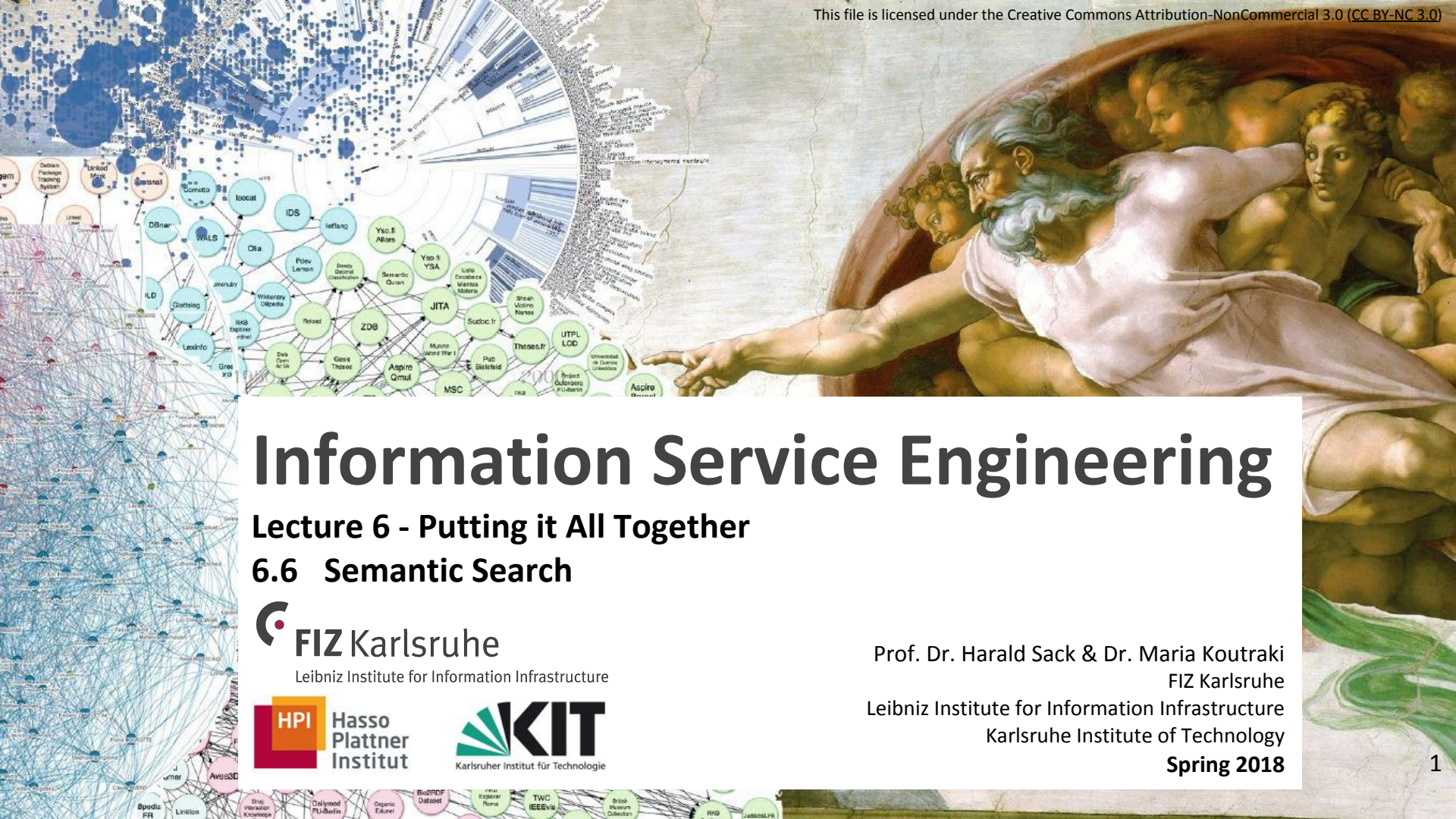




Information Service Engineering

Lecture 6 - Putting it All Together

6.6 Semantic Search



FIZ Karlsruhe

Leibniz Institute for Information Infrastructure



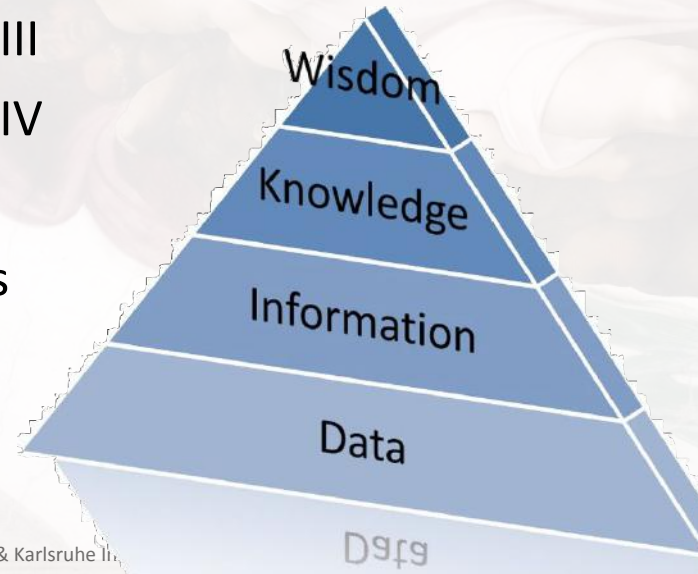
Prof. Dr. Harald Sack & Dr. Maria Koutraki
FIZ Karlsruhe

Leibniz Institute for Information Infrastructure
Karlsruhe Institute of Technology
Spring 2018

Information Service Engineering

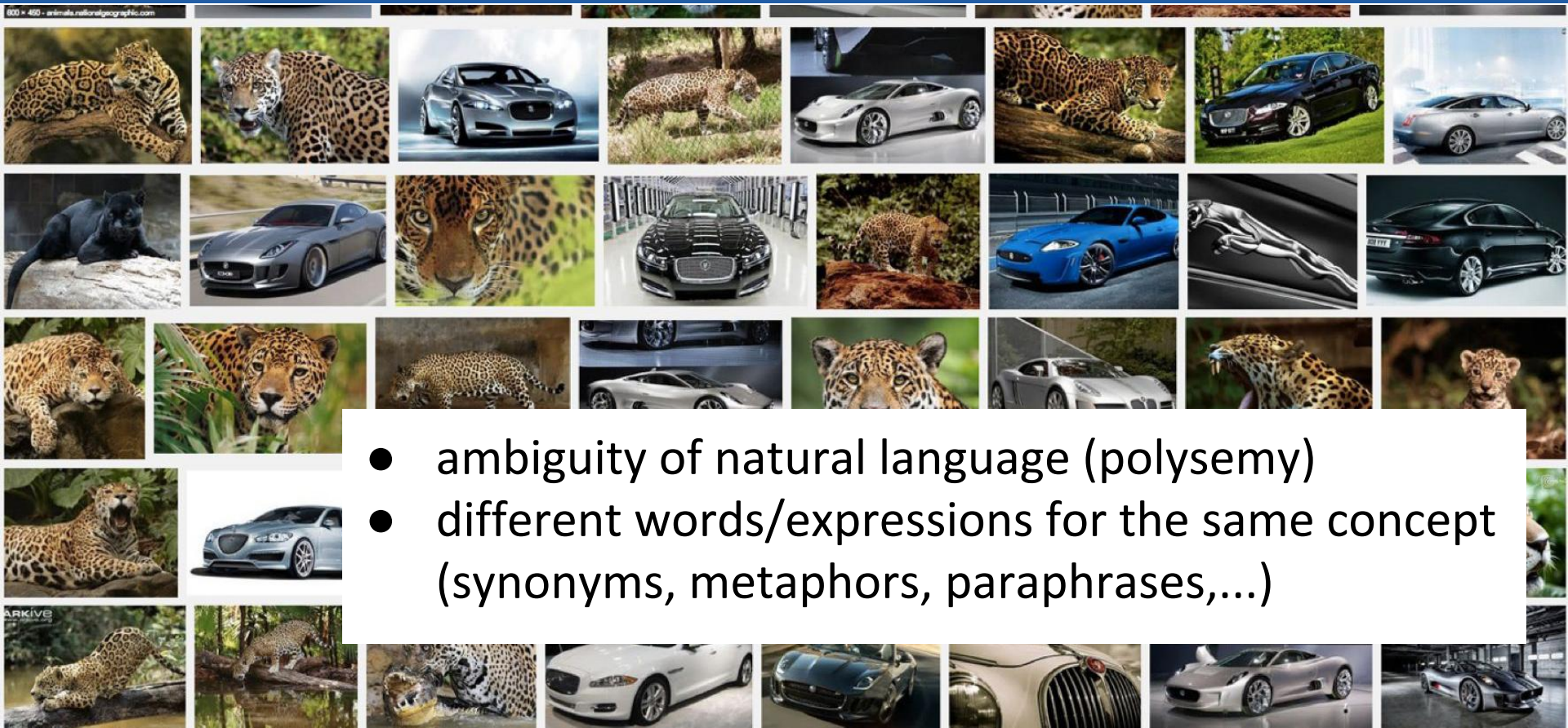
Lecture 6: Putting it All Together

- 6.1 Linked Data Programming
- 6.2 Knowledge Mining and Information Extraction
- 6.3 Knowledge Mining and Information Extraction II
- 6.4 Knowledge Mining and Information Extraction III
- 6.5 Knowledge Mining and Information Extraction IV
- 6.6 Semantic Search**
- 6.7 Exploratory Search and Recommender Systems



The Information Retrieval Dilemma

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search



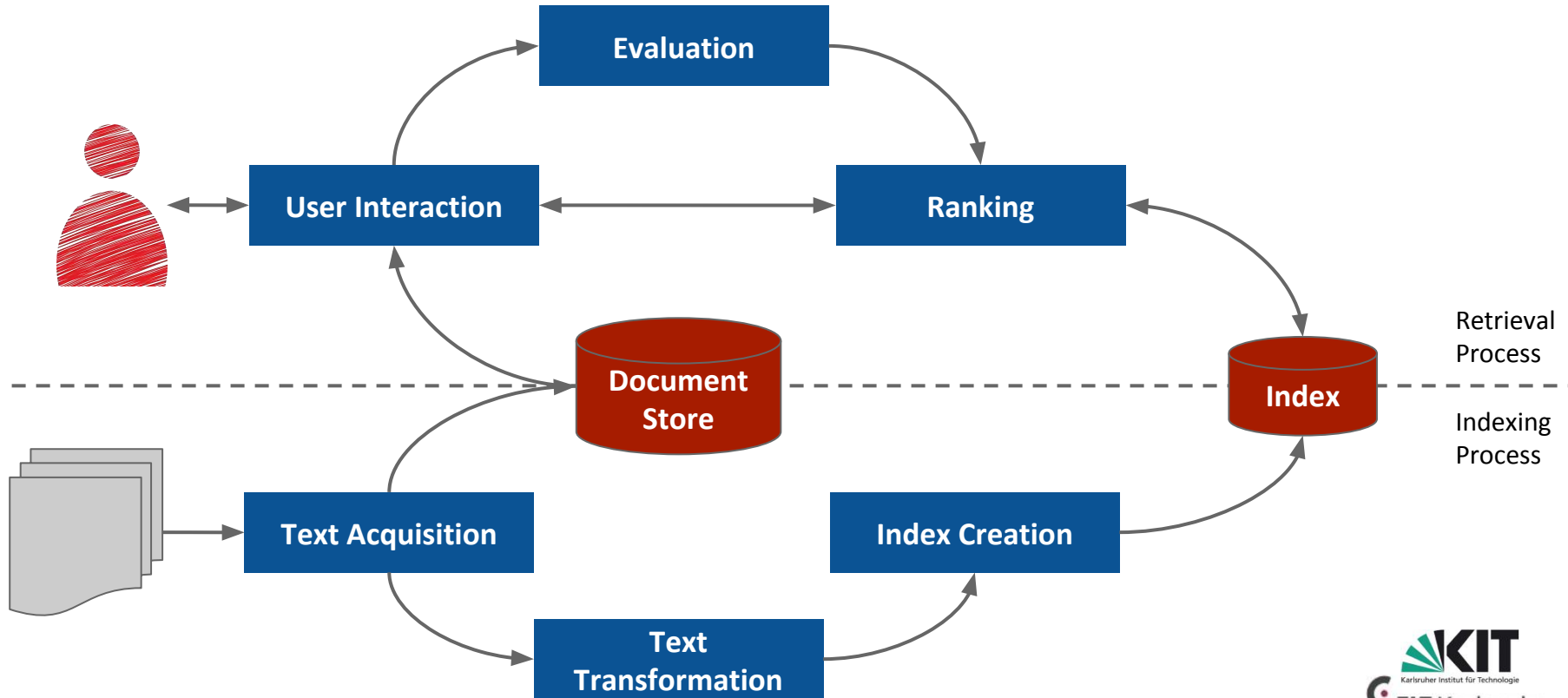
- ambiguity of natural language (polysemy)
- different words/expressions for the same concept (synonyms, metaphors, paraphrases,...)



How can Semantic Technology support IR?

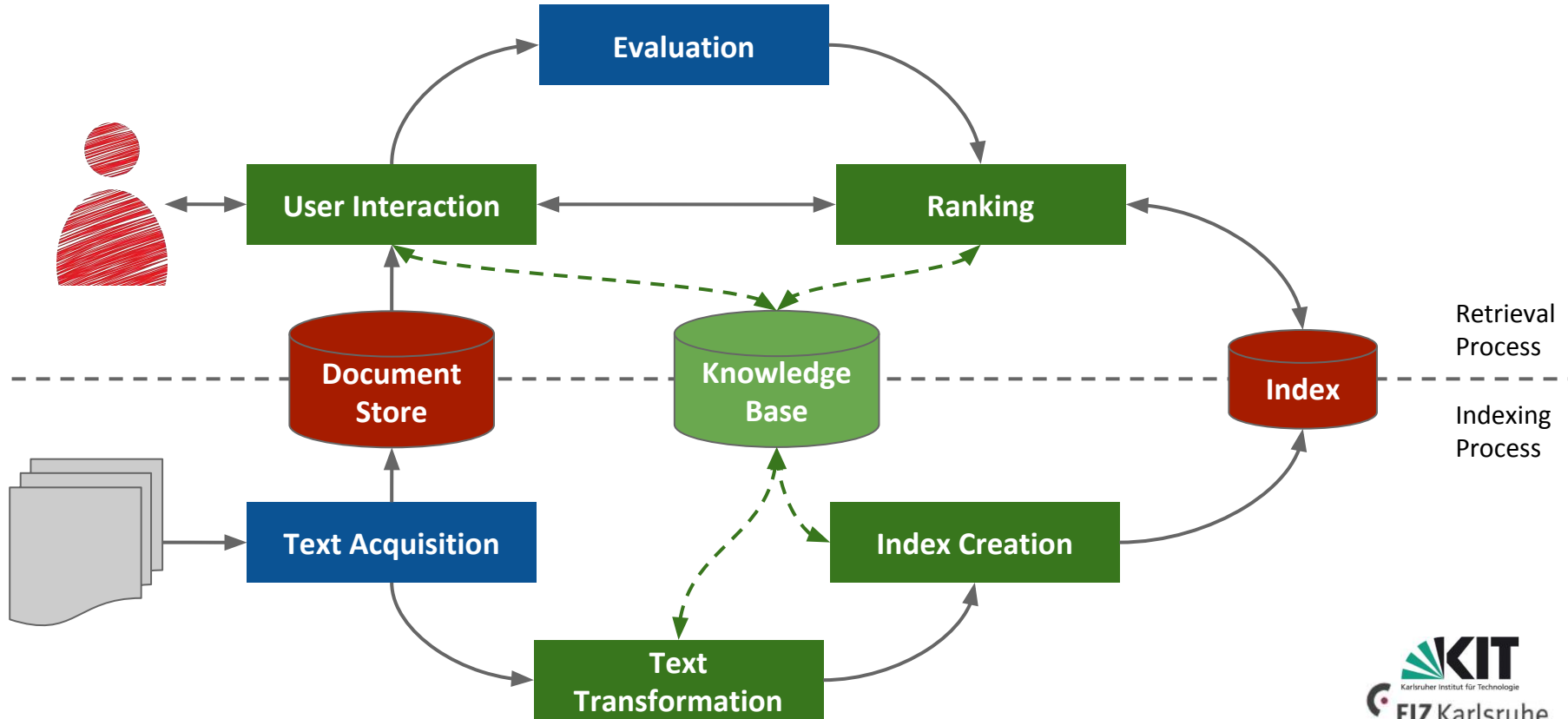
The Information Retrieval Process

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search



Semantic Technology Supported IR

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search



Semantic Technology Supported IR

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search

- **Prerequisite:**
Document Annotation with explicit semantics, e.g. semantic entities

On July 26, 1928, American film director, screenwriter, producer, cinematographer, and editor **Stanley Kubrick** was born. He is regarded as one of the greatest filmmakers of all time. His or short stories, are noted for their unique cinematography, n and an inspired use of music scores. Clockwork Orange, e Shining, 'I'm Spartacus' or Dr. Strangelove – they all are movies ever made.

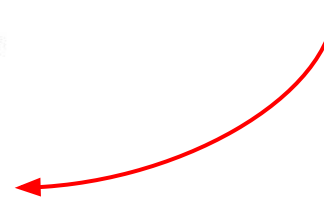
Stanley Kubrick
Stanley Kubrick (/ˈkʊːbrɪk/; July 26, 1928 – March 7, 1999) was an American film director, screenwriter, producer, cinematographer, and

birth year	1928
death year	1999
is producer of	2001: A Space Odyssey (film)
ethnicity	American Jews
is spouse of	Christiane Kubrick
occupation	Cinematographer

ink the thing was probably helped me the most was that do in the early 50s for someone who actually go and as impossible. It really is terribly easy. All anybody er, and some imagination." (Stanley Kubrick in an

district of New York, as the first of two children into a family with Jewish ancestry. As a child, Stanley was considered intelligent, but he did not achieve particularly high

Example for
Linked Data
Based Document
Annotation



<http://scih.org/breaking-new-grounds-in-cinematography-stanley-kubrick/>

- Enables **entity-based Information Retrieval**
 - Language independent

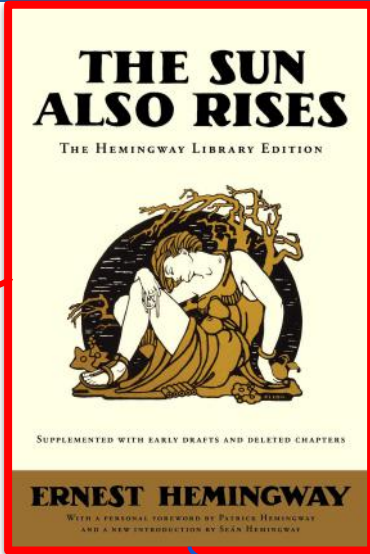
Semantic Annotation

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search

- The term “**annotation**” implies to **attach data to some other piece of data**. It establishes a **(typed) relation between the annotated data and the annotating data**
- The term “**annotation**” can denote both the **process of annotating** and the **result of that process**
- **Formal Model:**
 - An **annotation** A is a tuple (a_s, a_p, a_o, a_c) , where
 - a_s is the **subject** of the annotation (the annotated data)
 - a_o is the **object** of the annotation (the annotating data)
 - a_p is the **predicate** (the annotation relation) that defines the type of relationship between a_s and a_o , and
 - a_c is the **context** in which the annotation is made.

Semantic Annotation

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search



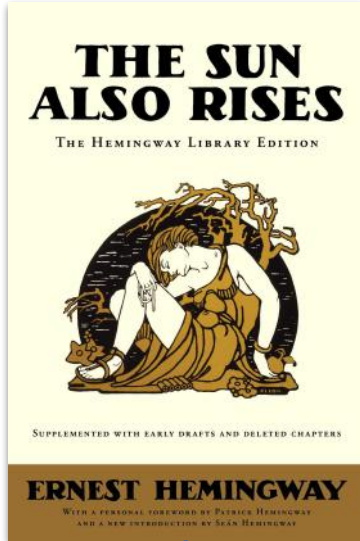
- **Formal Model:**

- An **annotation** A is a tuple (a_s, a_p, a_o, a_c) , where
 - a_s is the **subject** of the annotation (the annotated data)
 - a_o is the **object** of the annotation (the annotating data)
 - a_p is the **predicate** (the annotation relation) that defines the type of relationship between a_s and a_o , and
 - a_c is the **context** in which the annotation is made.

```
PREFIX dbr: <http://dbpedia.org/resource/> .  
PREFIX dbo: <http://dbpedia.org/ontology/> .  
  
:9787532717071 dbo:author dbr:Ernest_Hemingway .
```

Semantic Annotation

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search



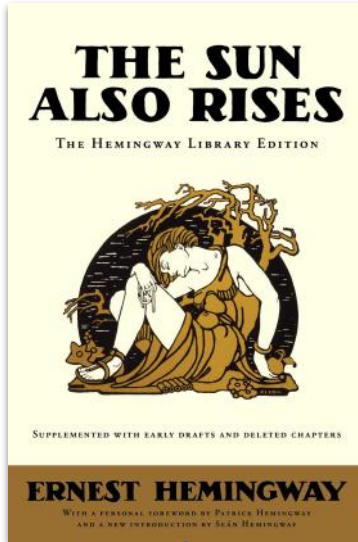
- **Formal Model:**

- An **annotation** A is a tuple (a_s, a_p, a_o, a_c) , where
 - a_s is the **subject** of the annotation (the annotated data)
 - a_o is the **object** of the annotation (the annotating data)
 - a_p is the **predicate** (the annotation relation) that defines the type of relationship between a_s and a_o , and
 - a_c is the **context** in which the annotation is made.

```
PREFIX dbr: <http://dbpedia.org/resource/> .  
PREFIX dbo: <http://dbpedia.org/ontology/> .  
  
::9787532717071 dbo:author dbpedia:Ernest_Hemingway .
```

Semantic Annotation

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search



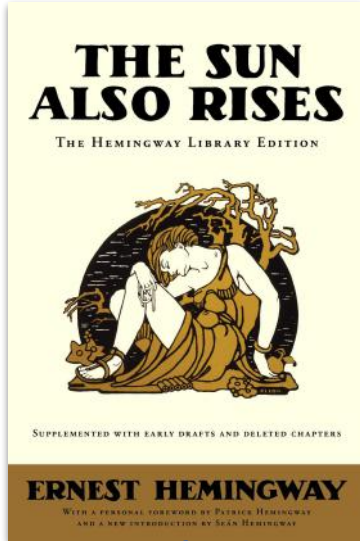
- **Formal Model:**

- An **annotation** A is a tuple (a_s, a_p, a_o, a_c) , where
 - a_s is the **subject** of the annotation (the annotated data)
 - a_o is the **object** of the annotation (the annotating data)
 - a_p is the **predicate** (the annotation relation) that defines the type of relationship between a_s and a_o , and
 - a_c is the **context** in which the annotation is made.

```
PREFIX dbr: <http://dbpedia.org/resource/> .  
PREFIX dbo: <http://dbpedia.org/ontology/> .  
  
::9787532717071 dbo:author dbr:Ernest_Hemingway .
```

Semantic Annotation

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search



- **Formal Model:**

- An **annotation** A is a tuple (a_s, a_p, a_o, a_c) , where
 - a_s is the **subject** of the annotation (the annotated data)
 - a_o is the **object** of the annotation (the annotating data)
 - a_p is the **predicate** (the annotation relation) that defines the type of relationship between a_s and a_o , and
 - a_c is the **context** in which the annotation is made.

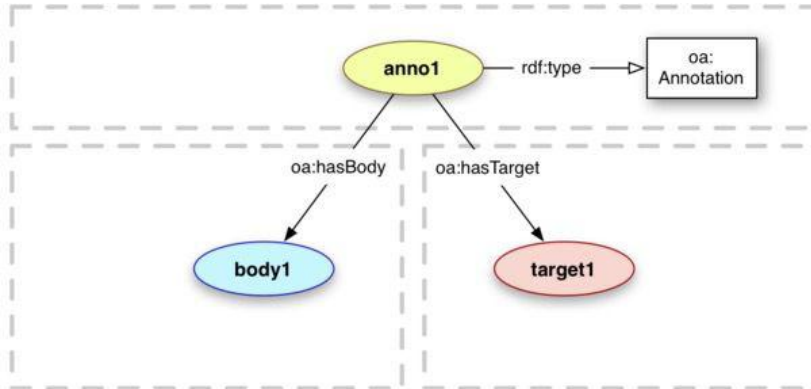
```
PREFIX dbr: <http://dbpedia.org/resource/> .  
PREFIX dbo: <http://dbpedia.org/ontology/> .  
  
::9787532717071 dbo:author dbr:Ernest_Hemingway .
```

- created at Dec 11, 2017
- created by Lysander07
- ...

Ontologies for Semantic Annotation

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search

- The **Open Annotation Ontology** is a W3C Standard (2017) [1]
 - “...specifies an interoperable framework for creating associations between related resources, annotations, using a methodology that conforms to the Architecture of the World Wide Web.”

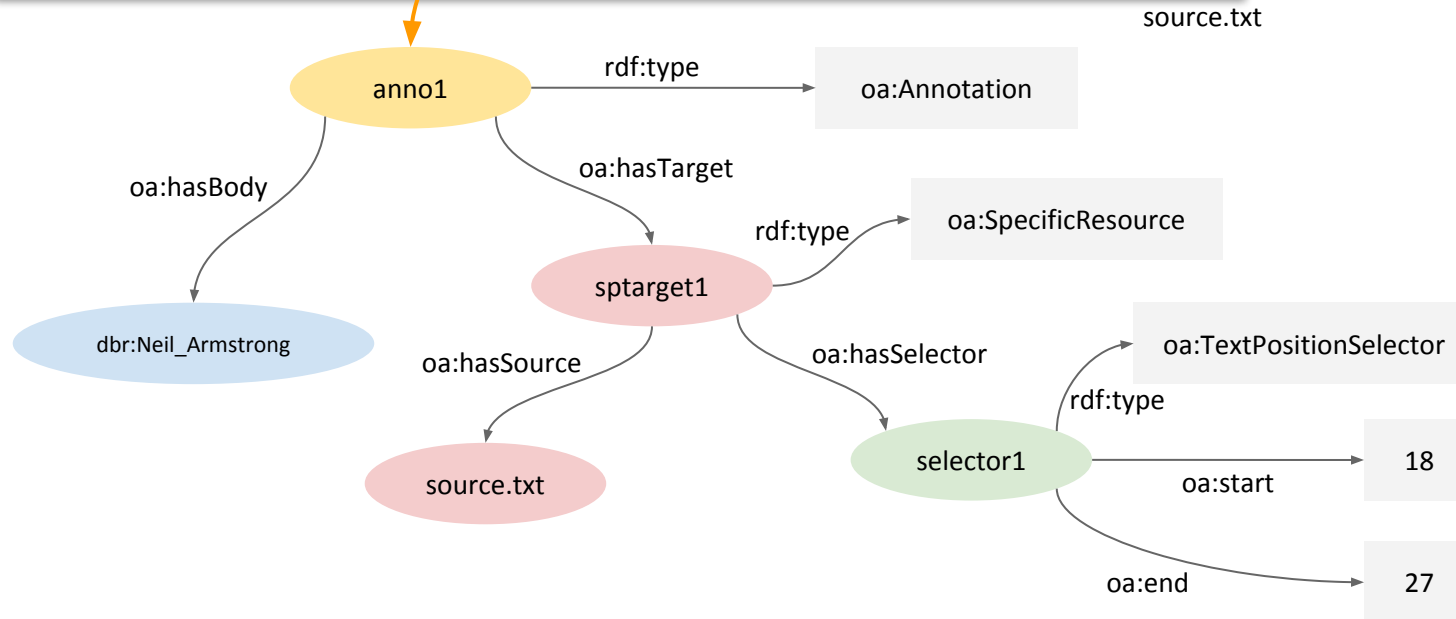


- The **Body** provides the information which is annotating the **Target**

Open Annotation Ontology Example

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search

On July 16, 1969, **Armstrong**, along with Edwin E. Aldrin, Jr., and Michael Collins, blasted off in the Apollo 11 vehicle toward the Moon.



Semantic Technology Supported IR

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search

- **Prerequisite:** Annotation with explicit semantics, e.g. semantic entities
- Enables **entity-based Information Retrieval**
 - Language independent
- Make use of underlying Knowledge Base, e.g.
 - content-based **similarities** among entities
 - content-based **relationships** between entities
- **Interoperable metadata** via semantic annotations
 - for content-based description
 - for structural and technical description
- **Content-Based Navigation** and **Result Filtering** (Search Facets)

Entity Based Search

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search

Query Processing:

Armstrong on the Moon

Named Entity Linking

dbr:Neil Armstrong

dbr:Moon

Indexing:

The first Man on the Moon

....
On the Moon, the 38-year-old civilian commander, radioes to earth and the mission control room here: "Houston, Tranquility Base here, The Eagle has landed."
....

dbr:Neil Armstrong

dbr:Moon

Named
Entity
Linking

Entity-Based Query Matching

- simple entity matching
- similarity-based entity matching
- relationship-based entity matching
- ...

Entity Based Search

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search

Query Processing:

Armstrong on the Moon

Named Entity Linking

dbr:Neil_Armstrong

dbr:Moon

Indexing

The 2nd Man on the Moon

dbr:Moon

....
Legendary astronaut

Buzz Aldrin has revealed some captivating pieces of Apollo 11 memorabilia on social media in the last few days.
....

dbr:Buzz_Aldrin

↑ semantic similarity ↓

dbr:Neil_Armstrong

Named Entity Linking

Entity-Based Query Matching

- simple entity matching
- **similarity-based entity matching**
- relationship-based entity matching
- ...

two entities are considered **semantically similar**,

- if they share property/value pairs
- if they share properties with similar values

Entity Based Search

Information Service Engineering - Lecture 6 Putting it All Together - 6.6 Semantic Search

Query Processing:

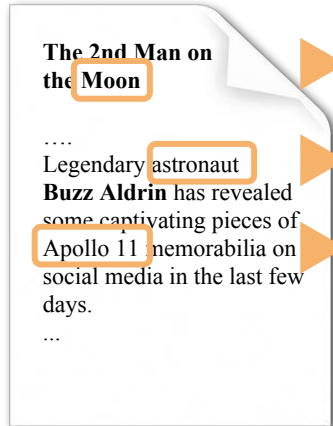
Armstrong on the Moon

Named Entity Linking

dbr:Neil_Armstrong

dbr:Moon

Indexing



dbr:Moon

dbo:Astronaut

dbr:Apollo_11

Entity-Based Query Matching

- simple entity matching
- similarity-based entity matching
- **relationship-based entity matching**
- ...

rdf:type

dbo:mission

dbr:Neil_Armstrong

Named Entity Linking



Exploratory Search

Picture References:

- P.4: By i ♥ happy!! from NY, NY (Flickr) [CC BY 2.0 (<http://creativecommons.org/licenses/by/2.0>)], via Wikimedia Commons from Wikimedia Commons
- P.19: St. Brendan's ship on the back of a whale, Scanned from a new book (2013) by Chet Van Duzer and published by the British Library: 'Sea Monsters on Medieval and Renaissance Maps', <https://www.flickr.com/photos/bibliodyssey/9606673293>