



Biotèa2Bioschemas

Enabling a knowledge graph from
scientific literature

Leyla Jael Garcia Castro (ljgarcia@ebi.ac.uk)

Knowledge and semantic web coordinator

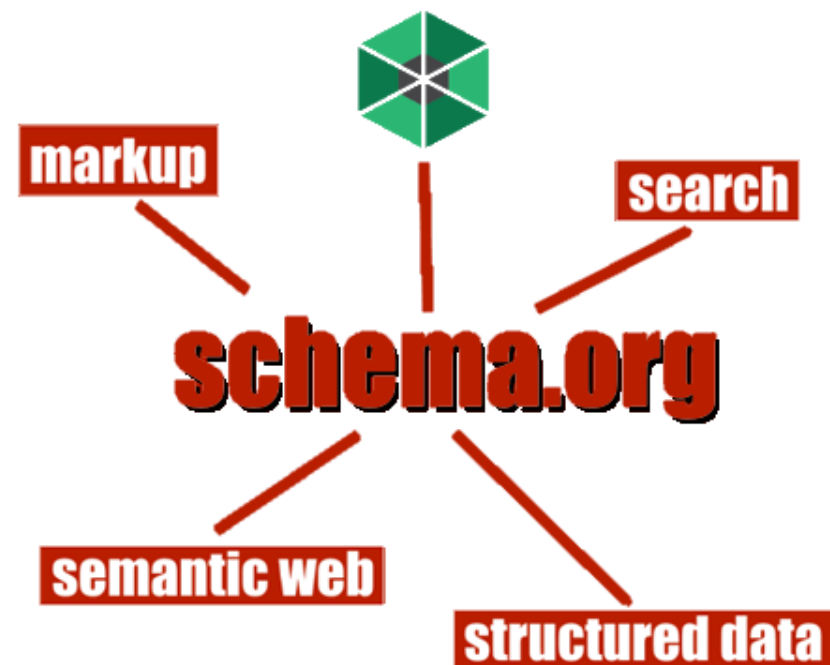
Elixir HUB

12-15 February 2019 - BLAH5, Kashiwa, Japan



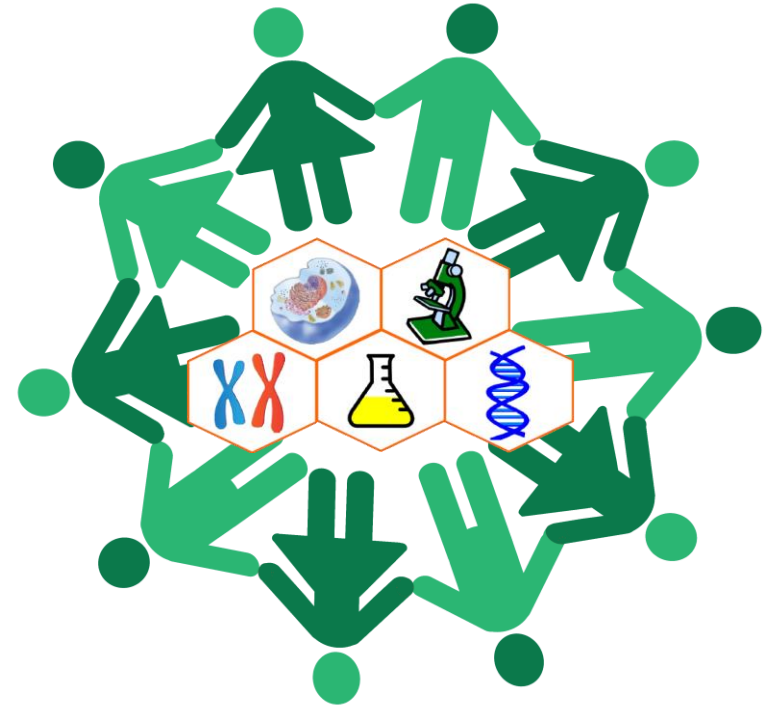
Background

- Collaborative, community activity
 - Schemas to add structure to web pages
 - Schema
 - Type, e.g., Book, Movie, Play and
 - Properties, .e.g., title, author, number of pages, duration, location
 - Search engines get better and better based on structured data retrievals



schema.org

- Community initiative built on top of **schema.org**
- Aim
 - Improve data discoverability and interoperability in Life Sciences
- How
 - Adding Life Science types to schema.org
 - Facilitating adoption and markup
 - Guidelines & examples → Profiles
 - Tools



Bioschemas

Bioschemas Profiles

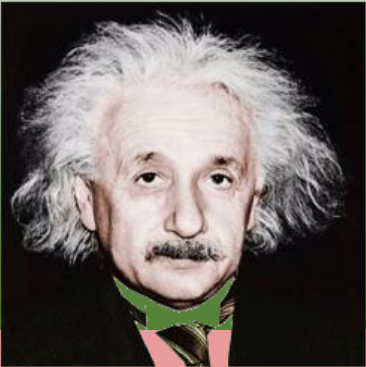
- Specific for Life Sciences
- Apply constraints to existing Schema.org types
- Minimum properties for finding and accessing data
- Best practices for selected properties
- Managed by Bioschemas

Birth date: 14
March 1879

Birth place: Ulm,
Kingdom of
Württemberg,
German Empire

email

Death date: 18
April 1955



Height: 1.72 cm

Alumni of: University
of Zurich

Weight: unknown

Work location: 1 Einstein Drive
Princeton, New Jersey
08540 USA

Know about: Physics,
relativity, mass energy
equivalence



Knows language:
German, English

Honorific prefix:
Dr. Prof.

Schema.org Types

- Generic data model
- Generous list of properties to describe data types
- Managed by Schema.org

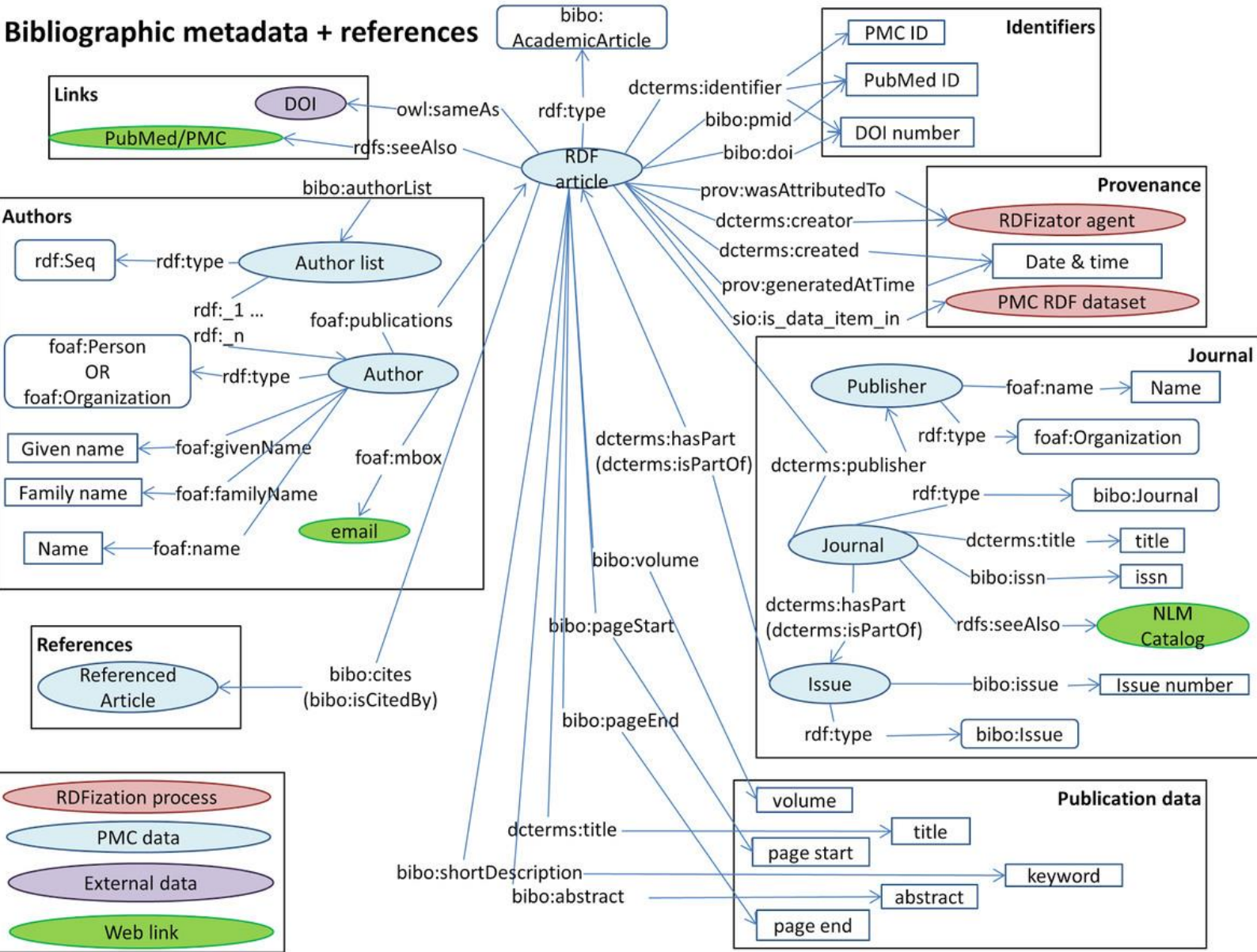
Award: Nobel
prize in physics

orcid

Siblings: Maria

isicv4: 7210

Bibliographic metadata + references



• Structured data model for scholarly publications

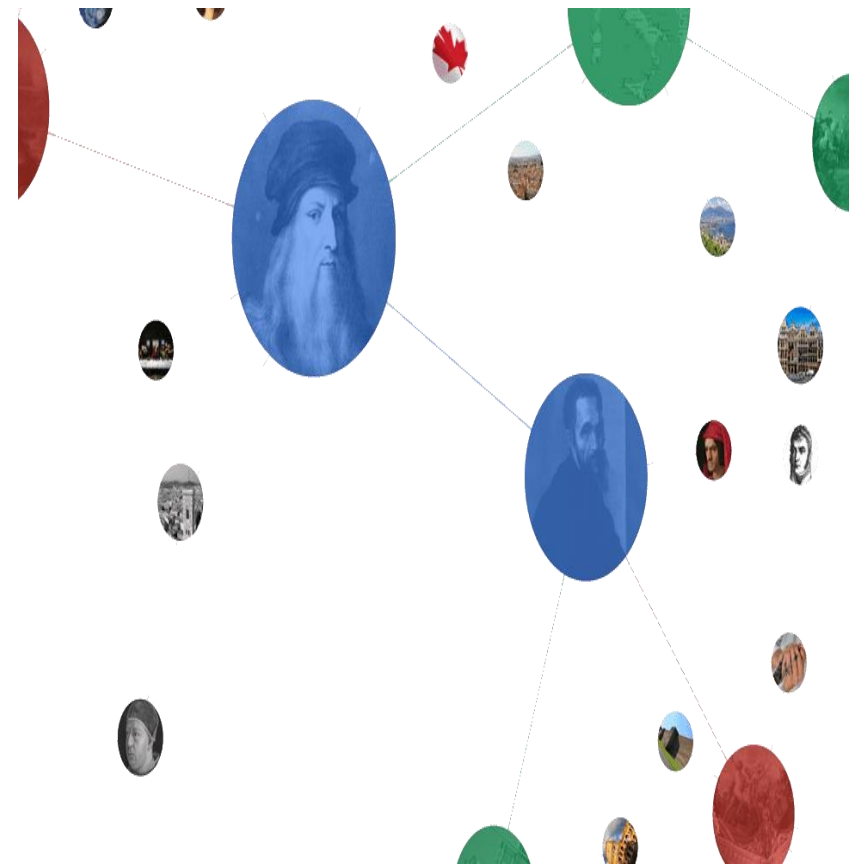
- Metadata
- References
- Content structure
- Full content
- Semantic annotations





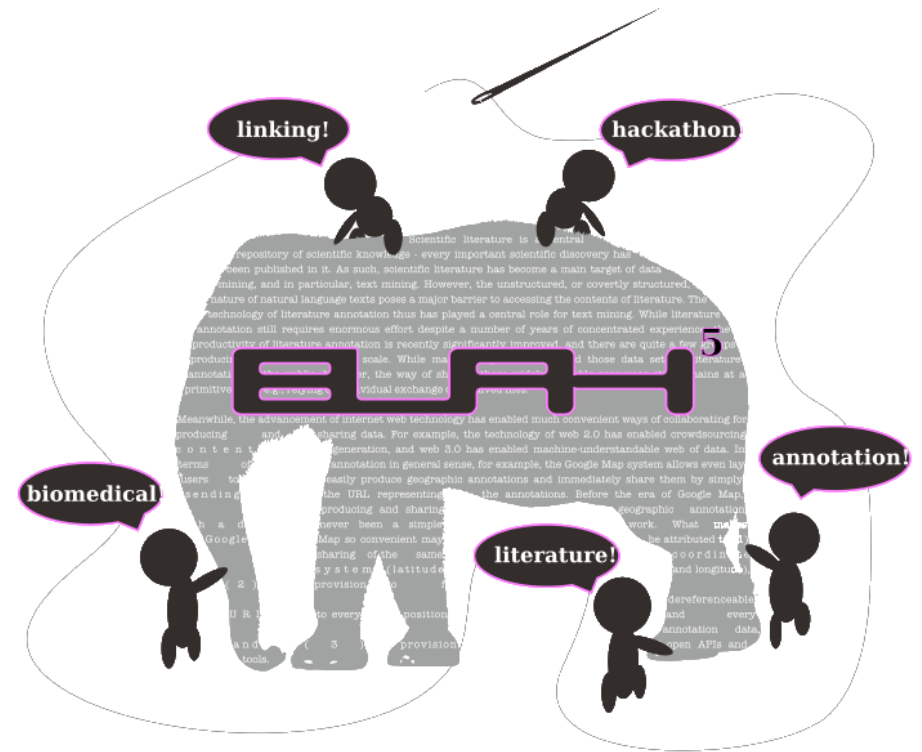
Biotea2Bioschemas

- Bioschemas profiles
 - Journal
 - PublicationVolume
 - PublicationIssue
 - ScholarlyArticle
 - SemanticScholarlyAnnotation
- Data → markup
- Knowledge graph
- Analyses and tools



Literature based knowledge graph

- Data → markup
 - On demand application + sitemap → crawlable
 - Batch application
- First day
 - Try out both options
 - Decide on one
- Second and third day
 - Hacking on the selected option



@BLAH5

*Thank
you*



BLAH5

 <http://bioschemas.org/>

 @bioschemas

 <https://github.com/bioschemas/>