Webinar

How to Prepare Your Content for Generative AI Search
Schema App is a scalable Knowledge Graph Solution for Enterprise SEO teams

We are a semantic technology company that help organizations build their marketing knowledge graph with our Schema Markup solution to accelerate business results and stand out in search.
Complete Small Business Tasks with Prompt Management & 1-Click Prompts for ChatGPT

AiPRM is a prompt management tool and community-driven prompt library. Use it to complete marketing, sales, operation, productivity, and customer support tasks in minutes that used to take hours with ready-to-use prompts for ChatGPT, Bard, Midjourney, and DALL-E 2.

Installs for Free

Already have the extension? Subscribe to a Premium Plan

Built for Small Businesses. Trusted by Big Businesses.
(If we showed the Small Business logos, you wouldn’t recognize them.)
How are search engines changing?

- Lexical to semantic search
- Context changing in SGE
- Knowledge Graphs
Two primary search models:
Lexical search vs. Semantic search
Lexical search is about counting the incidence rate of words.

Semantic search about deriving meaning from words and their relationships.
Words are converted to multi-dimensional coordinates in the vector space

"How old are you?"  [0.3, 0.2, …]  
"What is your age?"  [0.2, 0.1, …]  
"My phone is good."  [0.9, 0.6, …]  
...
Vector Space Model

- Documents and queries are converted to vector representations and plotted in multi-dimensional space.

- The query and document vectors are then compared based on cosine similarity and the ones that are closest to the query are the most relevant.

- The main takeaway here is that relevance is a quantitative value. This is perhaps the most important concept to understand about how search works.
This allows for mathematical operations

- Comparisons of content and keywords become linear algebraic operations.
Relevance is a Function of Cosine Similarity

- When we talk about relevance, the question of similar is determined by how similar the vectors are between documents and queries. This is a quantitative measure, not the qualitative idea of how we typically think of relevance.

- Angle \( \theta \) close to 0
- Cos(\( \theta \)) close to 1
- Similar vectors

- Angle \( \theta \) close to 90
- Cos(\( \theta \)) close to 0
- Orthogonal vectors

- Angle \( \theta \) close to 180
- Cos(\( \theta \)) close to -1
- Opposite vectors
Vector Density Has Improved Through Transformers (the T in GPT)

- Search engines have improved their ability to understand content by increasing the amount of information captured in their vectors.
Entities Have Vector Representations

- Information from knowledge graphs are converted to vector representations and used in a variety of tasks that inform search engine use cases.
Content & SEO teams must now prioritize crafting comprehensive, holistic content over narrow keyword-focused pieces.
Context changing in SGE & the New Bing

In psychology, a schema is a cognitive framework or concept that helps organize and interpret information. It describes patterns of thinking and behavior that people use to interpret the world. Schemas are developed and based on life experiences and provide a guide to one's cognitive processes and behavior. Schemas are also used to describe the specific actions or behaviors that babies, toddlers, and young children engage in as they explore the world around them and try to find out how things work.

Hello, this is Bing! I'm the new AI-powered chat mode of Microsoft Bing that can help you quickly get summarized answers and creative inspiration.

- Get a question? Ask me anything - short, long, or anything in between.
- Dive deeper. Simply ask follow-up questions to refine your results.
- Looking for things to do? I can help you plan trips or find new things to do where you live.
- Feeling stuck? Ask me to help you plan any project from meal planning to gift ideas.
- Need Inspi...?
Queries are Longer and the Featured Snippet is Bigger

1. The query is more natural language and no longer Orwellian Newspeak. It can be much longer than the 32 words that is has been historically in order.

2. The Featured Snippet has become the “AI snapshot” which takes 3 results and builds a summary.

3. Users can also ask follow up questions in conversational mode.
This is Called “Retrieval Augmented Generation”

- Neeva (RIP), Bing, and now Google’s Search Generative Experience all use pull documents based on search queries and feed them to a language model to generate a response.
Google’s Version of this is called Retrieval-Augmented Language Model Pre-Training (REALM)
There’s a Lot of Synergy Between KGs and LLMs

There are three models gaining popularity:

1. **KG-enhanced LLMs** - Language Model uses KG during pre-training and inference
2. **LLM-augmented KGs** - LLMs do reasoning and completion on KG data
3. **Synergized LLMs + KGs** - Multilayer system using both at the same time

Source: Unifying Large Language Models and Knowledge Graphs: A Roadmap
Organizations are doing RAG with Knowledge Graphs

- Anyone can feed their data into an LLM as a fine-tuning measure to improve the output.
- People are currently using their knowledge graphs to support this.
Rank Tracking Will Be More Complex

- As an industry, we'll need to decide what is considered the #1 result.
- Based on this screenshot positions 1-3 are now the citations for the AI snapshot and #4 is below it.
- However, the AI snapshot loads on the client side, so rank tracking tools will need to change their approach.
Context Windows Will Yield More Personalized Results

- SGE maintains the context window of the previous search in the journey as the user goes through predefined follow questions.
- This will need to drive the composition of pages to ensure they remain in the consideration set for subsequent results.
But, the Follow Up Questions Are Different From the PAA

- There is some overlap, but the follow up questions featured in the AI Snapshot typically differ from those in the PAAs.
- So, we’ll need to find another way to get this data.
Neural Matching

- Google's ranking algorithm has evolved to neural matching from keyword matching.

- It now recognizes and connects **related entities** across entire websites, emphasizing the importance of their interrelationships.
  - Even if a page lacks specific search terms, the concept can be connected to related entities and their attributes

- Content creators and SEO experts must now prioritize crafting comprehensive, holistic content over narrow keyword-focused pieces.
Fact Verification

- Although Google has historically said they do not verification of facts.
- LLM + KG integrations make this a possibility and Google needs to combat the wealth of content being produced with LLMs. So, it’s likely they will use this functionality.

Source: Fact Checking in Knowledge Graphs by Logical Consistency

Source: FactKG: Fact Verification via Reasoning on Knowledge Graphs
What is an Entity?

An entity is a thing with specific attributes.
Your Webpage  =  Entity Home
Entity authority and context

- In the world of SGE, you are the authority of your own content
  - 3rd party should not be the entity definition, you link to it
- You define your entities and how they are connected to other entities on your website and in the world
- The entity definition should live on the page that primarily speaks about that topic
- Use Schema Markup to articulate what the page is about and what linked entities exist in context
Linked Entities within your Knowledge Graph

Linked Entities to Wikipedia or Google Knowledge Graph

Entities relate topics with context (triples)

Topics are a group of keywords

Keywords
How Schema App helps you manage your knowledge graph

- Schema App platform creates entities for your content through the application of Schema Markup where each “@id” is a URI
  - You can link entities on your site using these URIs
- Linked Entity Recognition feature automatically connects your entities to external authoritative knowledge bases like Wikidata & Google Knowledge Graph at scale
Schema Markup builds Knowledge Graphs
What is Semantics?

Semantics is a term for the study of language, how it expresses meaning, and the rules it must abide by in order to convey this meaning.
What is Semantic Technology?

Semantic technology are the tools and technologies we use to articulate the meaning and relationship between things.
Schema Markup = Language to define Entities on your website
Do search engines care about semantic Schema Markup?

In short, yes.

Google

“Also, over time richer/correct semantics will favour more connected graphs.”

– Ryan Levering, Google (Source: Mastodon)

Bing

“Write great content annotated with schema markup... use Semantic Markup to convey information about the page.”

– Fabrice Canel, Bing (Source: Pubcon Austin 2023)
What is a Knowledge Graph?

A collection of relationships between things defined using a standardized vocabulary, from which new knowledge can be gained through inferencing.
Semantic Schema Markup = Marketing Knowledge Graph

- All of your entities have URIs, and are connected using the Schema.org vocabulary.
- Connections between entities and properties of things (i.e., name, MedicalSpecialty, aggregateRating)
- Connections and Properties are expressed in RDF. RDF is a data model that expresses everything as triples.
A semantic triple, commonly used in the context of the semantic web and knowledge graphs, represents a fundamental piece of information as a relationship between two entities in the form of:

subject-predicate-object

- The subject denotes the entity being described
- the predicate specifies the type of relationship or property
- the object represents the value or another entity that the subject is related to

In knowledge graphs, these triples form interconnected networks of data, allowing machines to understand and reason about complex relationships and hierarchies in a structured manner.
Knowledge Graphs are Networks of Triples

Knowledge graphs represent information as a **web of interconnected nodes and edges formed by semantic triples**, where nodes correspond to subjects and objects, and edges represent predicates.

By linking these triples together, knowledge graphs facilitate understanding and querying of complex relationships and structures in a coherent and structured manner.

Search engines extract this implicit information through parts of speech tagging. **Structured data allows you to provide this data more explicitly.**
Why should you implement Semantic Schema Markup?

- Develop your knowledge graph by connecting the various entities to your organization.
- Help AI search engines understand and contextualize your content.
- Train AI search engines to provide more accurate results by grounding the LLM with your knowledge graph.
Schema Markup to achieve Rich Results

Schema Markup to build a Knowledge Graph
How to incorporate Schema Markup into your Content Strategy

1. Design your content to talk about specific topics / entities
2. Use Schema Markup to describe the entity and its relationship to other things on the page and on your site for greater context
3. Connect known entities on your page with Wikidata or Google Knowledge Graph
Example

Emily Schroeder, MD
Internal Medicine
⭐⭐⭐⭐⭐ 4.5 (178 Reviews)

Emily Schroeder fulfilled her internal medical residency at Washington Health for specialty training in Endocrinology. Emily Schroeder, MD is a board certified internist and endocrinologist. Dr Schroeder's areas of expertise include:

- NEW: online doctor visits
- Osteoporosis
- Obesity
- Hormonal Disorders
- Postmenopausal Concerns
- Cholesterol Issues
- Electrolyte Imbalances
- Diabetes & Thyroid Disorders

Primary Office
US Health Systems Medical Center
1111 E Columbia St #107
+1 206-296-6300

Office Hours
Monday 7:30 AM – 4:00 PM
Tuesday 7:30 AM – 4:00 PM
Wednesday 7:30 AM – 4:00 PM
Thursday 7:30 AM – 4:00 PM
Friday 7:30 AM – 4:00 PM

FAQ
Will Dr. Schroeder accept my insurance?
How do I make an appointment with Dr. Schroeder

SchemaApp
**Leverage Schema.org properties to describe entity**

<table>
<thead>
<tr>
<th>Property</th>
<th>Expected Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>availableService</td>
<td>MedicalProcedure or MedicalTest or MedicalTherapy</td>
<td>A medical service available from this provider.</td>
</tr>
<tr>
<td>hospitalAffiliation</td>
<td>Hospital</td>
<td>A hospital with which the physician or office is affiliated.</td>
</tr>
<tr>
<td>medicalSpecialty</td>
<td>MedicalSpecialty</td>
<td>A medical specialty of the provider.</td>
</tr>
<tr>
<td>currencyAccepted</td>
<td>Text</td>
<td>The currency accepted.</td>
</tr>
<tr>
<td>openingHours</td>
<td>Text</td>
<td>The general opening hours for a business. Opening hours can be specified as a weekly time range, starting with days, then times per day. Multiple days can be listed with commas, separating each day. Day or time ranges are specified using a hyphen (e.g., Mon-Thu 9am-5pm). For example, 3pm is specified as 15:00, 10:00am as 10:00. Here is an example: <code>&lt;time itemprop=&quot;openingHours&quot; datetime=&quot;Tu, Th 16:00-22:00&quot;/&gt;Wednesdays and Thursdays 4-8pm/time&gt;. If a business is open 7 days a week, then it can be specified as </code>&lt;time itemprop=&quot;openingHours&quot; dateInput=&quot;Mo-Su&quot;&gt;Monday through Sunday, all day&lt;/time&gt;.`</td>
</tr>
<tr>
<td>paymentAccepted</td>
<td>Text</td>
<td>Cash, Credit Card, Cryptocurrency, Local Exchange Trading System, etc.</td>
</tr>
<tr>
<td>priceRange</td>
<td>Text</td>
<td>The price range of the business, for example $$.</td>
</tr>
</tbody>
</table>
Properties used to describe the Physician entity and its relationship to other entities on the page/the website:
Example

Connect entities on a page to other entities within your organization / website

@type: Physician
URL
name
telephone
description
openingHours
parentOrganization
  @type: MedicalOrganization
  name: Washington Health
  url: www.washingtonhealth.com
  description: Washington Health is a healthcare...
sameAs: www.twitter.com/washington-health
sameAs: www.facebook.com/washington-health
areaServed: https://www.wikidata.org/wiki/Q61
medicalSpecialty
availableService
Example

Connect the primary entity on a page to the relevant rich result

@type: Physician
URL
name
telephone
description
openingHours
parentOrganization
medicalSpecialty
availableService
aggregateRating
@type: AggregateRating
bestRating: 5
ratingValue: 4.5
reviewCount: 178
worstRating: 1
Example

Nesting links to external authoritative knowledge bases

@type: Physician
  URL
  name
  telephone
  description
  openingHours
  parentOrganization
  medicalSpecialty
    @type: MedicalSpecialty
      name: plastic surgery
      url: www.washingtonhealth.com/service/plastic-surgery
      sameAs: https://www.wikidata.org/wiki/Q182442
  availableService
What can you do to prepare your content for Generative AI search?

- Design your content to talk about specific topics / entities with depth
- Ensure content on your site covers all the relevant subtopics related to your entity
- Implement Schema Markup across your entire site to describe the entity and their relationship to other things – on page, on site or external knowledge bases
Questions?

Thanks for tuning in!

Martha van Berkel
CEO & Co-Founder, Schema App
Martha@SchemaApp.com
www.schemaapp.com

Mike King
CEO & Founder, iPullRank | CMO, AIPRM
Mike@iPullRank.com
https://ipullrank.com/

Ready to implement semantic Schema Markup at scale and build your marketing knowledge graph?

Get in touch with us!

Discover the latest in generative AI and how it can be a force multiplier for your business.

Install the AIPRM Chrome Extension for Free.
This model is a lot stronger in the neural network environment because Google can capture more meaning in the vector representations.
Sentence Embeddings Are the Answer to Scaling

- The more related keywords will be closer in 10-dimensional space than the unrelated keywords.

- Notice how the “nike sneakers” keyword is far away from the “[state] bike trails” keywords.

Try it out with the Sentence Transformer visualization: https://huggingface.co/spaces/radames/sentence-embeddings-visualization