Semantic Triples, Graphs and Schema

How machines turn sentences into facts, how those facts connect into a knowledge graph, and how to write content and schema that state them clearly and correctly.

Teaching series · Lesson 2 · intermediate

EVERY FACT IS THREE PARTS

SUBJECT	PREDICATE	OBJECT
Wayne Gretzky	played for	Edmonton Oilers

Answers the question: **“What team did Wayne Gretzky play for?”**

50	6	1
minutes	rewrite rules	schema exercise

Ameneh Saeednia
Senior SEO and Generative Search Strategist

Before you start

What this lesson covers

Search engines and AI systems store much of what they know as **simple facts with three parts**. Writing sentences and markup that express those facts clearly is one of the most practical semantic SEO skills.

By the end of this lesson you can

- Define a **semantic triple** and identify the subject, predicate and object.
- **Extract triples** from a sentence and find the main one.
- **Rewrite** vague sentences so the main entity sits in the first, clearest triple.
- Spot **wrong triples**, which are worse than missing ones.
- Connect triples into a **small knowledge graph** and express it in **valid schema.org JSON-LD**.

Before this lesson

Lesson 1, *What Is a Classifier?*, is helpful but not required. You should know what an **entity** is: a uniquely identifiable thing, such as a person, place or organization.

How to use it

Read Parts 1 to 6, do the **exercise on page 9**, then take the **quiz on page 10**. Answers and sources are on the last page.

Part 1

What a triple is

A semantic triple is a fact in three parts: a **subject**, a **predicate** (the relationship) and an **object**. It's the basic unit of the W3C's RDF data model, and the building block of knowledge graphs. [1]

Wayne Gretzky	→ played for →	Edmonton Oilers
Pantheon	→ located in →	Rome
The cat	→ has color →	orange

Why machines like them

Each triple answers a question directly. **"Which team did Wayne Gretzky play for?"** has a clean answer because the fact is stated as subject, relationship, object.

Subject and object are entities or values

The subject is usually an **entity**. The object can be another entity (Edmonton Oilers) or a **value** (orange, 3 hours, €180). The predicate is what makes the fact precise.

Part 2

Finding triples in a sentence

A single sentence often holds several triples. Start by marking the **nouns, verbs and adjectives**, then find the **main verb**: the one that isn't inside a side clause.

Example

"The quick brown fox jumps over the lazy dog."

Triple	Where it comes from	Importance
The fox → jumps over → the dog	Main verb connecting the two nouns	**Main triple**
The fox → is → quick	Adjective on the subject	Supporting
The fox → is → brown	Adjective on the subject	Supporting
The dog → is → lazy	Adjective on the object	Supporting

What machines do here

NLP systems use **dependency parsing** to find the root verb and what attaches to it. You don't need the software to apply the idea: ask **"who does what to whom?"** for each sentence, and check that the answer is the fact you meant to state.

Part 3

Rewriting for a clearer main triple

Opening sentences are often written for mood, not meaning. That leaves the main entity buried.

Before: main triple is “[you] welcome to a new era”

```
Welcome to a new era of Rome travel; an unforgettable
experience that offers everything you love and
nothing you don't.
```

After: main triple is “walking tours → show → sights”

```
Private walking tours in Rome show families the
city's main sights in three hours, with a local guide
who adapts the pace for children.
```

1

Put the entity first

The main entity should be the **subject of the first triple** on the page.

2

One claim per sentence

Long sentences with three ideas produce **three weak triples**.

3

Choose precise verbs

“Offers,” “is located in,” “is led by” say more than “provides an experience.”

4

Cut the fluff

“Unforgettable” and “everything you love” create **no usable fact**.

5

State values, not vibes

“**3 hours**,” “**€180**,” “**ages 6 and up**” are objects a system can store.

6

Keep the same name

Call it the same thing each time, so every triple points at **one entity**.

Part 4

A wrong triple is worse than none

In a 2023 teaching example, a sports-betting page was summarized as a list of “type of” triples. Several were **factually wrong**, and machines can't tell.

Triple as written (2023)	Problem	Corrected triple
NFL → is a type of → sports betting	The NFL is a league, not a bet type	NFL → is → a professional American football league
Prop bets → are a type of → NFL betting odds	Prop bets exist in every sport	Prop bets → are a type of → wager on a specific event in a game
Odds Shark's NFL Bible → is a guide on → NFL betting	Fine, but vague	[Guide] → covers → NFL betting lines, totals and futures

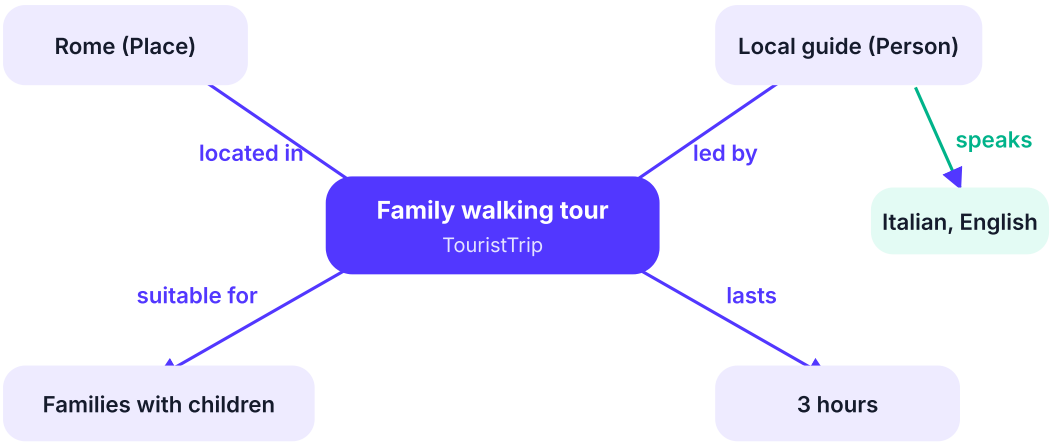
The rule

Before you add a triple to content or markup, ask: **would an expert agree this is true as stated?** “Is a type of” is the most misused predicate in SEO writing. Use it only for genuine categories.

Part 5

From triples to a knowledge graph

When triples share entities, they **connect**. That network is a knowledge graph: entities as nodes, relationships as edges.



Reading the graph

Five triples about one tour: **located in** Rome, **led by** a local guide, **suitable for** families, **lasts** 3 hours, and the guide **speaks** Italian and English. Each one answers a likely question.

Where this graph lives

On your site, the same graph can appear twice: in **sentences** people read, and in **structured data** machines read. Google says it uses structured data to understand page content and information about the world. [2]

Part 6

Expressing the graph in schema

Schema.org markup is a vocabulary for writing triples in JSON-LD. **It only works if you use real types and properties.**

2023 example: looks structured, isn't valid

```
{
  "@context": "http://schema.org",
  "@type": "WebPage",
  "headline": "NFL Betting Guide",
  "articleBody": "This guide provides information
...",
  "properties": {
    "nfl": {
      "@type": "Thing",
      "name": "NFL",
      "description": "A type of sports betting"
    }
  }
}
```

Valid: TouristTrip with real properties

```
{
  "@context": "https://schema.org",
  "@type": "TouristTrip",
  "name": "Family walking tour of Rome",
  "description": "A 3-hour family walking tour of
Rome.",
  "touristType": "Families with children",
  "provider": {
    "@type": "Person",
    "name": "[Guide name]",
    "knowsLanguage": ["it", "en"]
  },
  "itinerary": {
    "@type": "ItemList",
    "itemListElement": [
      {"@type": "ListItem", "position": 1,
       "item": {"@type": "Place", "name":
"Pantheon"}},
      {"@type": "ListItem", "position": 2,
       "item": {"@type": "Place", "name": "Trevi
Fountain"}}
    ]
  },
  "offers": {"@type": "Offer", "price": "180",
             "priceCurrency": "EUR"}
}
```

What was wrong

Schema.org has no **“properties”** property, so that whole block means nothing to a parser. **articleBody** belongs to Article, not WebPage. And the markup repeated the wrong triple about the NFL.

What makes it right

Every key is a **defined schema.org property** of its type. [3] Values match the **visible page**, as Google's guidelines require. [4] TouristTrip may not create a rich result, but it still describes the entities clearly.

Exercise

Map one page into triples and schema

Time: 20 minutes. Use one of your own service, product or tour pages.

Step	What to do
1	Copy the **first three paragraphs**. Write down every triple you can find, up to 10.
2	Mark the **main triple** of each sentence. Is the page's main entity the subject of the first one?
3	Flag any triple that is **vague, wrong or unverifiable**.
4	Rewrite **three sentences** using the six rules from Part 3.
5	Draw the triples as a **small graph** like the one on page 7.
6	Write JSON-LD for the main entity using **real schema.org types**. Test it in the **Schema Markup Validator** and **Rich Results Test**.

What good looks like

The main entity is the subject of the **first triple**. No triple is wrong. Every schema property is **defined** and matches visible text.

Discussion question

Which triple on your page would a customer most need to be correct? Is it stated in **both the text and the markup**?

Check your understanding

Quick quiz

Answers are on the next page.

1. In “The Pantheon is located in Rome,” the predicate is:

a) The Pantheon
b) located in
c) Rome

2. The best first sentence for a page about Rome food tours is:

a) Welcome to an unforgettable culinary journey!
b) Rome food tours take small groups to family-run trattorias in Trastevere.
c) We love food and so will you.

3. “The NFL is a type of sports betting” is a problem because:

a) It is too long
b) It is factually wrong
c) It has no predicate

4. In JSON-LD, a block called “properties” inside WebPage will:

a) Create a rich result
b) Be ignored, because it isn't a schema.org property
c) Define new entity types

5. Schema markup should describe:

a) What the page visibly says
b) Extra keywords you'd like to rank for
c) Competitors' products

Answers and sources

Answer key, glossary and further reading

Answer key

1: **b**.
2: **b**. Entity first, precise verb, concrete detail.
3: **b**. A wrong fact stated clearly is still wrong.
4: **b**. Undefined properties carry no meaning.
5: **a**. Markup must represent visible content.

Glossary

Triple: subject, predicate, object.
Predicate: the relationship in a triple.
Knowledge graph: connected entities and relationships.
JSON-LD: a format for writing structured data.
Schema.org: a shared vocabulary of types and properties.
Dependency parsing: finding how words in a sentence relate.

01 **RDF 1.1 Concepts and Abstract Syntax. W3C**
The formal definition of triples as subject, predicate and object.
https://www.w3.org/TR/rdf11-concepts/

02 **Introduction to structured data markup. Google Search Central**
How Google says it uses structured data to understand pages.
https://developers.google.com/search/docs/appearance/structured-data/intro-structured-data

03 **TouristTrip. Schema.org**
Type and properties used in the valid example.
https://schema.org/TouristTrip

04 **General structured data guidelines. Google Search Central**
Markup must represent the page's visible content.
https://developers.google.com/search/docs/appearance/structured-data/sd-policies

Lesson by Ameneh Saeednia, adapted from 2023 training notes in my archive and prepared with AI assistance. Examples are illustrative. Related reading on seogirl.ca: *Semantic Triples: the Tiny Grammar Habit* and *Entity Types, Classes, and Attributes*.