within legacy CMS platforms. However, there has been a noticeable shift towards using RDFa for **navigation elements**, such as `listitem` and `breadcrumblist`, which are now prevalent on a significant portion of web pages. This reflects an industry-wide emphasis on enhancing **structured navigation data** for better user experience, particularly on mobile platforms.

In contrast, traditional RDFa types like `foaf:image` and `foaf:document` have seen declining usage, as newer formats like **JSON-LD** and **Open Graph** offer more flexible solutions for image and document metadata. The adoption of **schema.org types** within

RDFa, such as `schema:webpage`, has shown modest but stable growth, further indicating a shift towards **Schema.org vocabularies**.

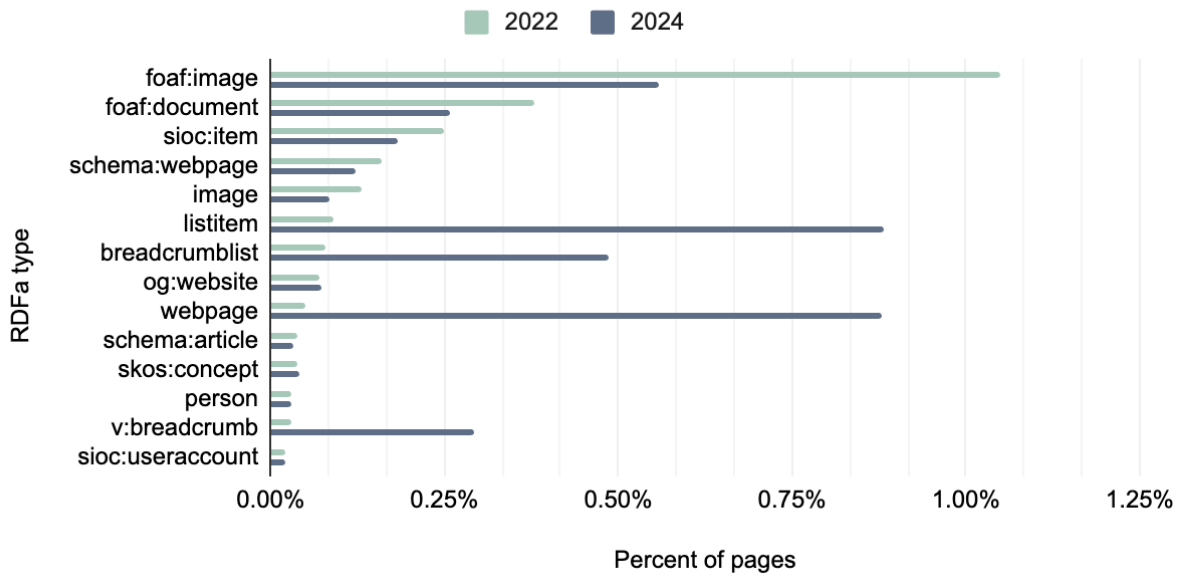
RDFa usage by year

Web Almanac 2024: Structured Data (mobile)



RDFa usage by year - Desktop

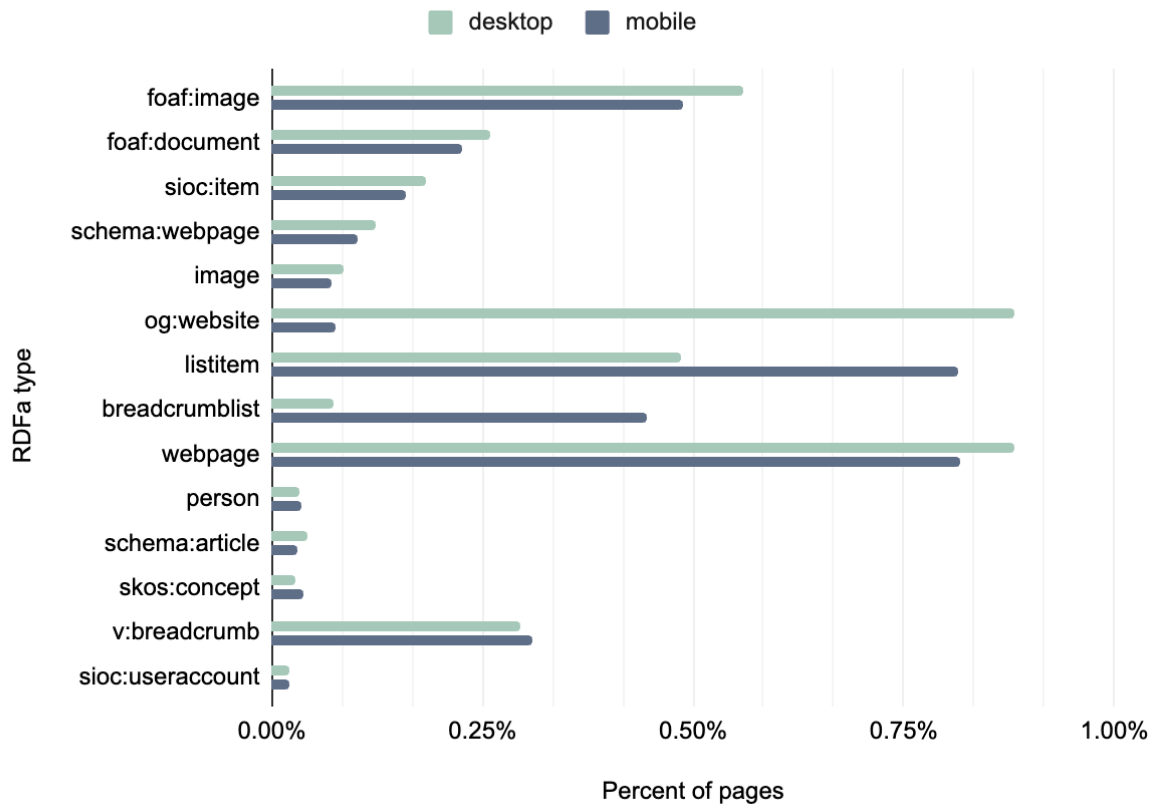
Web Almanac 2024: Structured Data



The data suggests that while RDFa remains a valuable tool, its dominance is gradually being overtaken by modern structured data formats like JSON-LD, particularly in dynamic content applications.

RDFa usage by device

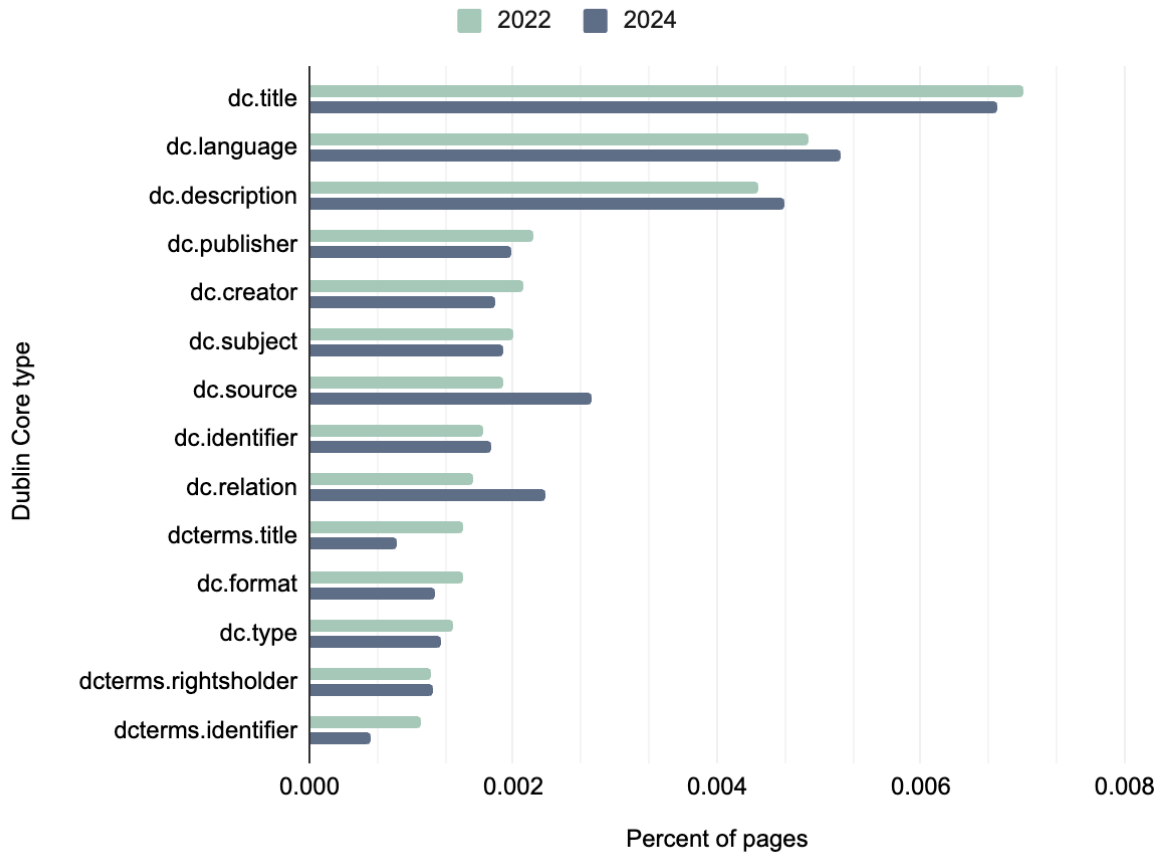
Web Almanac 2024: Structured Data



Dublin Core

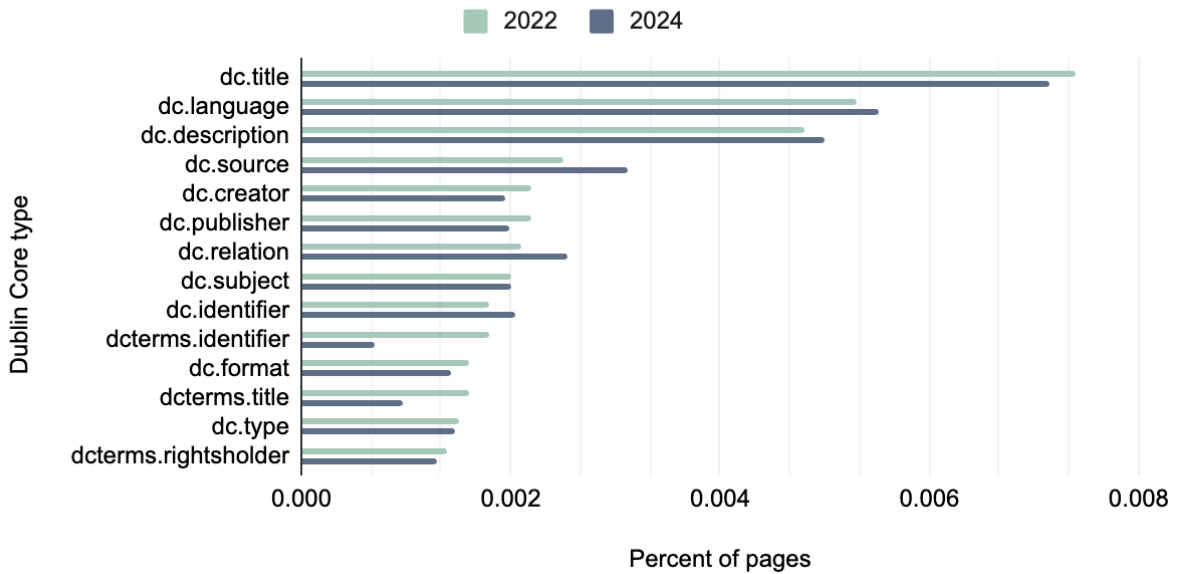
Dublin Core usage by year

Web Almanac 2024: Structured Data (mobile)



Dublin Core usage by year - Desktop

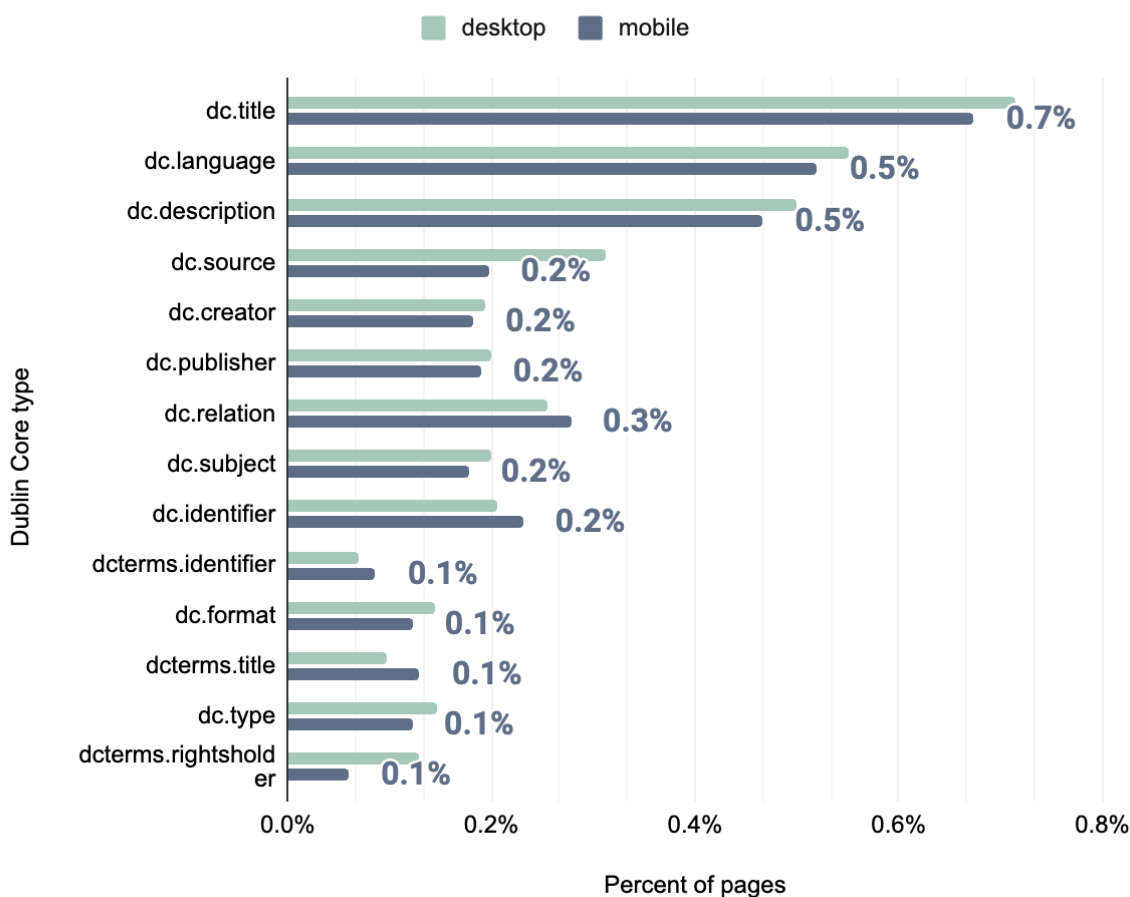
Web Almanac 2024: Structured Data



Dublin Core remains a stable but less frequently used format for metadata, especially when compared to modern formats like JSON-LD and Open Graph.

Dublin Core usage by device

Web Almanac 2024: Structured Data



Its key fields, such as `dc.title` and `dc.language`, show **minimal year-on-year changes**, maintaining a consistent presence primarily in **academic and legacy web projects**.

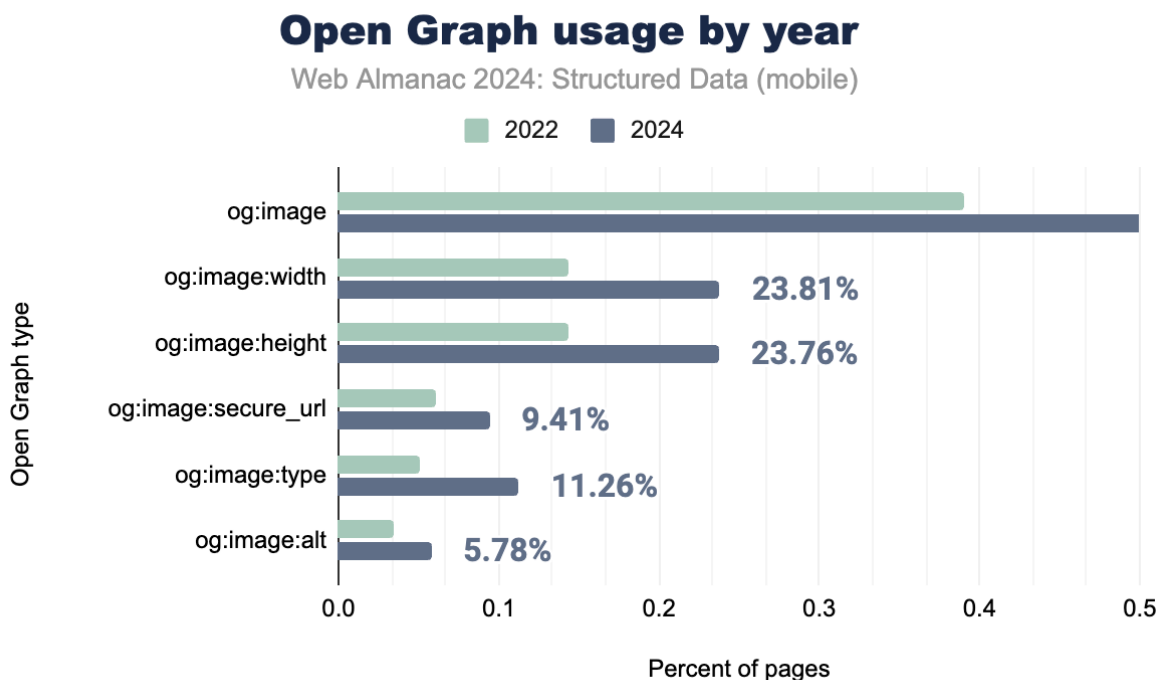
An increase in the use of `dc.source` reflects a growing emphasis on **citing original sources**, while fields like `dc.identifier` continue to be crucial for **resource identification**. However, specialized fields such as `dcterms.identifier` have seen declining adoption, signaling that Dublin Core is less central in today's web environments.

Interestingly, **Dublin Core** retains relevance in **multilingual document management**, particularly through the `dc.language` field, which is essential for managing and categorizing content in multiple languages. This makes it a valuable tool in contexts where **document metadata** needs to support internationalization and localization efforts.

Overall, while Dublin Core is being gradually outpaced by more versatile formats like JSON-LD, it continues to serve **niche needs** where **structured document metadata** and **multilingual support** are critical.

Open Graph

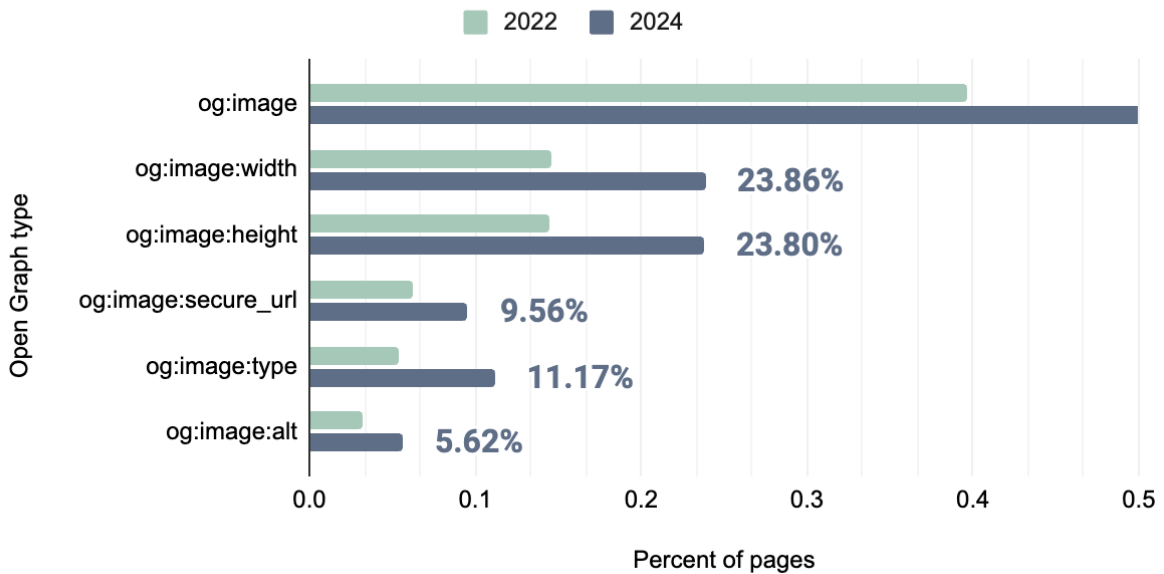
Open Graph continues to be one of the most widely implemented structured data formats, particularly in the context of **social media sharing**. The `og:image` tag remains the most frequently used property, reflecting the growing emphasis on **visual content optimization**. Other image-related tags, such as `og:image:width` and `og:image:height`, have also seen a steady increase in adoption as websites strive to enhance the **presentation of shared content** across platforms.



A key development in 2024 is Google's update to its search documentation, now including the `og:title` meta tag as a source for generating **title links** in search results. This update allows Google to consider the `og:title` tag alongside traditional sources, such as the HTML `<title>` tag, when determining how clickable titles are displayed in search results. As a result, the `og:title` tag has gained renewed significance, not only for **social media visibility** but also for **SEO**.

Open Graph usage by year - Desktop

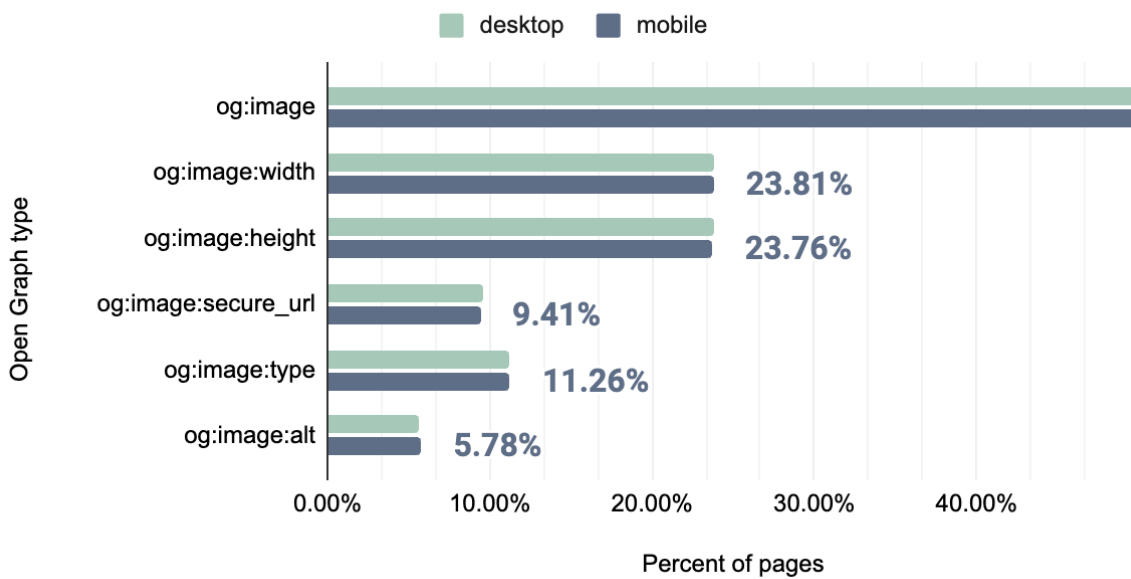
Web Almanac 2024: Structured Data



This dual role of Open Graph in **social sharing** and **search engine optimization** makes it a critical tool for webmasters looking to improve both **user engagement** on social platforms and **visibility** in search results.

Open Graph usage by device

Web Almanac 2024: Structured Data



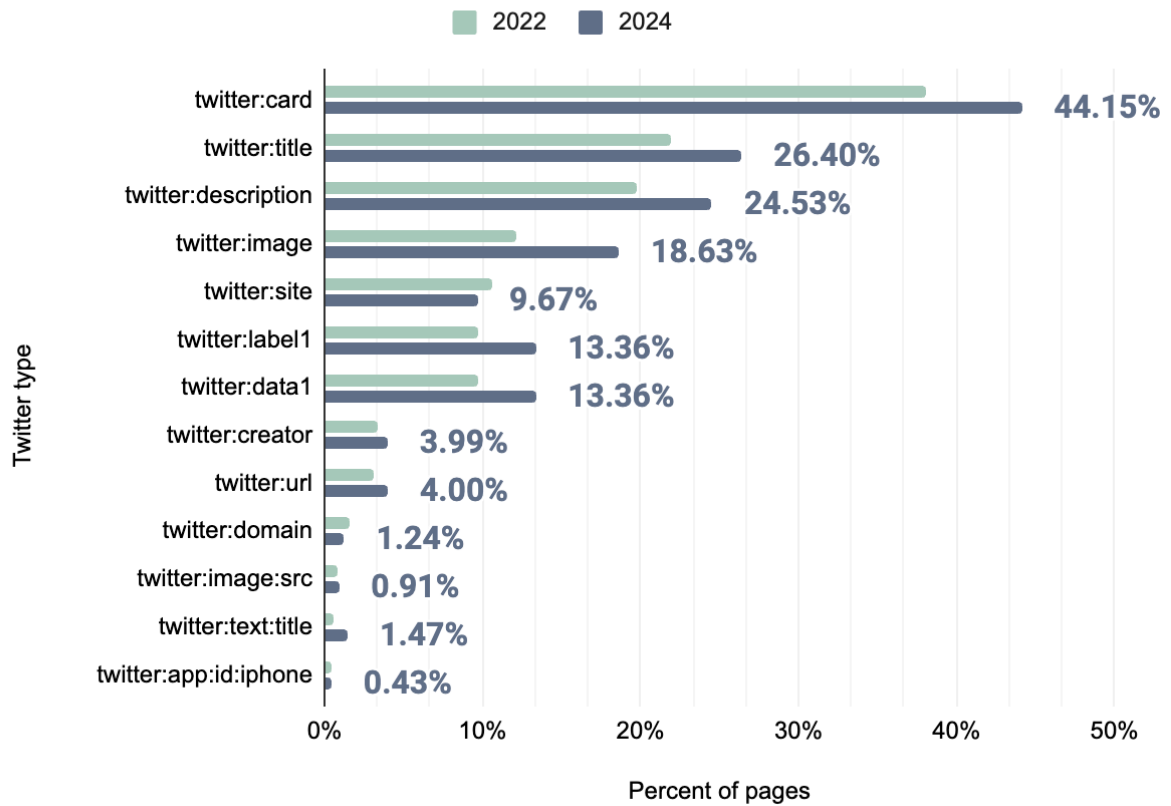
Security and type-related properties have also gained traction. The `og:image:secure_url` property, which ensures image URLs are served over **secure HTTPS connections**, has increased to **9.41% on mobile** and **9.56% on desktop**. Similarly, `og:image:type`, which specifies the MIME type of the image, has grown to **11.26% on mobile** and **11.17% on desktop**. These properties help ensure consistent and secure media delivery across devices and platforms.

Twitter

Despite the platform's transition to **new ownership** and its rebranding as **X**, Twitter's meta tags remain a vital part of the **structured data landscape**, particularly in the realm of **social media optimization**. The `twitter:card` tag continues to dominate, showing significant growth across mobile and desktop pages, as it plays a key role in defining how content is displayed when shared on the platform.

Twitter meta tag usage by year

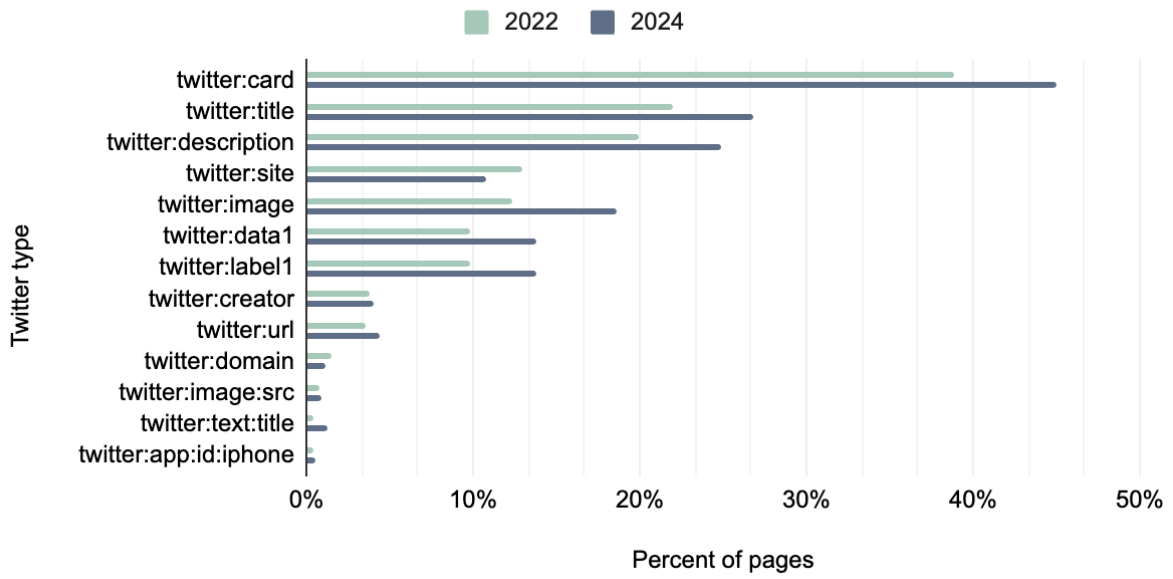
Web Almanac 2024: Structured Data (mobile)



Core descriptive tags like `twitter:title` and `twitter:description` have also seen widespread adoption, appearing on approximately **26% of mobile pages** and **24% of desktop pages**. These tags are essential for **content previews**, enhancing how web pages appear when shared on **social media**, and ensuring key information is highlighted.

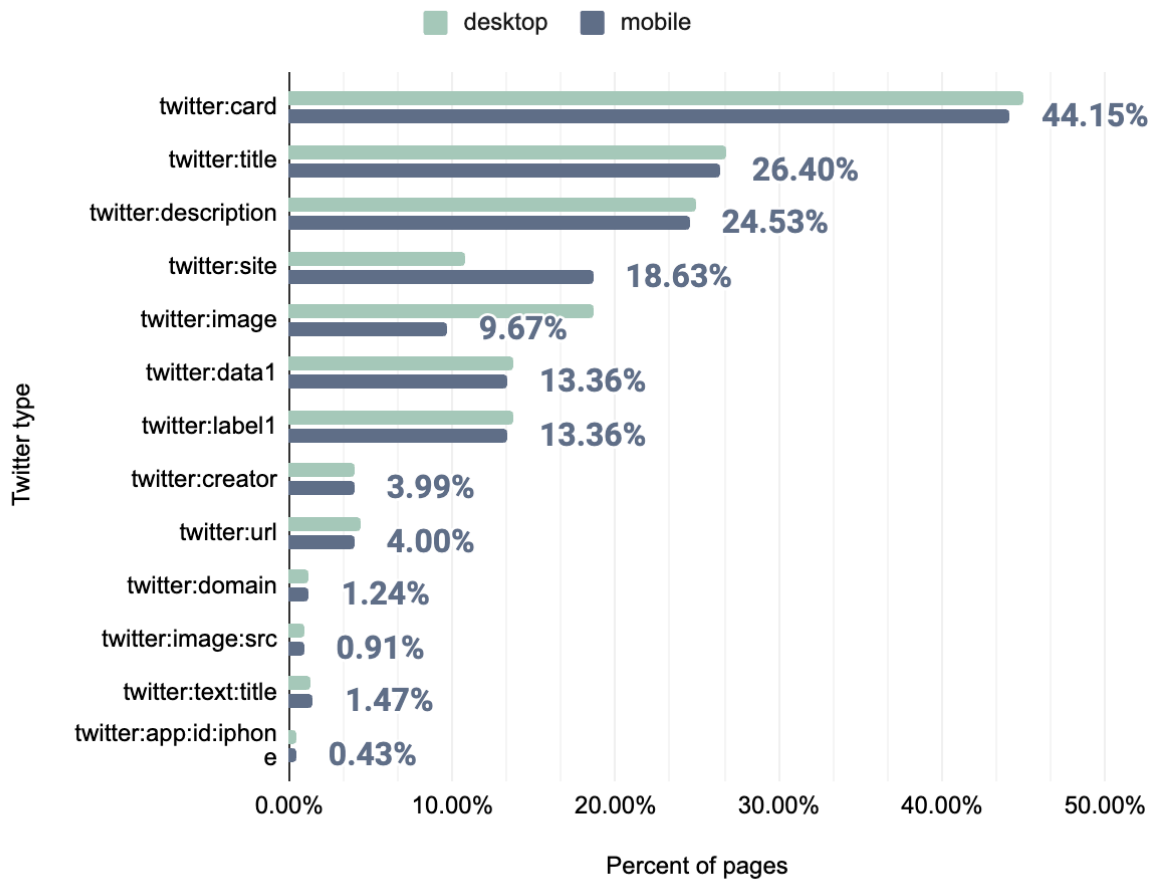
Twitter meta tag usage by year

Web Almanac 2024: Structured Data (desktop)



Twitter meta tag usage by device

Web Almanac 2024: Structured Data



The platform's **enhanced metadata** properties, such as `twitter:data1` and `twitter:label1`, which support **rich card features**, have seen coordinated growth, now appearing on **13.36% of mobile pages**. This indicates the increasing use of **Twitter Cards** for more detailed content representations, such as for **product listings** or **event details**.

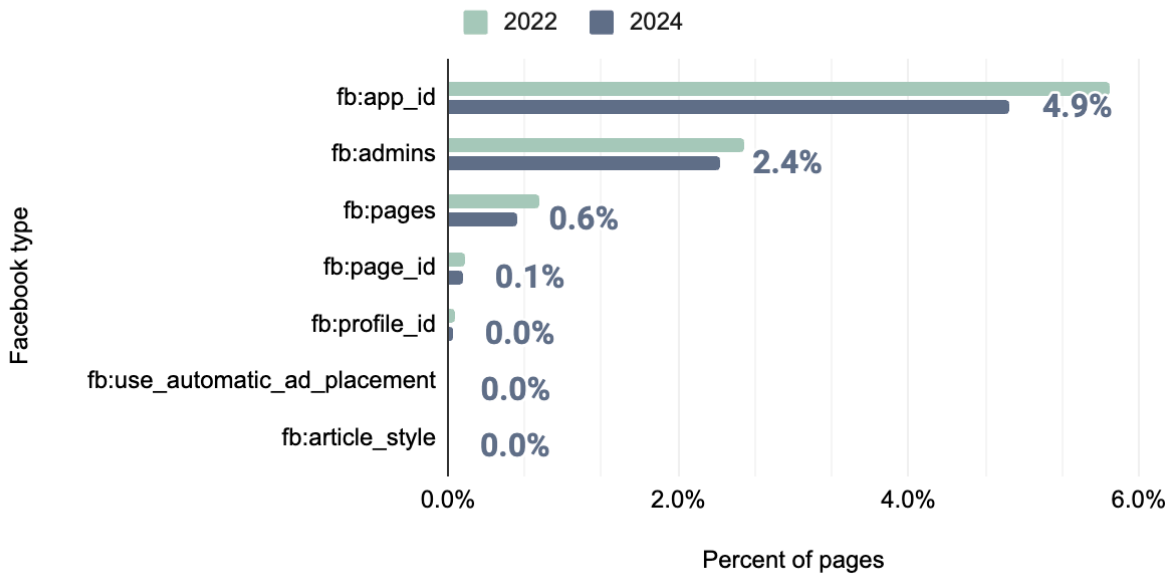
While **X** has undergone major branding changes, the **metadata architecture** it introduced remains critical for webmasters and SEO professionals, ensuring content shared on social media is **engaging**, **informative**, and **optimized** for interaction. This highlights the platform's enduring importance in the **social media and metadata ecosystem**.

Facebook

Facebook-specific meta tags have seen a marked decline in usage between 2022 and 2024, reflecting the broader industry shift toward **Open Graph** as the preferred format for **social sharing metadata**. The `fb:app_id` tag, once widely used to integrate apps with the Facebook platform, now appears on only **4.9% of mobile pages**, down from previous years. Similarly, administrative tags like `fb:admins` have dropped to just **2.4%**, serving primarily for backend management rather than enhancing content visibility.

Facebook meta tag usage by year

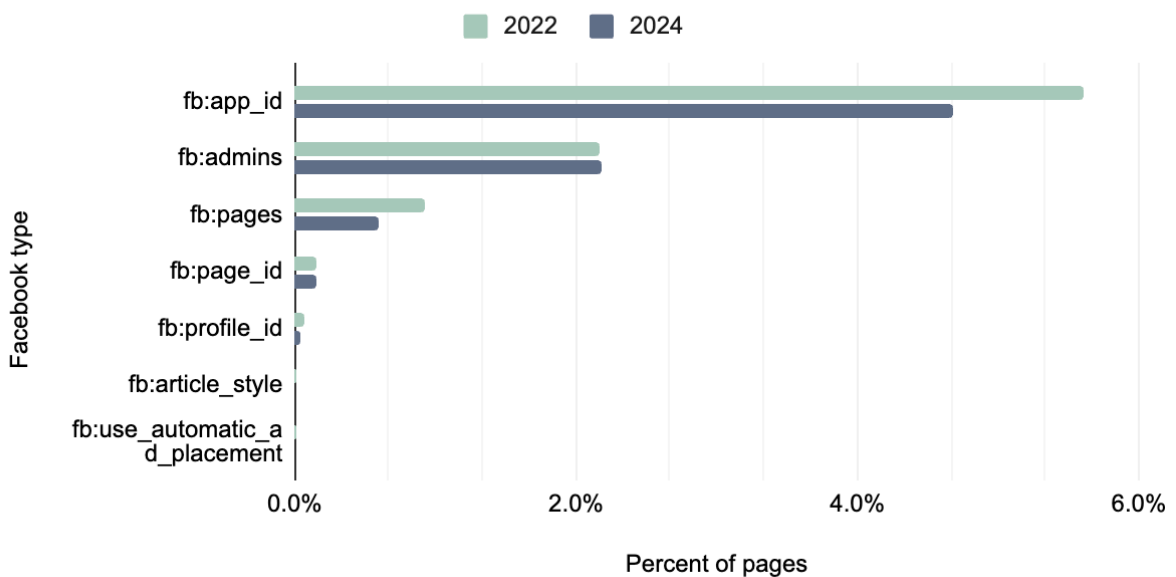
Web Almanac 2024: Structured Data (mobile)



This decline underscores a strategic move by developers and webmasters to adopt **Open Graph**, which originated with Facebook but has since become the **standard for social media sharing** across platforms. The **Open Graph format** offers greater flexibility and interoperability, making it the go-to choice for content optimization on Facebook as well as other social networks.

Facebook meta tag usage by year

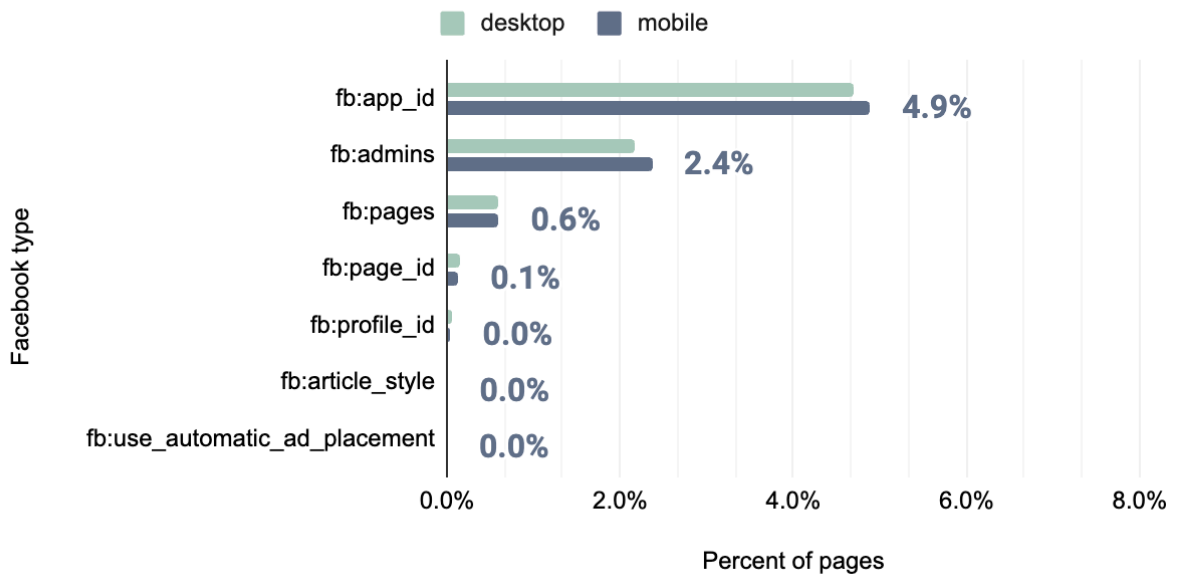
Web Almanac 2024: Structured Data (desktop)



Despite the decreasing adoption of Facebook-specific tags, **Facebook** itself remains a key player in the **social media landscape**, with Open Graph handling most of its metadata needs. This trend reflects the consolidation of social sharing standards, where **platform-agnostic tags** provide greater reach and functionality.

Facebook meta tag usage by device

Web Almanac 2024: Structured Data



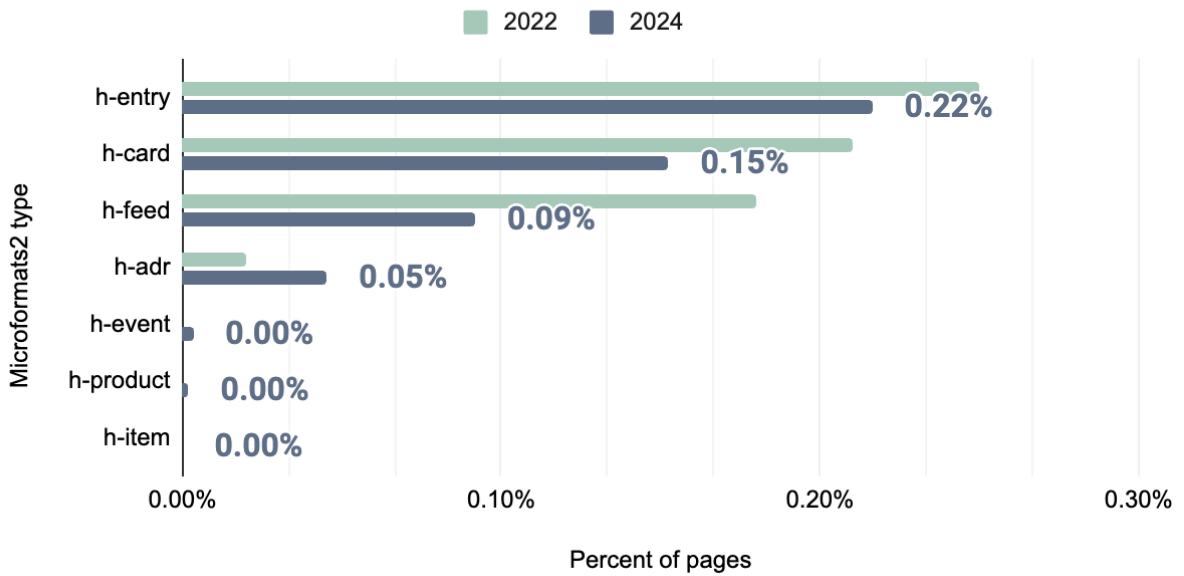
Microformats and Microformats2

Microformats continue to show **limited adoption**, primarily in niche use cases where simple, semantic data is required. The **adr** tag, used for address-related data, remains the most widely adopted Microformats type, appearing on approximately **0.4% of pages** across both mobile and desktop platforms. Other tags, such as **geo** and **hReview**, have minimal usage, as more sophisticated formats like JSON-LD and Open Graph have become more prevalent.

Microformats2, while still relatively niche, has seen slightly higher adoption than its predecessor. Tags like **h-entry** and **h-card**, which are used for blogging and personal identity data, now appear on **0.22% of mobile pages** and **0.15% of desktop pages**. These tags continue to serve specific needs, particularly for **address data** and **simple content structures**.

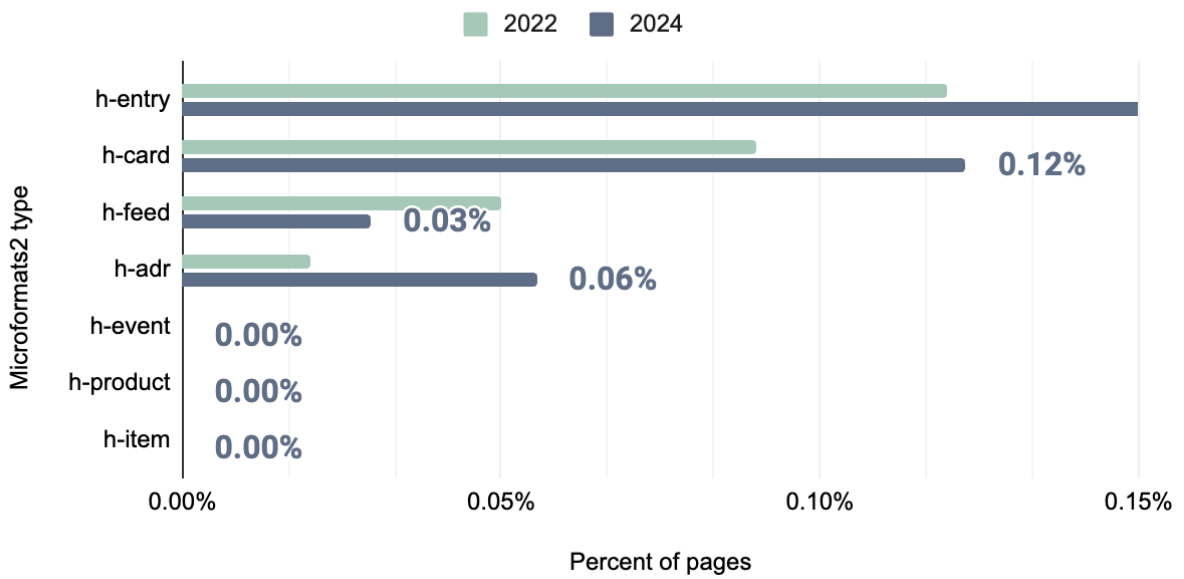
Microformats2 usage by year

Web Almanac 2024: Structured Data (mobile)



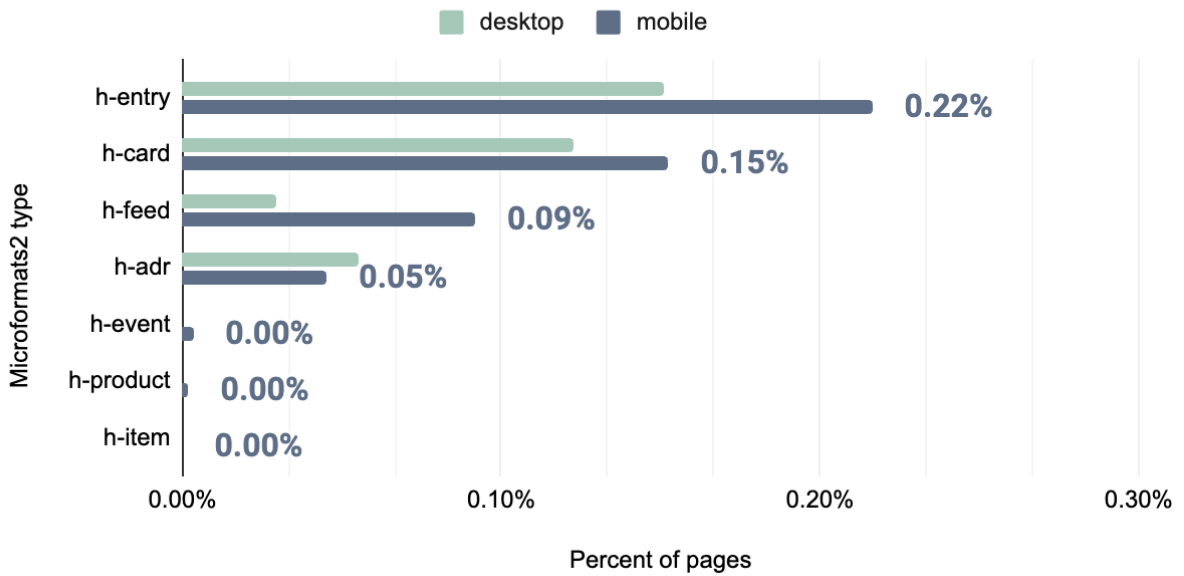
Microformats2 usage by year

Web Almanac 2024: Structured Data (desktop)



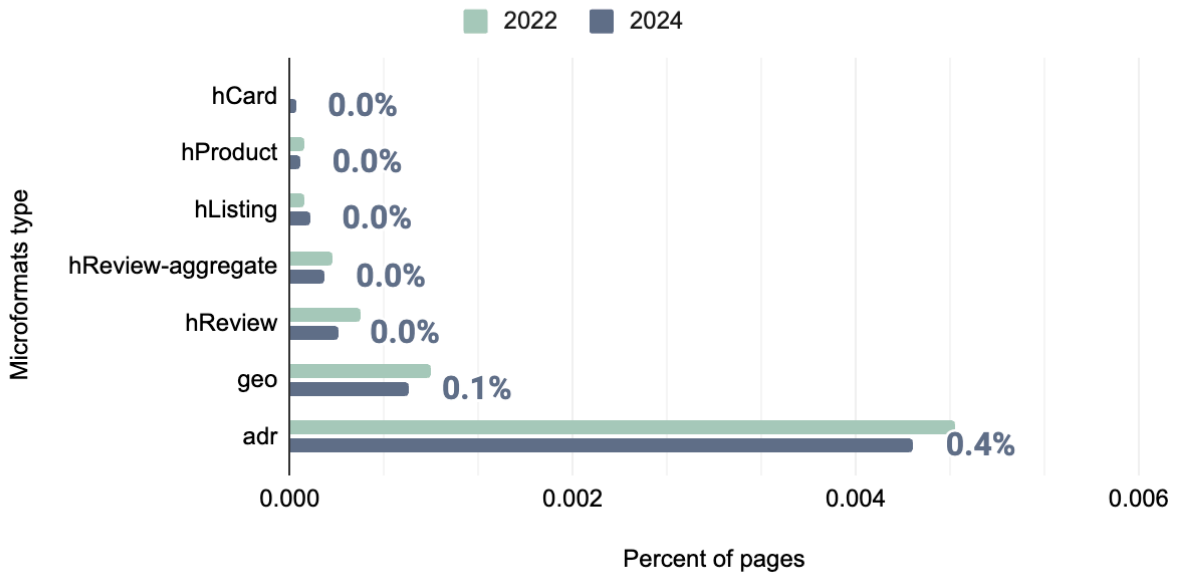
Microformats2 usage by device

Web Almanac 2024: Structured Data



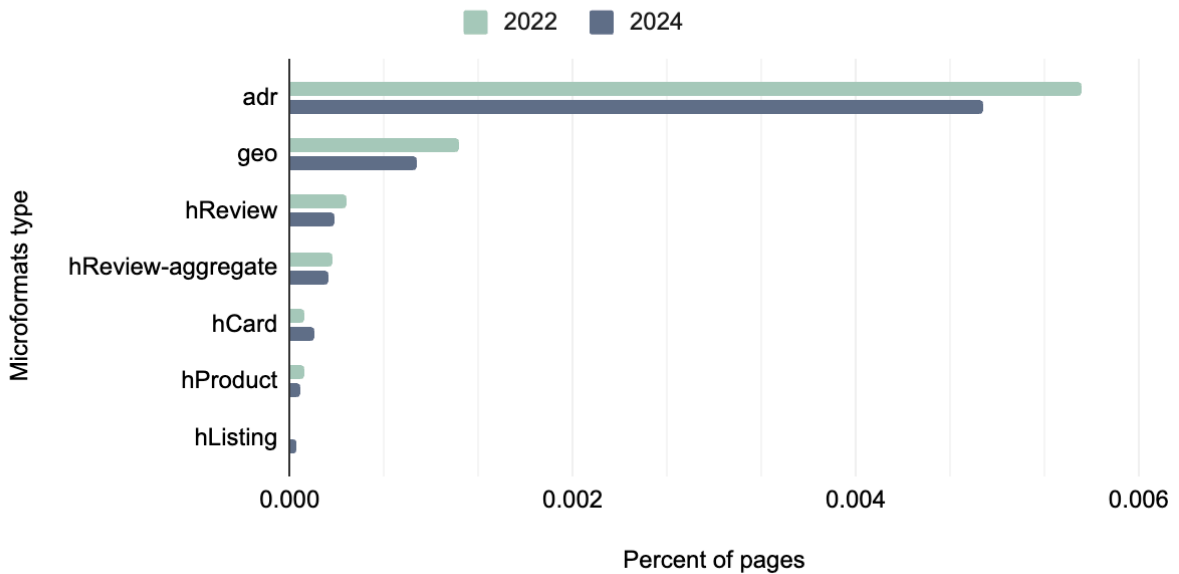
Microformats usage by year

Web Almanac 2024: Structured Data (mobile)



Microformats usage by year

Web Almanac 2024: Structured Data (desktop)

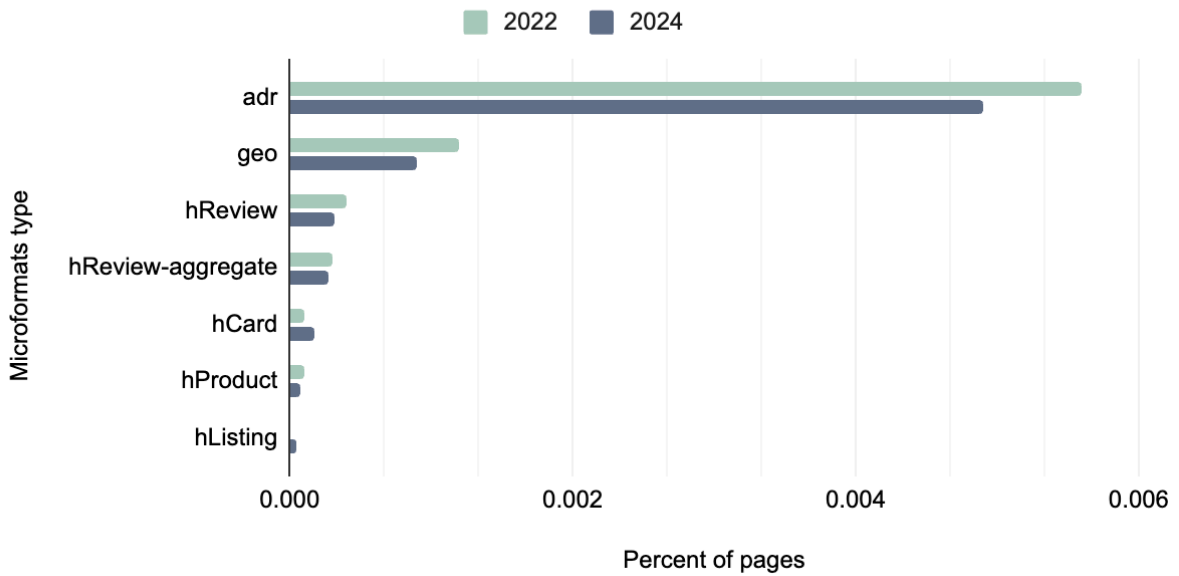


Cross-device implementation remains relatively consistent, though with some variation between mobile and desktop. The data shows a **general decline in traditional Microformats usage** from 2022 to 2024, particularly in **review-related properties** like [hReview](#) and [hReview-aggregate](#). This decline reflects the **industry's shift toward more modern structured data formats** like JSON-LD and RDFa, which offer **broader functionality** and better **integration with current web standards**.

Despite this decline, **Microformats** and **Microformats2** remain useful in specific contexts where lightweight, human-readable semantic data is needed. However, their overall presence continues to be eclipsed by more versatile formats like JSON-LD, which dominate the structured data landscape.

Microformats usage by year

Web Almanac 2024: Structured Data (desktop)

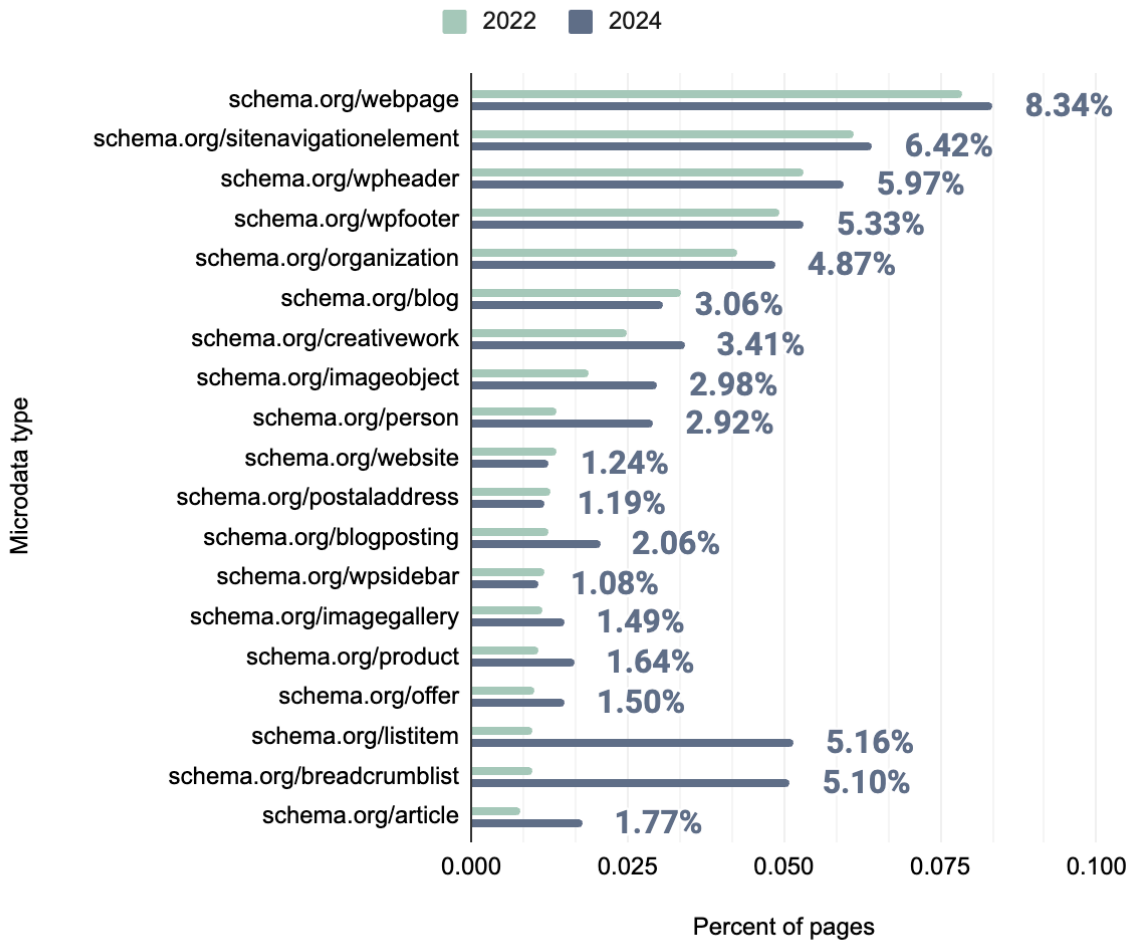


Microdata

Microdata continues to be widely used for **structural elements** and **navigation data**, particularly within legacy platforms and sites where simpler, static page structures are required. The most frequently implemented types include schema.org/webpage (appearing on **8.34% of mobile pages**) and schema.org/sitenavigationelement (used on **6.42% of mobile pages**), indicating the format's enduring relevance for **webpage structure** and **site navigation**.

Microdata usage by year

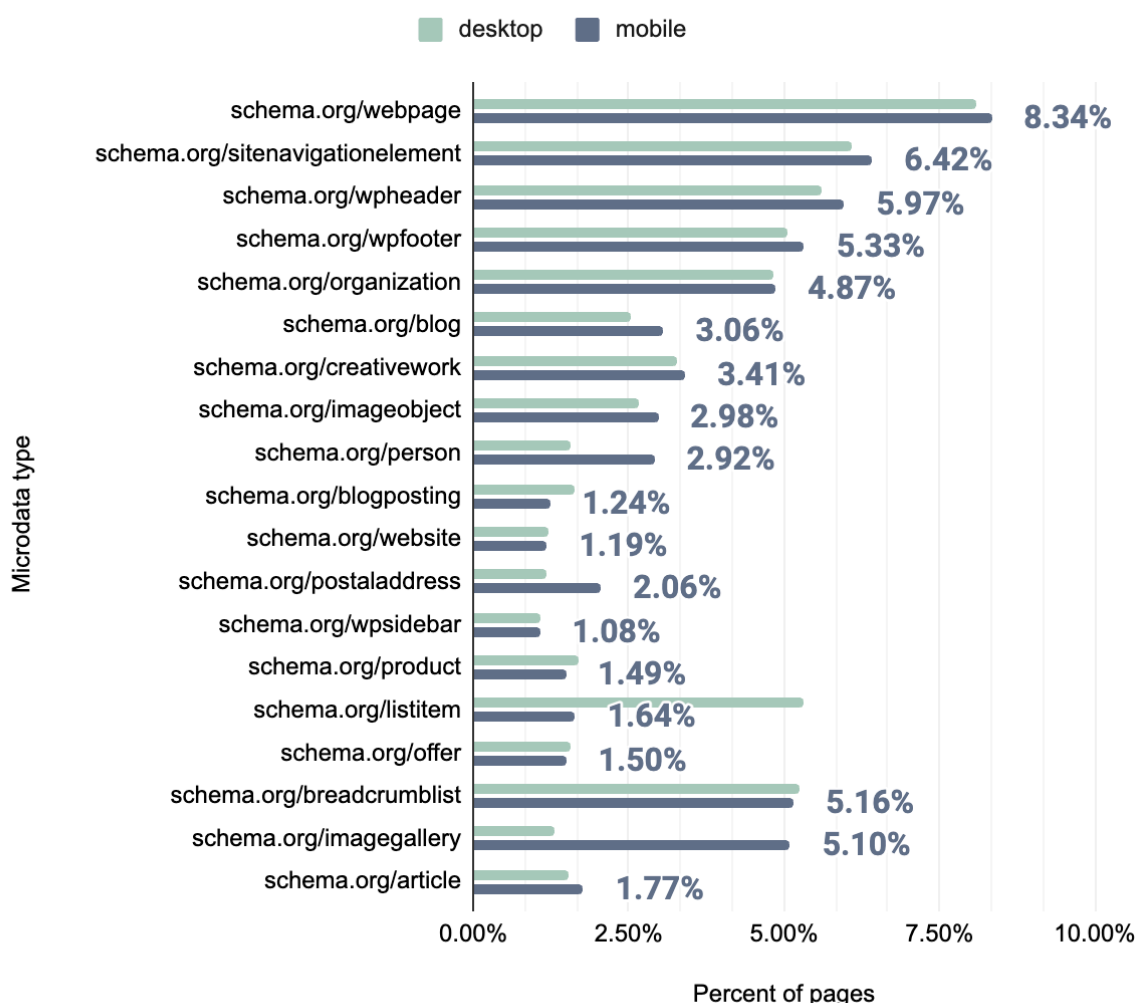
Web Almanac 2024: Structured Data (mobile)



Navigation-related types like [listitem](#) and [breadcrumblist](#) have also seen steady growth, reflecting the need for more organized and structured navigation data, particularly on mobile devices. However, **content-specific types** such as [schema.org/article](#) and [schema.org/product](#) remain less common, with adoption rates of **1.77%** and **1.50%** respectively, as developers increasingly turn to JSON-LD for more flexible and scalable implementations.

Microdata usage by device

Web Almanac 2024: Structured Data



While **Microdata** remains a significant format for **fundamental webpage structure**, its use in **dynamic content** and **e-commerce** applications has been gradually overtaken by more modern formats like **JSON-LD**, which offer broader support for content enrichment and structured data scaling across large websites.

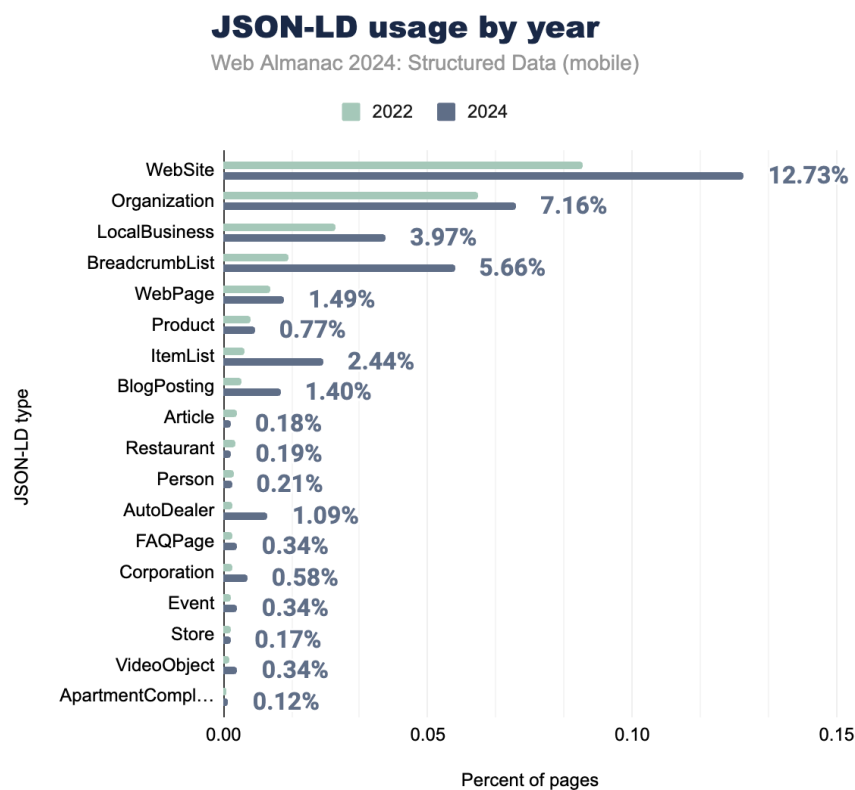
JSON-LD

JSON-LD types continue to show diverse implementation patterns across websites, with the **WebSite** schema leading adoption at **12.73%** of mobile pages, significantly higher than other types. **Organization** and **LocalBusiness** types maintain a strong presence at **7.16%** and **3.97%**, respectively, reflecting their importance in establishing **entity identity**.

BreadcrumbList implementation has seen notable growth, appearing on **5.66%** of pages, suggesting an increased focus on **structured navigation data**. The **WebPage** schema shows steady adoption at **1.49%**, while the **Product** schema appears on **0.77%** of pages. Content-specific types like **BlogPosting** (**1.40%**) and **Article** (**0.18%**) maintain consistent presence, though at lower levels.

Specialized business types such as **Restaurant** (**0.19%**), **AutoDealer** (**1.09%**), and **Store** (**0.17%**) demonstrate the growing adoption of **industry-specific markup**, corresponding to **Google's increased support** for these schemas. Supporting content types including **VideoObject**, **FAQPage**, and **Event** each appear on approximately **0.34%** of pages, indicating steady but modest implementation of **specialized content markup**.

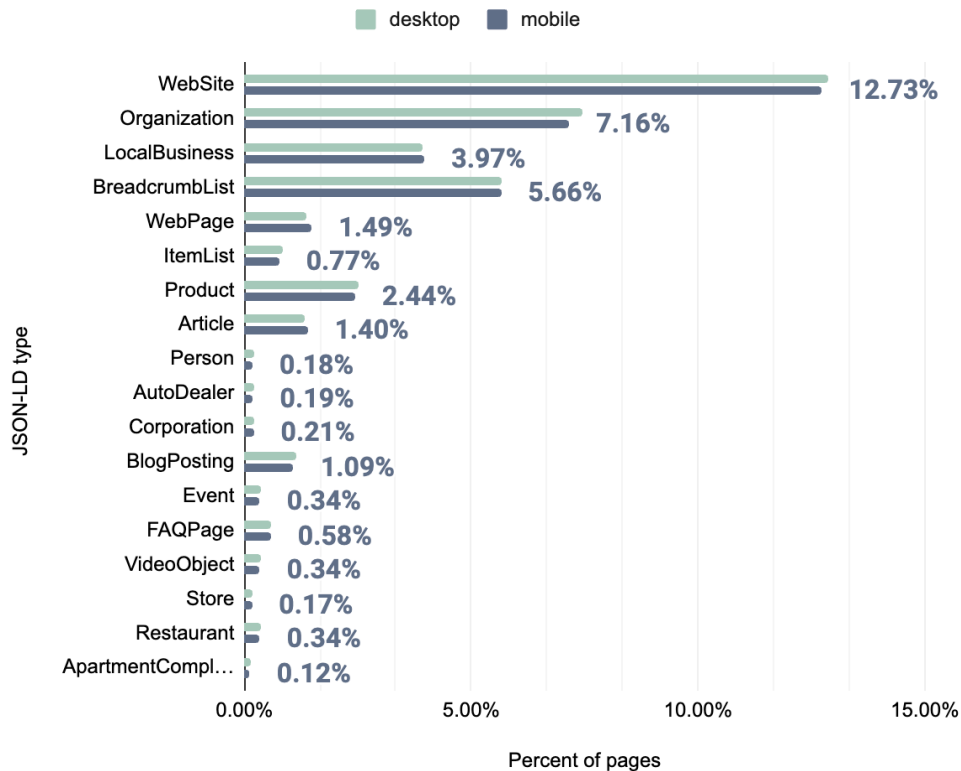
ItemList schema shows healthy adoption at **2.44%**, suggesting increased use of **structured listing data**. The overall distribution of **JSON-LD types** reflects a maturing ecosystem where **fundamental entity types** dominate, while specialized schemas serve specific **business and content needs**.



The consistency in implementation across devices indicates a mature approach to structured data deployment, where developers are ensuring uniform markup regardless of the target platform. This alignment between mobile and desktop implementations suggests that organizations are following best practices for responsive design while maintaining consistent structured data across all user experiences.

JSON-LD usage by device

Web Almanac 2024: Structured Data



JSON-LD Relationships

When evaluating JSON-LD relationships in structured data implementations, several key patterns emerge in how entities are connected in a graph. These relationship patterns reflect how structured data is used to **create comprehensive, interconnected entity descriptions** that help search engines better understand content context and relationships. The most successful implementations leverage these connections to provide rich, detailed information while maintaining logical content relationships.

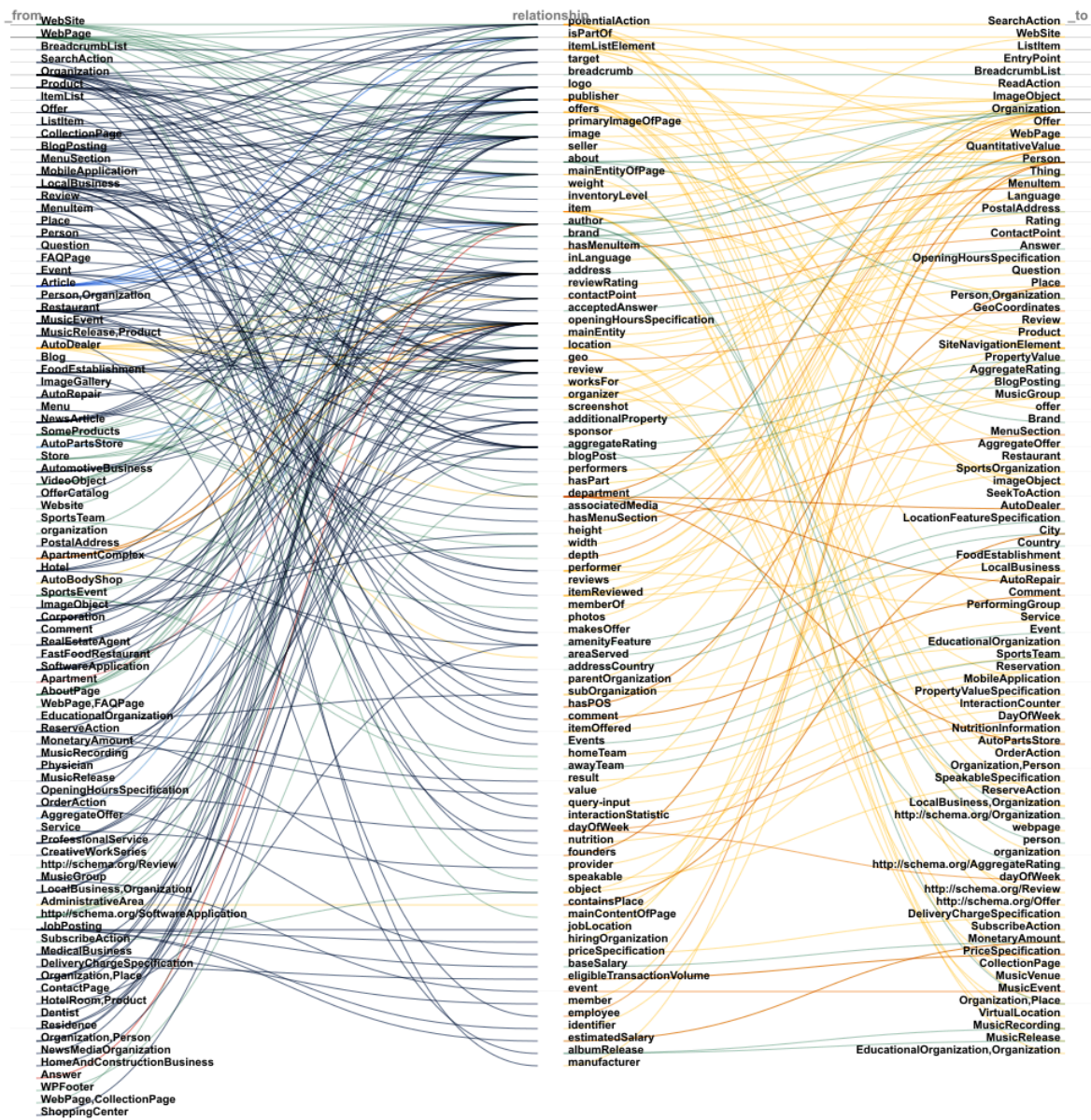
Let's review the most critical patterns from the JSON-LD relationship analysis:

- Local Business Ecosystem.** The most sophisticated structured data implementations are occurring in the local business sector, where we see rich interconnections between LocalBusiness, OpeningHoursSpecification, PostalAddress, and GeoCoordinates. This suggests businesses are creating comprehensive digital identities that go beyond basic location information to include detailed operational data. This aligns with Google's increasing focus on local search and the growing importance of location-based services.
- Content Organization.** Maturity There's a clear pattern of publishers implementing more sophisticated content structures. The relationships between Article, BlogPosting, and WebPage entities consistently link to ImageObject, author

attributes, and publishing details. This isn't just about marking up individual pieces of content – it's about creating proper content graphs that establish clear relationships between content, creators, and organizational entities.

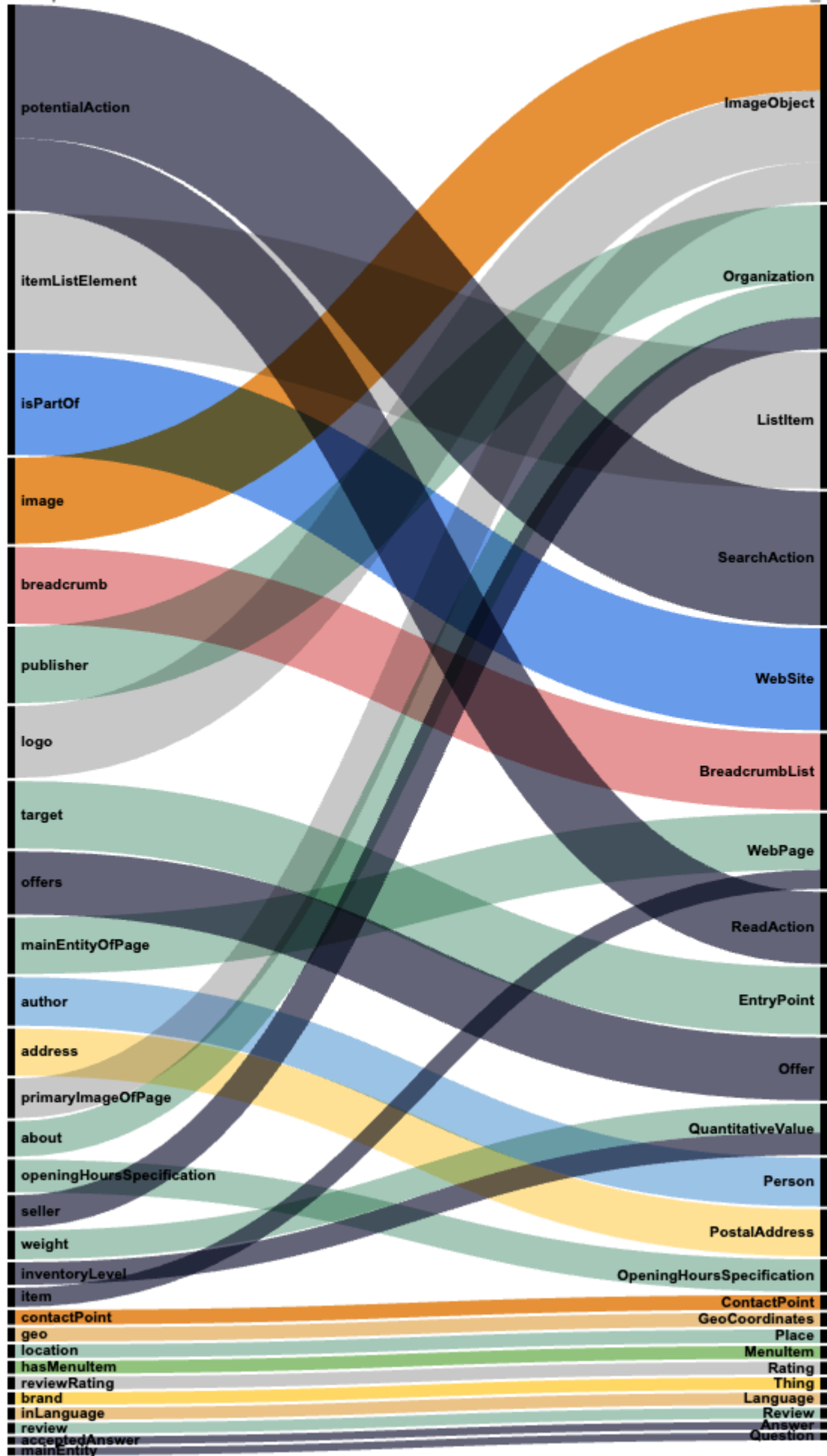
3. **E-commerce Integration.** The product-related relationships show an interesting evolution. Beyond basic product markup, we're seeing more connections to ReviewRating, AggregateOffer, and PriceSpecification entities. This suggests e-commerce sites are building more comprehensive product knowledge graphs that can support advanced features like price tracking and inventory status.

Most notably, these patterns indicate that **structured data implementation is moving beyond simple SEO markup toward creating true knowledge graphs** that can support AI-powered search experiences and rich data integrations.



relationship

_to



As seen also in the previous chart the most frequent JSON-LD property relationships reveal several critical implementation patterns across websites. PotentialAction emerges as a dominant property, showing strong connections to SearchAction and WebSite, indicating **widespread implementation of site search functionality** (we expect this to decrease as Google is removing support for this feature snippet). **Image-related properties** form another major cluster, with ImageObject frequently connected to Organization and WebPage entities, demonstrating the importance of visual content attribution. The publisher and logo properties frequently link to Organization entities, establishing clear brand identity.

Navigation structures show clear patterns through BreadcrumbList and itemListElement properties, typically connecting to WebPage entities. Content relationships are evidenced by mainEntityOfPage connections, while business-specific information flows through address, openingHoursSpecification, and geo properties.

Particularly noteworthy is the consistent implementation of **contact and location information**, with PostalAddress, ContactPoint, and GeoCoordinates forming a well-defined cluster. This suggests businesses are prioritizing local presence markup. The presence of review-related properties (reviewRating, rating) connected to various entities indicates strong focus on reputation management through structured data.

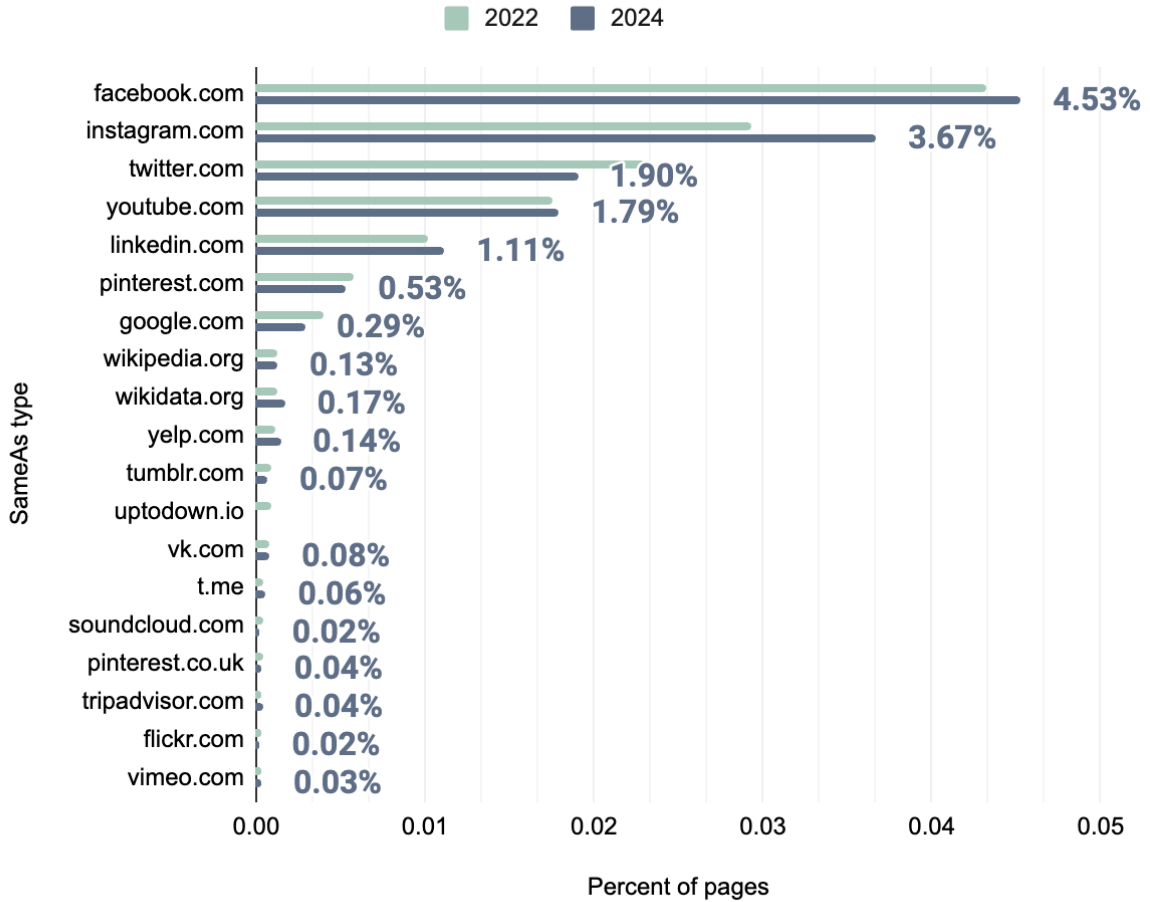
sameAs

The sameAs property plays a crucial role in **entity disambiguation and knowledge graph development**, extending far beyond simple social media profile linking. While our data shows strong implementation for major platforms (Facebook at 4.53%, Instagram at 3.67%), the true strategic value lies in how sameAs helps search engines understand and validate entity relationships.

When properly implemented, sameAs serves as a powerful tool for entity disambiguation, particularly for organizations and persons. By linking to authoritative sources like Wikidata (0.17%) and Wikipedia (0.13%), brands can establish unambiguous entity identification. This creates what we might call a "*entity fingerprint*" that helps search engines confidently associate various online presences with the correct entity.

SameAs usage by year

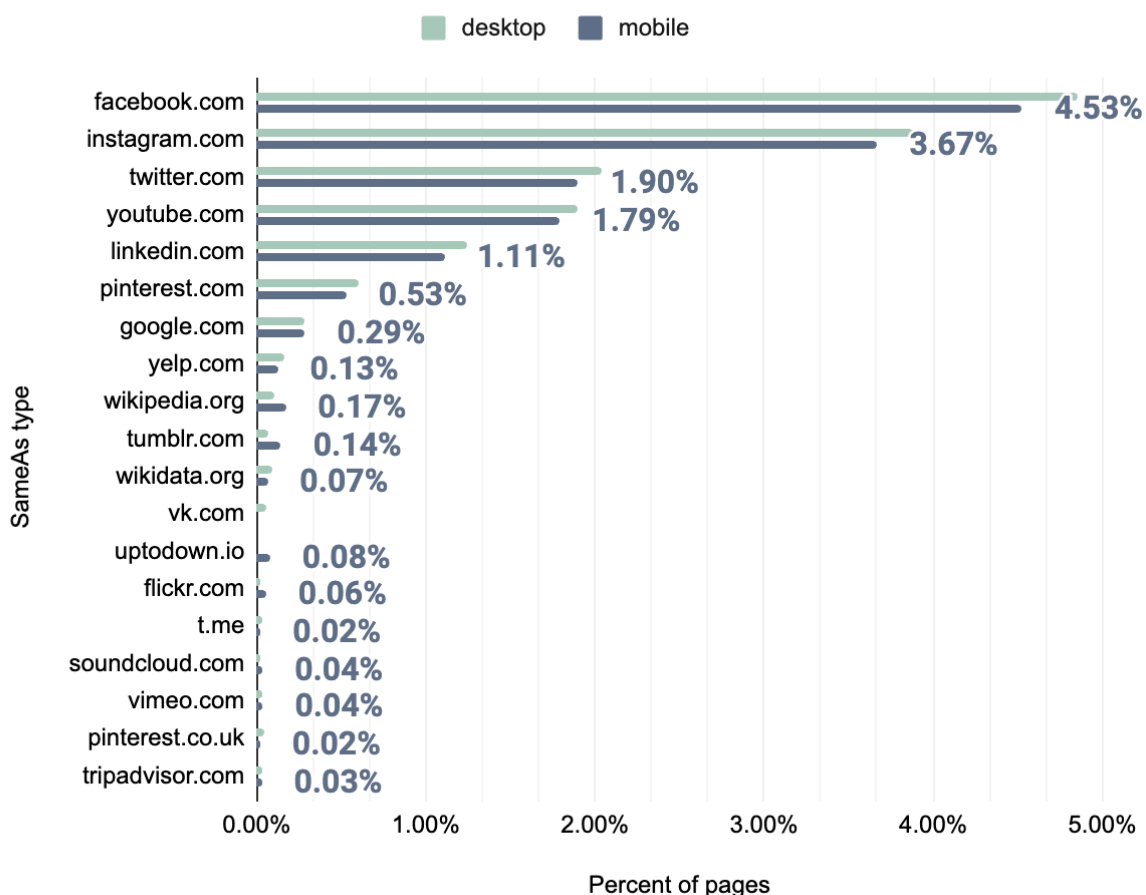
Web Almanac 2024: Structured Data (mobile)



For personal entities (executives, authors, experts), sameAs similarly helps establish authority and credibility by connecting professional profiles (LinkedIn at 1.11%) with other authentic entity markers. This becomes particularly valuable for E-E-A-T signals and knowledge panel generation.

SameAs usage by device

Web Almanac 2024: Structured Data



This parity between mobile and desktop implementations represents a significant advancement in structured data deployment. It suggests that organizations are increasingly using consistent templating systems or automated solutions to manage their structured data, rather than maintaining separate implementations for different devices.

JSON-LD Context

Schema.org remains the dominant force in JSON-LD context implementation with over 20 million instances, far exceeding all other contexts. This dominance (20,960,693 implementations versus the next highest at 11,973) reflects its position as the industry standard for structured data markup.

Among secondary implementations, contao.org leads with 11,973 instances, primarily within its CMS ecosystem, followed by googleapis.com (3,743) and baidu.com (1,409). Educational institutions show consistent adoption patterns around 25-50 implementations each, while regional variations appear through implementations like schema.org.cn and schema.gov.sg, indicating global adoption of structured data standards.

The vast gap between Schema.org and other contexts underscores its critical role in structured data standardization and reflects strong alignment with search engine requirements.

Emerging Trends and Future Outlook

The structured data landscape is rapidly evolving, marked by Google's introduction of specialized schemas for vehicles, courses, and 3D product models, alongside increased support for Digital Product Passports through GS1 Digital Link. The growing adoption of JSON-LD (now at 41% of pages) and sophisticated entity relationships through sameAs properties indicates a maturing ecosystem focused on comprehensive knowledge graph development.

The data shows a clear shift toward more specialized implementation patterns, particularly in e-commerce and local business contexts. Entity disambiguation has become increasingly critical, with organizations leveraging structured data to establish clear digital identities across platforms and knowledge bases.

Looking Ahead: The Future of Structured Data

As we analyze current trends, we also cast our gaze forward to emerging developments:

- **AI and Structured Data Symbiosis**

The growing interdependence between **AI systems** and **structured data** is becoming crucial for delivering **grounded, hallucination-free content generation** and enhancing **conversational search experiences**. As AI relies increasingly on structured data for accurate and context-rich information, this symbiosis is redefining how **AI-powered tools** interact with content across the web.

- **Data Commons and Knowledge Graph Integration**

The expansion of open data initiatives, such as **Google's Data Commons**, which leverages **Schema.org** for structuring and linking public datasets, is further fueling the evolution of **knowledge graph-based systems**. These initiatives provide a rich, unified foundation for **AI-driven data enrichment** and **exploration**, creating new possibilities for scalable and reliable data integration across platforms.

- **SEOntology and Specialized Vocabularies**

In parallel, the development of [SEOntology](#) and other specialized vocabularies is addressing the need for **SEO-specific structured data** that can improve **content discoverability** and **search engine optimization**. By creating vocabularies tailored

to the unique requirements of SEO, we can further enhance the alignment between structured data and AI, driving more targeted and efficient search experiences.

- **Regulatory Impacts**

Finally, regulations such as the **EU's Digital Product Passport** are poised to reshape future structured data standards. These initiatives will likely influence how structured data is applied, especially in domains like **e-commerce** and **product traceability**, encouraging more structured and transparent data practices.

By examining these aspects, we aim to provide a comprehensive overview of the state of structured data in 2024, its recent evolution, and its future trajectory. Whether you're a seasoned SEO professional, a web developer, an eCommerce strategist, or simply interested in the evolution of the web, this chapter offers valuable insights into how structured data is reshaping our digital world and paving the way for a more connected, transparent, and intelligent online experience.

Conclusion

The analysis of structured data in 2024 highlights a clear shift from its SEO roots toward a broader, more strategic role in AI and semantic metadata. The dominance of RDFa and Open Graph on over 60% of pages, combined with JSON-LD's growth (now on 41% of pages, particularly in eCommerce), points to a maturing technology. But the true impact lies in how structured data is transforming AI discovery and enhancing machine understanding.

This year, we've seen significant changes in how search engines handle structured data. While Google has deprecated certain rich results, such as FAQs, HowTos, and SiteLinks, they've simultaneously introduced new types for vehicles, courses, 3D product models, product groups, loyalty cards, and certifications, expanding the scope of structured data. Even more importantly, structured data is now essential for AI systems, supporting tasks from fact-checking to improved search capabilities and training large language models (LLMs).

The advent of Digital Product Passports and increased adoption of GS1 standards underlines the growing importance of structured data in commerce and regulatory compliance. As AI-driven search becomes the norm, businesses are realizing that structured data is no longer just about search visibility—it's key to ensuring content is machine-readable and future-proof.

For businesses developing their structured data strategy, the way forward is clear: implement it comprehensively, maintain it rigorously, and adapt continuously. New projects should focus on JSON-LD, while legacy formats should be preserved where appropriate. Systems must be built to scale and evolve alongside emerging technologies and standards.

In conclusion, the future of the web is structured, semantic, and increasingly intelligent. Organizations that invest on structured data today won't just improve their search visibility – they are building the foundation for success in AI Discovery.