

# Entity Life Cycle In Search Centric Knowledge Graphs

---

By Amgad Madkour

Data & Applied Scientist  
Microsoft



---

## Outline

1

### Introduction

Introduce search centric knowledge graphs and highlight how different they are from other graphs.

2

### Entity Life Cycle

Present the life cycle of an entity including crawling, extraction, ingestion, and publishing.

3

### Showcases

Demonstrate some of the different experiences generated for an entity.



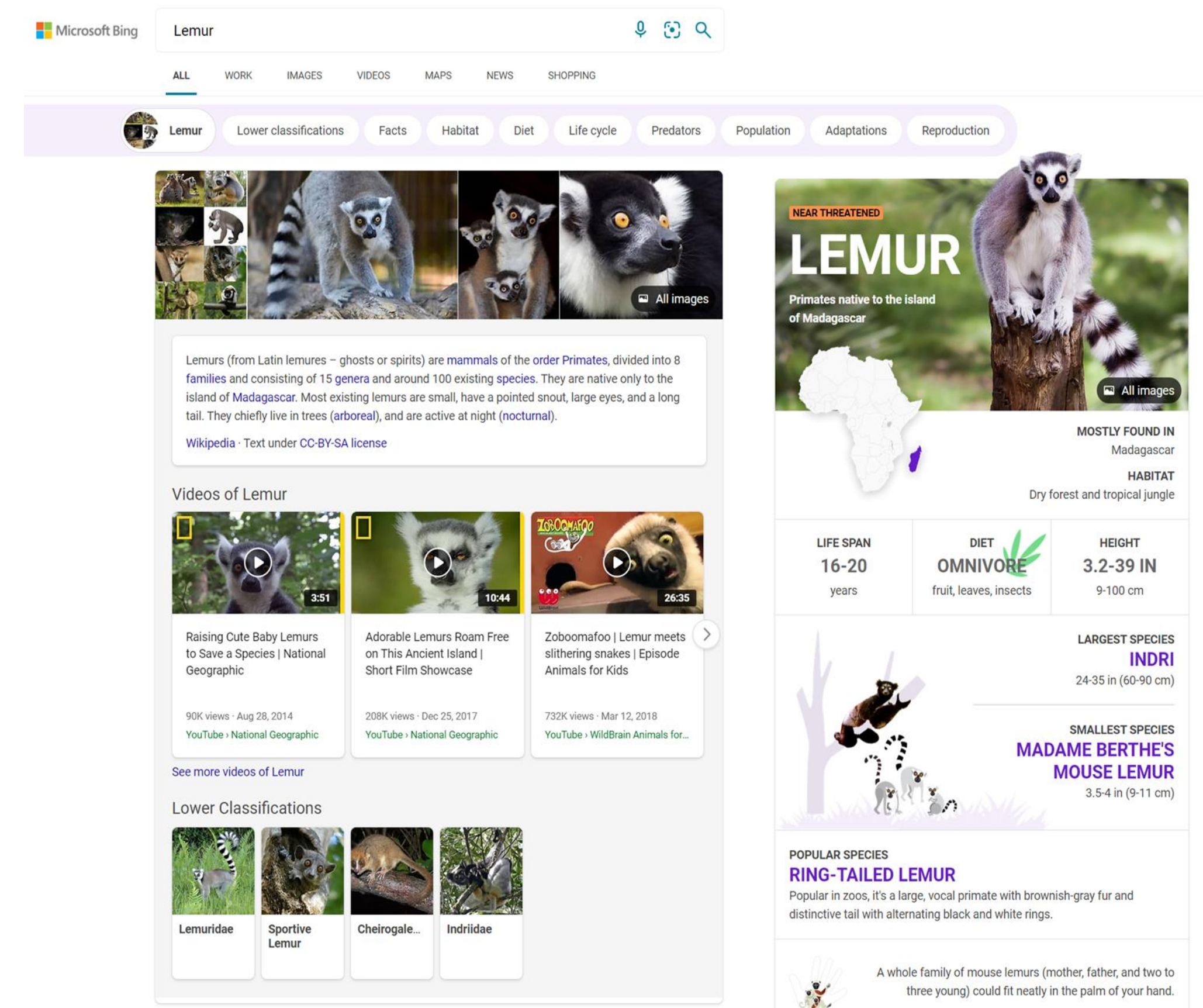


# Introduction



# Entity Experience

- **Entity-based results** are becoming an **integral** part of the search experience.
- Search companies highly rely on knowledge graphs for building **rich search experiences**.
- An entity experience includes **factual, textual, and video** information.



**Figure:** Entity experience for the Lemur that includes description, videos and a knowledge card



---

# Customer Obsession

- Search centric knowledge graphs are built around entities that bring **value** to the search **customer**.
- Knowledge graphs are considered a **mean** and not a **goal** for many companies.
- The **relevance** and **quality** of the data is more important than the **amount of information** about an entity.





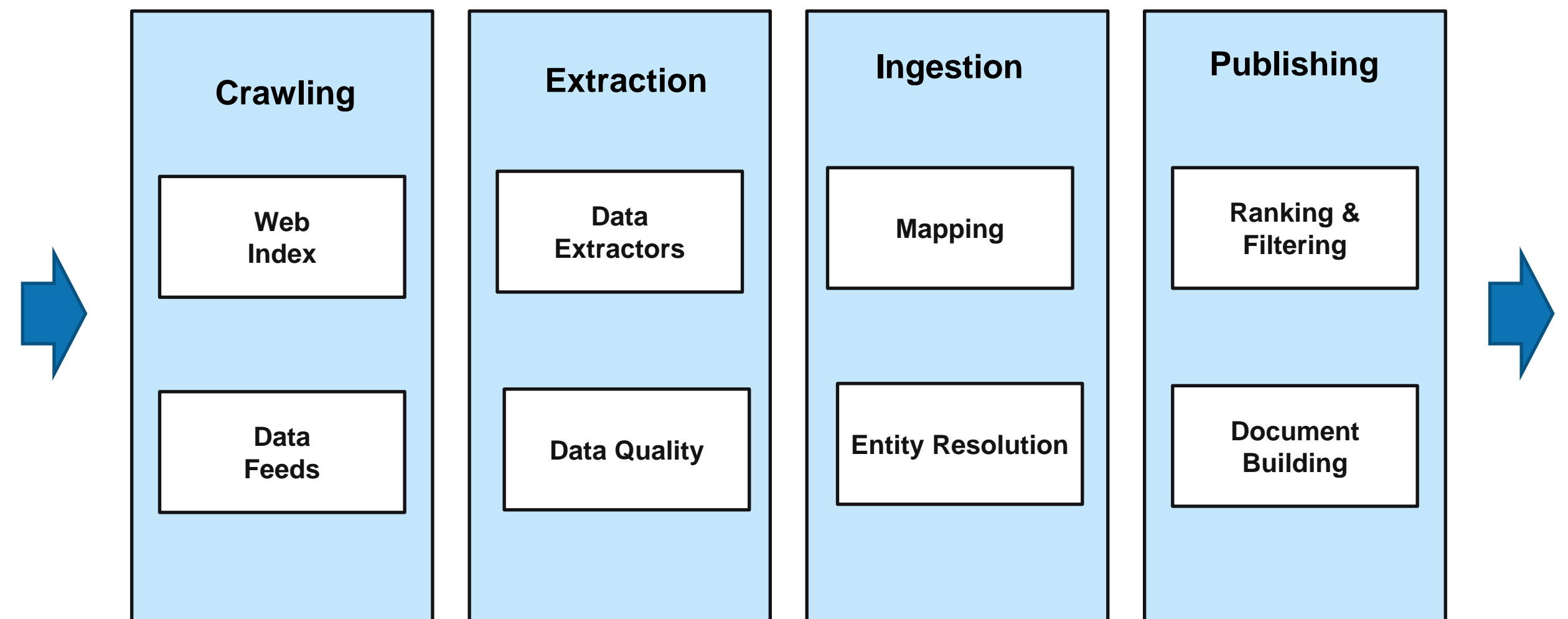
---

# Entity Life Cycle



# Overview

- An entity passes through a **series of stages** to be onboarded into a knowledge graph.
- The life cycle stages are **crawling**, **extraction**, **ingestion**, and **publishing**.
- In this talk, the focus will be on the fundamental components of the entity life cycle only.

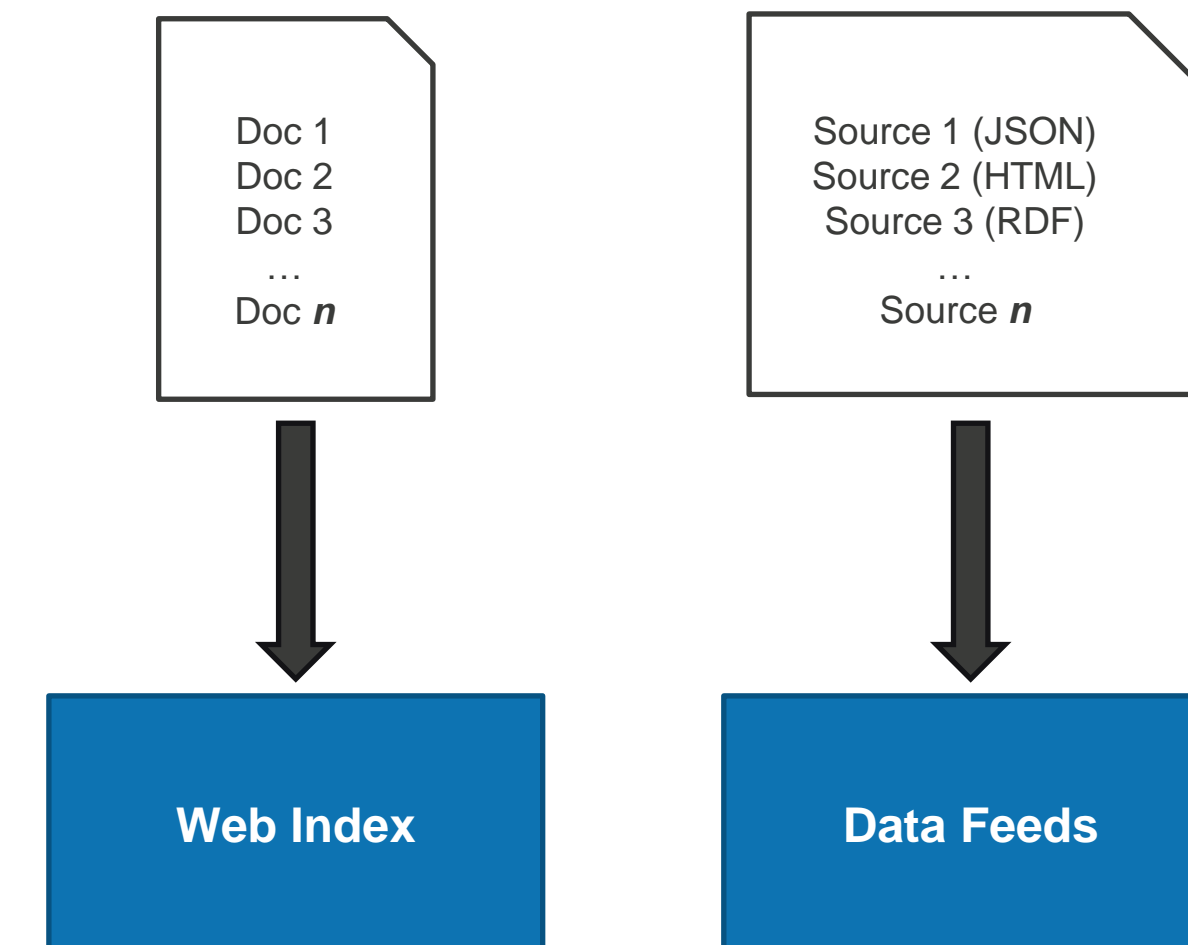


**Figure:** The fundamental stages of an entity's life cycle



# Crawling

- Entities are discovered from **hundred of sources**.
- The **web index** is one rich source for entities.
- Entities not in the index can be **crawled separately**.
- **Data feeds** provided by **partners** are another sources of entity information.

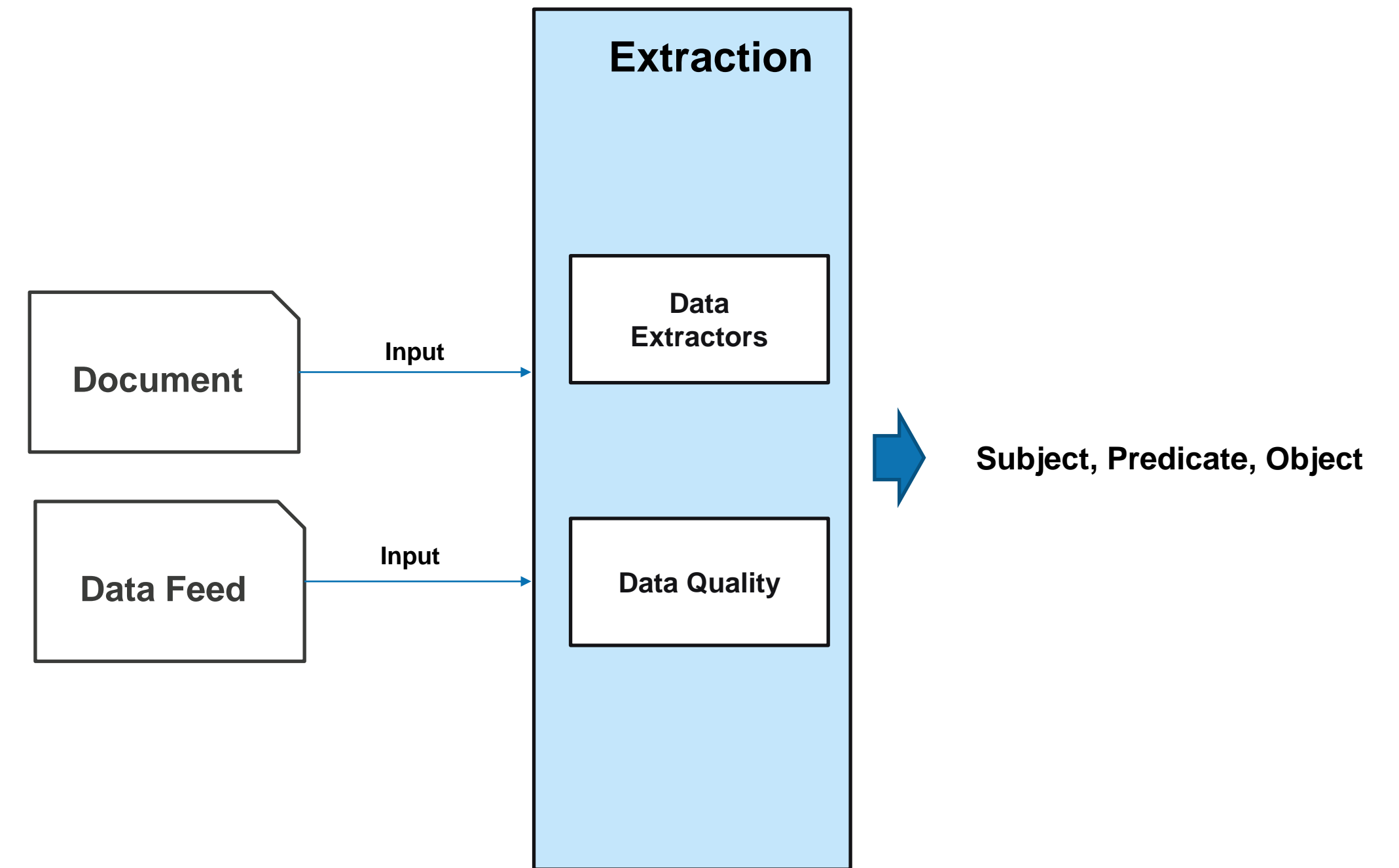


**Figure:** The input sources could be from a web index (documents), or data feeds (JSON,HTML, RDF)



# Extraction

- Once a source is available, the extraction converts data into **RDF** (i.e. **Subject, Predicate, Object**).
- **Data extractors** utilize **machine learning** techniques to handle different sources.
- **Data quality** is *mainly* employed during the extraction stage where **pre-specified rules** and **machine learning** are used to normalize and clean the data.
- Some **pre-specified rules** are expressed in **SPARQL**.

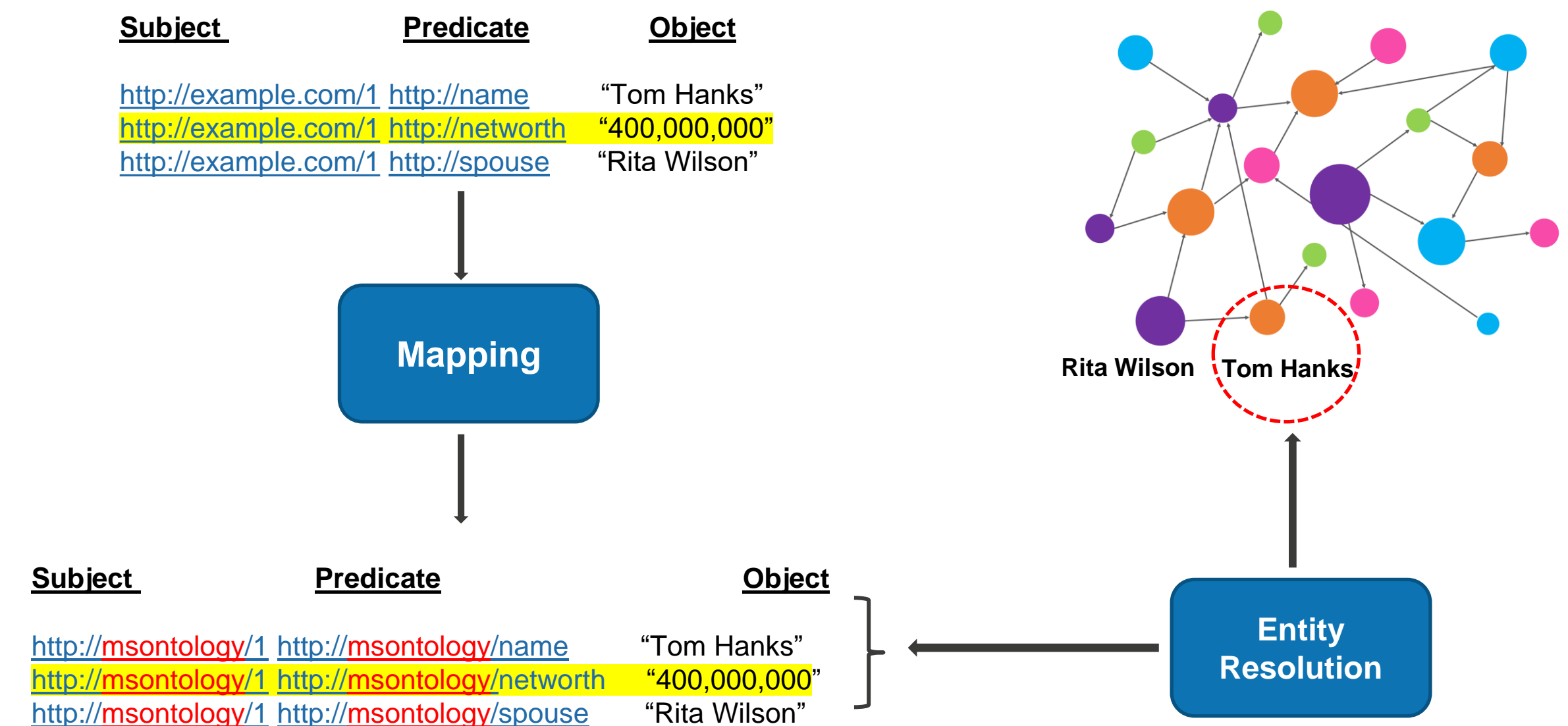


**Figure:** The extraction stage takes documents and feeds sources as input and creates triples as output



# Ingestion

- The ingestion phase includes **mapping** and **entity resolution**.
- The mapping component connects the **predicates of the source** to ones defined by the **ontology** of the knowledge graph.
- Entity resolution uses the mapping to **match entities** from the **source** against the **knowledge graph**.
- Hybrid approaches (**rule-based** and **machine learning**) are usually employed by both components.

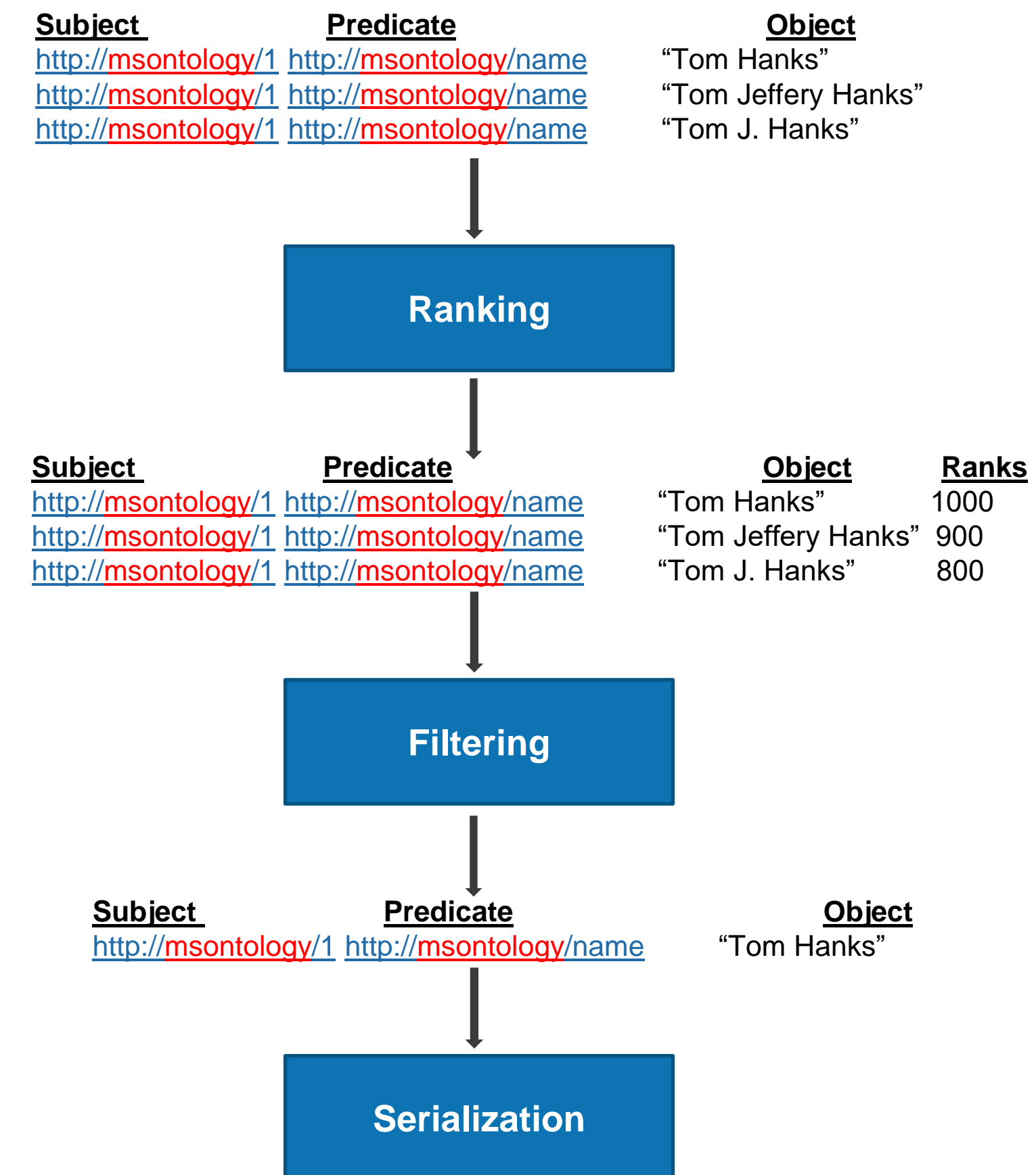


**Figure:** Example of adding Tom Hanks net worth from one of the sources to the knowledge graph using mapping and entity resolution



# Publishing

- Entities are **ranked** and **filtered** according to multiple aspects  
(e.g., source, predicate, market information)
- Entities are published as a **serialized** and **size bounded** document of facts using a **document builder**.
- This **document** is **replicated** across multiple **data centers** in order to be served in a matter of milliseconds.



**Figure:** Entity facts pass through ranking, filtering, and serialization stages



# Multilingual Aspect

- A **source** (e.g. Wikipedia) can supply facts in **multiple languages**.
- There are **hundreds** of languages that can have graphs indicating **different** or the **same** set of **facts** about an entity.
- **Selectivity** about the information in the graph and processed in every update is **mandatory**.
- Consolidation techniques on an entity level can help reduce the consumed resources of an entity.

English Market			Arabic Market		
<u>Subject</u>	<u>Predicate</u>	<u>Object</u>	<u>Subject</u>	<u>Predicate</u>	<u>Object</u>
<a href="http://example.com/1">http://example.com/1</a>	<a href="http://name">http://name</a>	"Tom Hanks"	<a href="http://example.com/1">http://example.com/1</a>	<a href="http://name">http://name</a>	"توم هانكس"
<a href="http://example.com/1">http://example.com/1</a>	<a href="http://networth">http://networth</a>	"400,000,000"	<a href="http://example.com/1">http://example.com/1</a>	<a href="http://networth">http://networth</a>	"400,000,000"

**Figure:** The same fact about an entity could be redundantly stored in two separate sources, resulting in two facts in the knowledge graph



---

# Showcases



\_\_\_\_\_



**Figure:** The lower classification of the Lemur entity is exposed in the Lemur entity experience



# Entity Carousel

Microsoft Bing

united states presidents

ALL

WORK

NEWS

IMAGES


VIDEOS

MAPS


SHOPPING

Also try: [list us presidents in order](#) · [list of all u s presidents in order](#)


United States - Presidents




Joe Biden  
Since 2021




Donald Trump  
2017 - 2021




Barack Obama  
2009 - 2017




George W. Bush  
2001 - 2009




Bill Clinton  
1993 - 2001




George Bush  
1989 - 1993




Ronald Reagan  
1981 - 1989



Jimmy Carter  
1977 - 1981



Gerald Ford  
1974 - 1977



Richard Nixon  
1969 - 1974

List of presidents of the United States - Wikipedia

[https://en.wikipedia.org/wiki/List\\_of\\_Presidents\\_of\\_the\\_United\\_States](https://en.wikipedia.org/wiki/List_of_Presidents_of_the_United_States)

84 rows · Apr 12, 2003 · The president of the United States is the head of state and head of government of the United ...

PRESIDENCY	PRESIDENCY	PRESIDENT	PARTY
46	January 20, 2021 – Incumbent	Joe Biden	Democratic
45	January 20, 2017 – January 20, 2021	Donald Trump	Republican
44	January 20, 2009 – January 20, 2017	Barack Obama	Democratic
43	January 20, 2001 – January 20, 2009	George W. Bush	Republican

See all 84 rows on [en.wikipedia.org](https://en.wikipedia.org)


Presidents | The White House

<https://www.whitehouse.gov/about-the-white-house/presidents>


- George Washington. The 1st President of the United States.
- John Adams. The 2nd President of the United States.
- Thomas Jefferson. The 3rd President of the United States.
- James Madison. The 4th President of the United States.

See full list on [whitehouse.gov](https://www.whitehouse.gov)

UNITED STATES



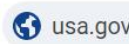
All images



COUNTRY IN NORTH AMERICA

The foremost military power in the world and a leading political, cultural, and scientific force internationally.

W




CAPITAL

WASHINGTON, D.C.

OFFICIAL LANGUAGE

ENGLISH




**Figure:** The carousel experience for the “United states presidents” query


KGCC

The Knowledge Graph Conference




# Fact Gamification

 **Tom Cruise** | Movies | Upcoming movies | Awards | Quotes | Recent highlights | Interviews | Fashion | Merchandise | Biography




All images


### News about Tom Cruise




**Tom Cruise's \$39.5M Colorado Ranch Is This Week's Most Popular Home**  
Hollywood legend Tom Cruise wants to sell his 320-acre Telluride, CO, ranch. Fans of the perennial A ...  
SFGate · 17h



Tom Cruise and Nicole Kidman's Son Connor Unveils Surprising New Career Venture  
E News · 2d




Tom Cruise and Nicole Kidman's rarely-seen son Connor, 26, launches new venture as a 'grill master'  
Daily Mail · 2d




Tom Cruise, Nicole Kidman's Son Connor Announces New Career Path  
ET Canada on MSN.co... · 2d

[See more news about Tom Cruise](#)




**TOM CRUISE**  
American Actor  
W | IG | IMDb | F | TW  
All images


An actor and producer and is one of the **highest-paid actors in the world**.

	<b>SPOUSE</b> Katie Holmes 2006 - 2012	<b>BORN</b> Jul 03, 1962 age 58	<b>HEIGHT</b> 5'7" 1.7 m
---	--	---------------------------------------	--------------------------------


### MOVIES AND TV SHOWS




**Top Gun: Maverick**  
2021




**Mission: Impossible ...**  
2018



**Top Gun**  
1986




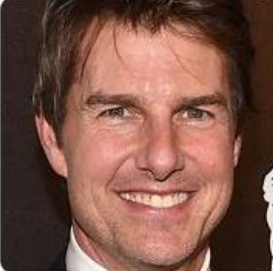
**Edge of Tomorrow**  
2014



**Mission: Impossible 7**  
2021




 **Tom Cruise** | **Height** | Movies | Upcoming movies | Awards | Quotes | Recent highlights





Tom Cruise · Height  
**5' 7"**  
(1.70 m)


### Trivia

Who's closest in height to Tom Cruise?

A   
Will Smith

B   
Sylvester Stallone

C   
Robert Downey Jr.

D   
Jennifer Connelly

[Feedback](#)

**Figure:** An entity fact is used to show facts about other entities (e.g. via trivia Q/A)



---

# #KGC2021

Join the Conversation



@KGConference

@**AmgadMadkour**



[linkedin.com/company/the-knowledge-graph-conference/](https://linkedin.com/company/the-knowledge-graph-conference/)



[youtube.com/playlist?list=PLAiy7NYe9U2Gjg-600CTV1HGypiF95d\\_D](https://youtube.com/playlist?list=PLAiy7NYe9U2Gjg-600CTV1HGypiF95d_D)

