Engineer

Strukturierte Daten

Fakten über eine Seite, für Maschinen geschrieben

Was es ist

Die Kontaktseite eines Unternehmens beschreibt in einem Satz, wo das Büro liegt; strukturierte Daten sagen es noch einmal als `PostalAddress` mit Straße, Postleitzahl und Land, damit ein Programm beim Lesen der Seite nicht erraten muss, welches Wort den Ort bezeichnet. Dabei wird wiederholt, was eine Seite in Prosa aussagt: als typisierte Aussagen in einem Vokabular, auf das sich Anbieter und Nutzer der Daten verständigen. So können Suchmaschinen und andere Programme die Fakten lesen, ohne das Layout auszulegen.

Das Vokabular ist meist schema.org. Googles Dokumentation nennt es als Grundlage für die meisten strukturierten Daten, die seine Suche verwendet. Google, Microsoft, Yahoo und Yandex haben es gegründet; seit April 2015 ist eine W3C Community Group sein wichtigstes Forum. Es definiert Typen wie `Organization`, `Person`, `Service`, `Review` oder `DefinedTerm` und die Eigenschaften, die jeder davon haben kann, etwa `name`, `address`, `author` oder `price`.

Die Aussagen lassen sich in drei Syntaxformen schreiben. Microdata und RDFa ergänzen Attribute direkt an den sichtbaren HTML-Elementen. JSON-LD hält die Aussagen getrennt: als JSON-Block in einem `script`-Element mit dem Typ `application/ld+json`. Seine Version 1.1 wurde am 16. Juli 2020 als W3C Recommendation veröffentlicht. Google akzeptiert alle drei gleichermaßen und empfiehlt JSON-LD, weil es sich im großen Maßstab am einfachsten pflegen lässt.

```
<script type="application/ld+json">
    {
        "@context": "https://schema.org",
        "@type": "Organization",
        "@id": "https://example.com/#organization",
```

```
        "name": "Example Ltd",
        "url": "https://example.com/"
    }
</script>
```

Die wichtigste Zeile in diesem Block ist `@id`. JSON-LD identifiziert damit einen Knoten. Ein Knoten ohne eine solche Kennung ist ein unbenannter Knoten, ein Blank Node, ohne eine vom Dokument unabhängige Kennung. Kennungen unbenannter Knoten können von Blöcken innerhalb desselben Dokuments gemeinsam verwendet werden. Mit einer global verwendbaren Kennung kann jeder andere Block auf jeder anderen Seite `"publisher": {"@id": "https://example.com/#organization"}` schreiben. Die Spezifikation nennt das einen Knotenverweis: einen Zeiger auf etwas, das anderswo beschrieben wird, statt einer neuen Beschreibung derselben Sache. Erst Kennungen machen aus verstreuten Blöcken einen gemeinsamen Graphen.

Eng verwandt ist der Begriff der Metadaten. Titel, Beschreibung und Open-Graph-Tags im Kopf einer Seite beschreiben die Seite als Dokument. Strukturierte Daten beschreiben dagegen die Dinge, um die es auf der Seite geht: das Unternehmen dahinter, die Person, die sie geschrieben hat, die angebotene Dienstleistung oder das erklärte Wort.

Das Konzept für eine Entscheidung nutzen

Nutzen Sie „Was sind strukturierte Daten?“ für eine konkrete Entscheidung. Wir können helfen, das Konzept mit dem System in Verbindung zu bringen, das Sie entwickeln oder einkaufen. Eine Website oder App, die die KI-Assistenten Ihrer Kunden finden, lesen und bedienen können — bei jeder Veröffentlichung geprüft statt einmal versprochen.

Unsere Leistungen entdecken Kontakt aufnehmen →

Die Annahmen hinter dem Begriff prüfen

Nutzen Sie die Erklärung und die verlinkten Arbeiten, um Fragen für Ihr Projekt zu formulieren. Wir können die Annahmen gemeinsam mit Ihnen prüfen, bevor ein Werkzeug ausgewählt wird.
Hugo-Website mehrsprachig in 3 Sprachen aufgebaut — 108 Vorlagendateien für drei Sprachen und vier Formate

Hugo-Website mehrsprachig in 3 Sprachen aufgebaut Kontakt aufnehmen →

Warum es wichtig ist

Die Entscheidung lautet, wie viel ausgesagt werden soll. Die ehrliche Antwort ist meist: weniger, als das Vokabular erlaubt, aber vollständig. Nichts in der Syntax hindert einen Anbieter daran, jede von schema.org angebotene Eigenschaft auszufüllen. Googles eigene Anleitung weist in die andere Richtung: lieber

weniger empfohlene Eigenschaften vollständig und korrekt angeben als jede mögliche mit unvollständigen oder falschen Daten. Nichts auszeichnen, was ein Leser auf der Seite nicht sehen kann, selbst wenn es stimmt.

Der Grund: Jede ausgezeichnete Tatsache ist eine zweite Fassung von etwas, das die Seite bereits sagt, und eine zweite Fassung kann abweichen. Ein Preis wird in der sichtbaren Tabelle geändert und im JSON vergessen, eine Bewertung erscheint nur für Maschinen, eine Adresse wird auf jeder Seite etwas anders geschrieben. In jedem Fall sagt die Seite Menschen etwas anderes als Programmen. Die Kosten strukturierter Daten liegen weniger im Schreiben als darin, ihren Inhalt korrekt zu halten. Die nützlichen Fragen sind deshalb, für welche Aussagen die Seite einstehen kann und ob diese aus derselben Quelle stammen wie der sichtbare Text.

Der Gewinn ist Lesbarkeit. Eine Suchmaschine oder ein Agent, der hundert Erwähnungen zu einem Unternehmen zusammenführen kann, weiß mehr als einer, der hundert unverbundene Namen liest. Seiten, die den Richtlinien folgen, können für Rich Results infrage kommen, also Suchergebnisse mit zusätzlichen Angaben. Die Eignung garantiert allerdings kein solches Ergebnis. Was sich ändert, ist, wie gut eine Maschine versteht, worum es auf der Seite geht.

Wie es in der Praxis aussieht

Die Quellenseite dieser Website verpflichtet sie dazu, nur das auszuzeichnen, was auf der Seite wahr und sichtbar ist. Die Geschichte ihrer strukturierten Daten zeigt, warum dieser Satz festgehalten werden musste.

Vor der Überarbeitung im Juli 2026 beschrieben 963 Blöcke das Unternehmen 1.671-mal, ohne irgendwo eine stabile Kennung zu verwenden. Ein Nutzer dieser Daten hatte deshalb keine Grundlage dafür, die Organisation auf einer Seite als dieselbe wie auf einer anderen zu behandeln. Der Graph beruht nun auf drei aus dem Origin der Website gebildeten Kennungen: für die Organisation, die Website und die Gründerin. Alle anderen Blöcke verweisen mit `@id` auf sie, statt sie erneut zu beschreiben. Die Kennungen sind in jeder Sprache gleich. Ein Kommentar in der Vorlage erklärt, dass eine eigene dänische Kennung eine zweite Organisation mit derselben Unternehmensregistrierungsnummer behaupten würde. In dieser damaligen Phase waren sechzehn Typen dokumentiert, und Listen nannten ihre Einträge allein per URL. Die damals 107 Einträge auf der Portfolioübersicht vollständig auszuschreiben hätte die komprimierte Seite um 52 Prozent vergrößert.

Mehrere Fakten werden aus derselben Quelle gelesen wie die sichtbare Seite. Dienstleistungspreise stammen aus derselben Datei wie die Preise auf dem Bildschirm; einer Maschine und einem Leser wird also derselbe Preis genannt. Der Zeitraum, den eine Arbeit umfasst, kommt aus demselben Partial, das den Satz darüber schreibt. Brotkrumennavigationen folgen dem sichtbaren Pfad, und Listenpositionen folgen der Reihenfolge auf dem Bildschirm, niemals einer nur für Maschinen veröffentlichten Rangfolge. Der eigene Knoten der Seite verwendet dieselbe Adresse wie ihr kanonischer Link. Markup und Dokumentkopf treffen damit dieselbe Aussage.

Die bewussten Auslassungen sind ebenso überlegt wie die Aussagen. Die eigenen Arbeitsbeschreibungen des Unternehmens werden nicht als Rezensionen ausgezeichnet, weil den eigenen Worten damit ein Verfasser zugeschrieben würde, den es nicht gibt. Die Vorlage für Empfehlungen verwendet Review-Knoten nur für Einträge, die mit erteilter Zustimmung veröffentlicht werden, und fügt keine Bewertungsnote hinzu. Jede Glossarseite, auch diese, enthält einen `DefinedTerm`, dessen Beschreibung dieselbe einzeilige Erläuterung ist wie in der Liste. Die Definition, die eine Maschine liest, ist damit dieselbe, die ein Leser sieht.

Eine Prüfung der gebauten Website stellt sicher, dass sich jeder Block parsen lässt, jede Kennung aufgelöst werden kann und jede im Graphen genannte URL gebaut wurde. Ein Graph, der auf eine nicht vorhandene Seite verweist, lässt das Qualitätsgate somit scheitern. Dieselbe Quellenseite nennt den Schema Markup Validator zum Lesen der Daten und Googles Rich Results Test, um zu sehen, was eine Suchmaschine daraus macht.

Wo es schiefgeht

Rich Results bleiben aus, obwohl alle Validatoren zufrieden sind. Das Markup ist gültiges JSON-LD und verwendet echte schema.org-Typen; trotzdem erscheint die Seite weiter als gewöhnliches Suchergebnis. Die Daten beschreiben Dinge, die ein Leser nicht sehen kann: eine nie abgegebene Bewertung, auf der Seite verborgene Fragen oder einen anderen Preis als den sichtbaren. Google warnt, dass Markup, das gegen seine Richtlinien verstößt, eine Seite ungeeignet machen kann, selbst bei fehlerfreier Syntax. Nur das auszeichnen, was die Seite zeigt, und beide Darstellungen aus einer Quelle erzeugen, damit sie nicht auseinanderlaufen.

Hunderte Blöcke beschreiben ein Unternehmen, und nichts verbindet sie. Jede Seite wiederholt Namen, Logo und Adresse der Organisation, mitunter leicht abweichend. Eine Korrektur bedeutet dann, alle Vorkommen zu finden. Die Knoten sind anonym, oder jede Seite und jede Sprache bildet für dieselbe Entität eine eigene Kennung. Jedem realen Objekt eine aus einer stabilen URL mit Fragment gebildete `@id` geben, es einmal beschreiben und überall sonst nur referenzieren.

Eine Änderung an der Vorlage beschädigt das Markup aller Seiten zugleich. Ein verirrtes Anführungszeichen, ein abschließendes Komma oder ein leerer Wert macht jeden Block ungültig; die Nutzer der Daten können dann nichts davon verwenden. Googles Anleitung warnt, dass Vorlagen das Markup nach der Bereitstellung beschädigen können. Das JSON wurde als Text zusammengesetzt, mit zwischen geschweifte Klammern eingefügten Werten. Ein fehlendes Feld hinterlässt deshalb eine leere Zeichenfolge oder eine beschädigte Struktur. Die Aussage als Datenstruktur aufbauen und serialisieren, wie es die Vorlagen dieser Website mit Hugos `jsonify` tun. Eigenschaften ohne Wert weglassen, statt sie leer auszugeben, und jeden Block in der gebauten Ausgabe parsen.

Der Graph nennt fehlende Seiten. URLs im Markup führen zu einer 404-Seite, einem Staging-Host oder einer gespiegelten Kopie. Die Adressen wurden von Hand geschrieben oder nach einer anderen Regel gebildet als Links und kanonische Adresse. Jede URL im Markup aus derselben Funktion ableiten, die die kanonische Adresse der Seite erzeugt, und prüfen, dass sie tatsächlich gebaut wurde.

Weiterführende Lektüre

- schema.org — das gemeinsame Vokabular aus Typen und Eigenschaften, mit einer Seite für jeden Typ.
- JSON-LD 1.1 — die W3C Recommendation für die Syntax einschließlich Knotenkennungen und Knotenverweisen.
- schema.org: DefinedTerm — der Typ, mit dem jede Glossarseite das erklärte Wort beschreibt.
- Google Search Central: Einführung in strukturierte Daten — die drei Syntaxformen, die Sichtbarkeitsregel und die Prüfung von Rich Results.
- Schema Markup Validator — liest die strukturierten Daten einer Seite anhand des schema.org-Vokabulars.

Glossar

Für Ihre tatsächlichen Rahmenbedingungen auswählen

Beschreiben Sie den Arbeitsablauf, die vorhandenen Schnittstellen und die Nutzer, die die Veränderung benötigen. Wir können eine prüfbare Lieferung definieren und die Übergabe so gestalten, dass Ihr Team die gelieferten Bestandteile warten kann.

Preise ansehen Kontakt aufnehmen →

Die Begriffe praktisch nutzbar machen

Fragen Sie, wie „Was sind strukturierte Daten?“ auf Ihre Organisation zutrifft. Ein konkretes Beispiel aus Ihrem bestehenden System reicht aus, um das Gespräch zu beginnen.

Unsere Leistungen entdecken

welcome@engineer.company +45 50 10 61 84 Kopenhagen, Dänemark

Nehmen Sie diese Seite mit

Kontaktkarte

Mit Sorgfalt und Liebe gestaltet von D.

CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N

Copyright © 2025 – heute · Engineer ApS · Alle Rechte vorbehalten.

Gedruckt von https://engineer.company/de/glossary/structured-data/