

**Building a Linked Open Data
Knowledge Graph**

Michele Pasin

Library Fair Forum 2017
November 2017

Springer Nature's Metadata Mission Statement



We understand metadata as the gateway to our content.

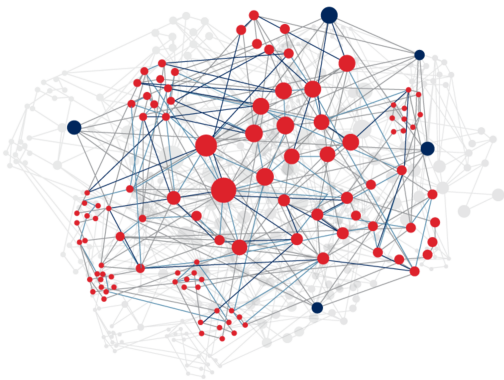
We provide best quality metadata with state-of-the-art enrichment in all key formats and flavors, available in all relevant delivery models.

Our bibliographic metadata is free, open and reusable.

Springer Nature SciGraph

A Linked Open Data platform for the scholarly domain

SN SciGraph



- > Collaborative effort between Springer Nature and Digital Science
- > Supporting internal use cases, but also contributing to an emerging web of **linked open science data**
- > Not just publications data but a **wealth** of further related information

Linked Open Data Publishing So Far

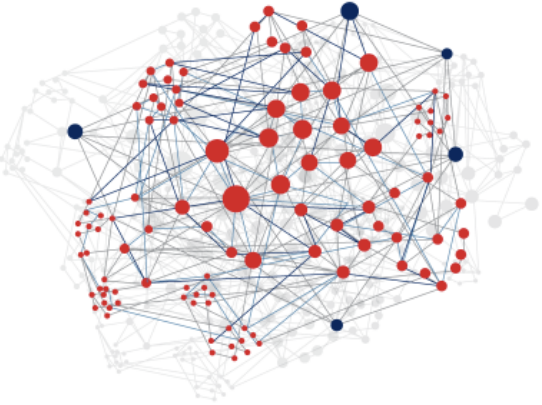
SPRINGER NATURE

Springer Nature SciGraph

A Linked Open Data platform for the scholarly domain

We are pleased to introduce Springer Nature SciGraph, the new Linked Open Data platform aggregating data sources from Springer Nature and key partners from the scholarly domain. The Linked Open Data platform will initially collate information from across the research landscape, such as funders, research projects, conferences, affiliations and publications. Additional data, such as citations, patents, clinical trials and usage numbers will follow over time. This high quality data from trusted and reliable sources provides a rich semantic description of how information is related, as well as enabling innovative visualizations of the scholarly domain.

By doing so, Springer Nature SciGraph overcomes former boundaries by relating comprehensive information about the research landscape. It represents a further step in data integration and it will continue to grow organically. This platform will increase the discoverability of high quality data as larger parts of our datasets will be made freely available under a CC BY-NC 4.0 license.



The data in Springer Nature SciGraph is projected to contain 1.5 to 2 billion triples. It will comprise metadata from journals and articles, books and chapters, organizations, institutions, funders, research grants, patents, clinical trials, substances, conference series, events, citations and reference networks, Altmetrics, links to research datasets and much more.

Any questions?
Please contact us.

Dataset Download

Licensing Information

Further Info

Conference Presentation 2016 (PDF, 11.56 MB)

At a glance:

- 150 M triples / 32G downloads
- CC-BY-NC license

Metadata about:

- Articles 2012-2016 (5M) + Abstracts
- Grants (200k)
- Journals (3k)
- Subjects (3k)
- Core Ontology

www.springernature.com/scigraph

Open Data Events: Hack Day June 2017

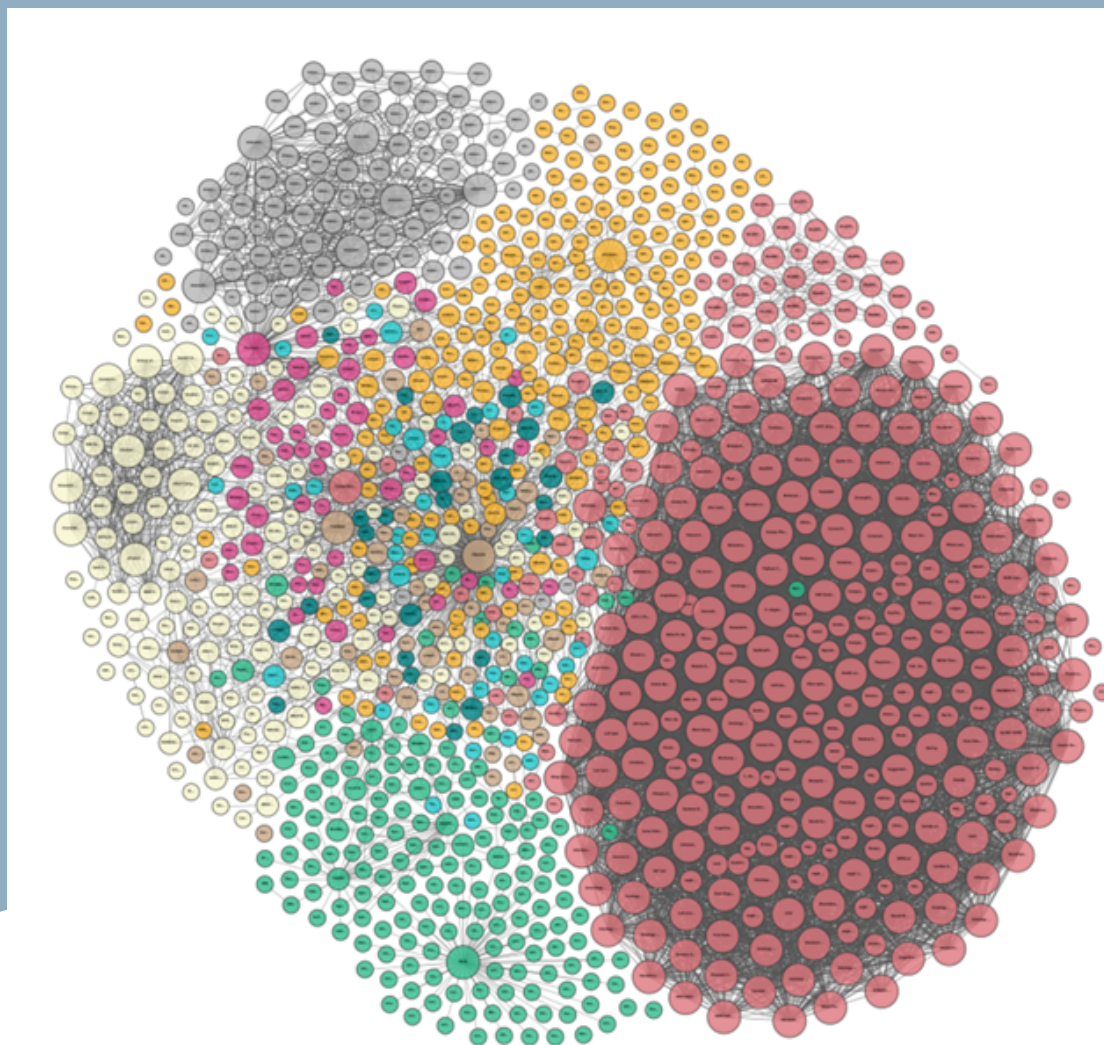
Aims and Scope

- Engagement with Linked Data Researcher Community
- Encourage developers to build tools with our data
- Position ourselves as Open Data research publisher
- Gather first-hand feedback from users of our data



Web of Data

- Be part of the LOD cloud!
- The future library and trade catalog is the Internet!



Linking Open Data cloud diagram 2017, by Andrej's Abele, John P. McCrae, Paul Buitelaar, Anja Jentzsch and Richard Cyganiak. <http://lod-cloud.net> (CC-BY-SA)

Libraries using Linked Data

The Library of Congress > Linked Data Service

LIBRARY OF CONGRESS LINKED DATA SERVICE

LC Linked Data Service
Authorities and Vocabularies

Search

Enter Keyword or Phrase

All
 LC Subject Headings
 LC Name Authority File
 LC Classification
 LC Children's Subject Headings

Available Datasets

The Linked Data Service provides access to commonly found standards and vocabularies promulgated by the Library of Congress. This includes data values and the controlled vocabularies that house them. The following are currently available:

- > [LC Subject Headings](#)
- > [LC Name Authority File](#)
- > [LC Classification](#)
- > [LC Children's Subject Headings](#)
- > [LC Genre/Form Terms](#)
- > [LC Medium of Performance Thesaurus for Music](#)
- > [MARC Relations](#)
- > [MARC Countries](#)
- > [MARC Geographic Areas](#)
- > [MARC Languages](#)
- > [MARC Genre Terms](#)
- > [ISO639-1 Languages](#)
- > [ISO639-2 Languages](#)
- > [Schemes](#)
- > [Identifiers](#)
- > [Carriers](#)
- > [Content Types](#)
- > [Media Types](#)
- > [Resource Types](#)
- > [Description Conventions](#)

Linked Data Service

- [About](#)
- [Main Dataset Descriptions](#)
- [Preservation Dataset Descriptions](#)
- [Small Datasets Descriptions](#)
- [Search](#)
- [Download](#)
- [Technical Center](#)
- [Contact Us](#)
- [Privacy Policy](#)

Library of Congress Linked Data Service (2009)

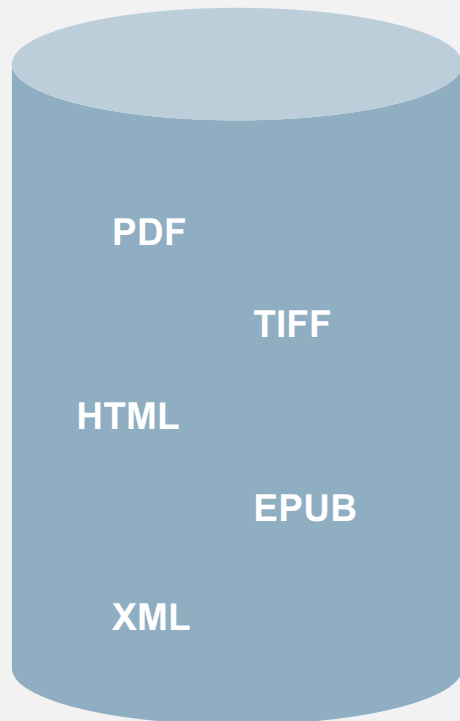
- A library catalog “must be designed by considering its context of the Web”
- Access to data at no cost.
- Ability to link to Library of Congress data values within your metadata via Linked Data.

Other libraries:

- British Library (**BL**)
- French National Library (**BNF**)
- German National Library (**DNB**)
- National Library of Spain (**BNE**)
- National Library of Sweden (**LIBRIS**)
- Hungarian National Library (**NSL**)

Vision: From Content to Data

Content base



We publish content



Knowledge Graph



We manage knowledge

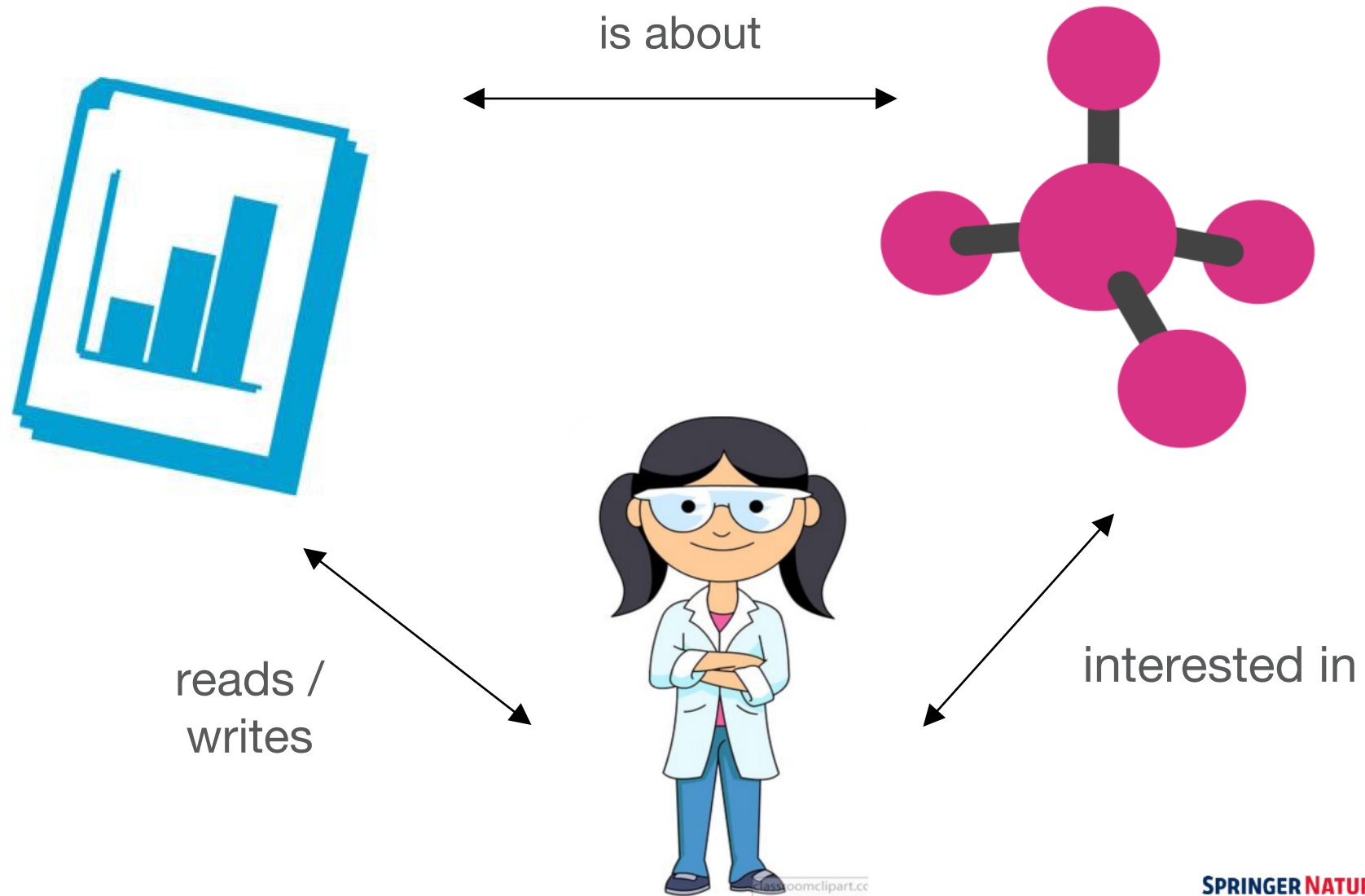
Vision: From Content to Data

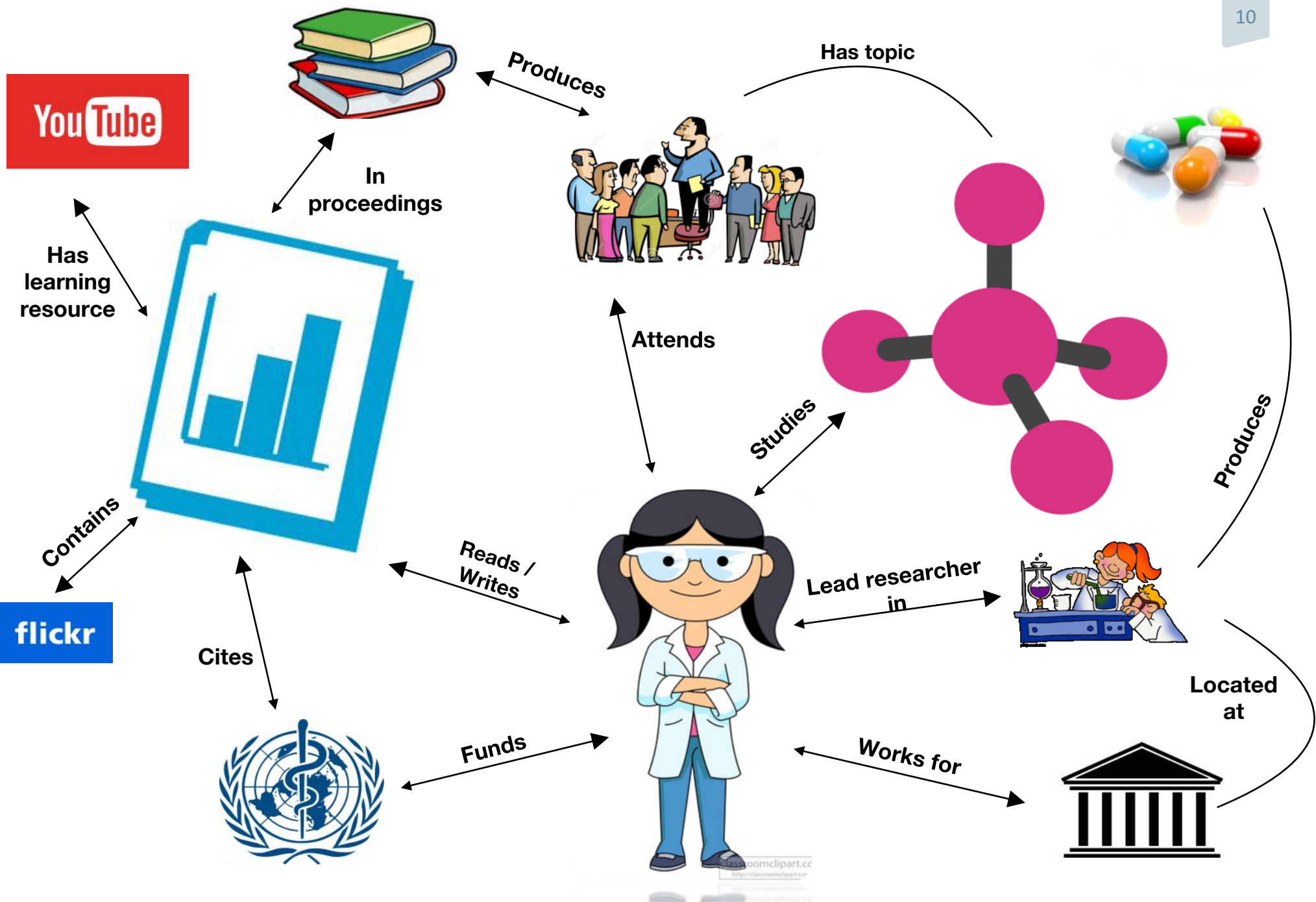
- We create the largest state-of-the-art linked open data aggregation platform for the scholarly domain.
- In doing so, we increase content discoverability and provide data tools and services for researchers, authors, editors, librarians, data scientists, funders, conference organizers, and many others by adding value across all content types.

We publish content

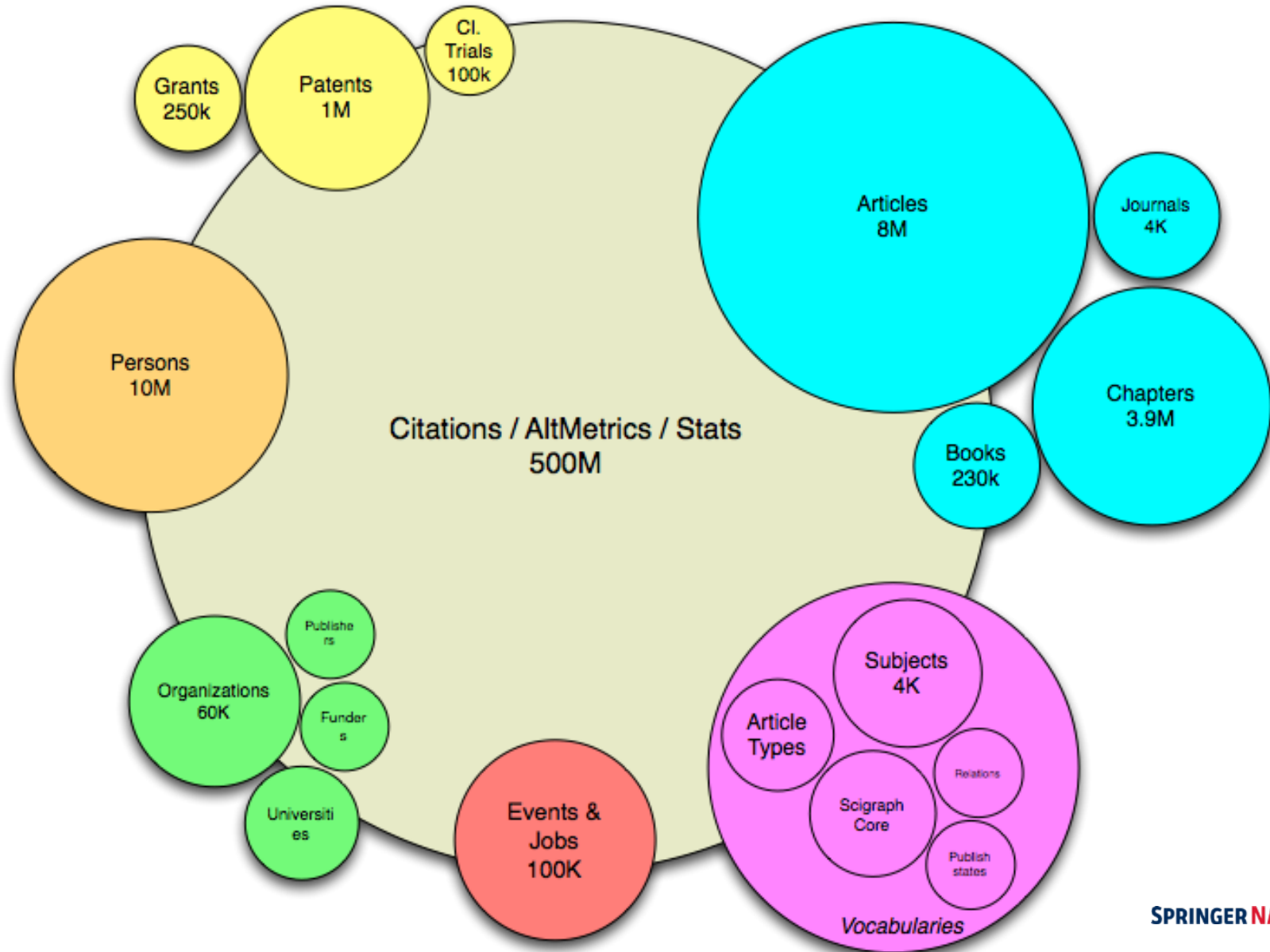
We manage knowledge

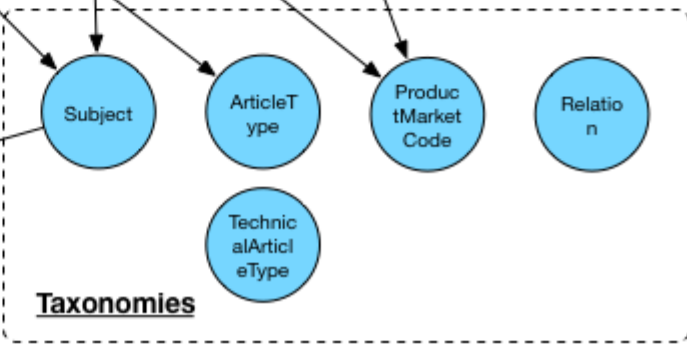
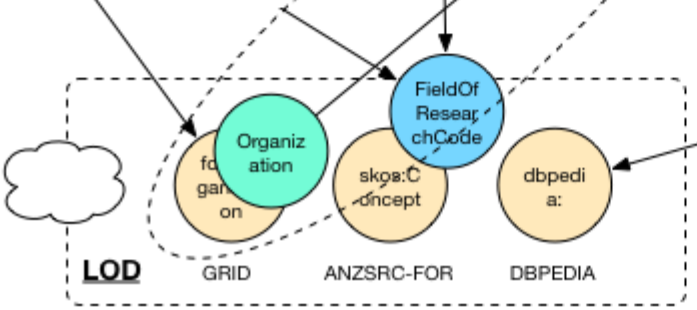
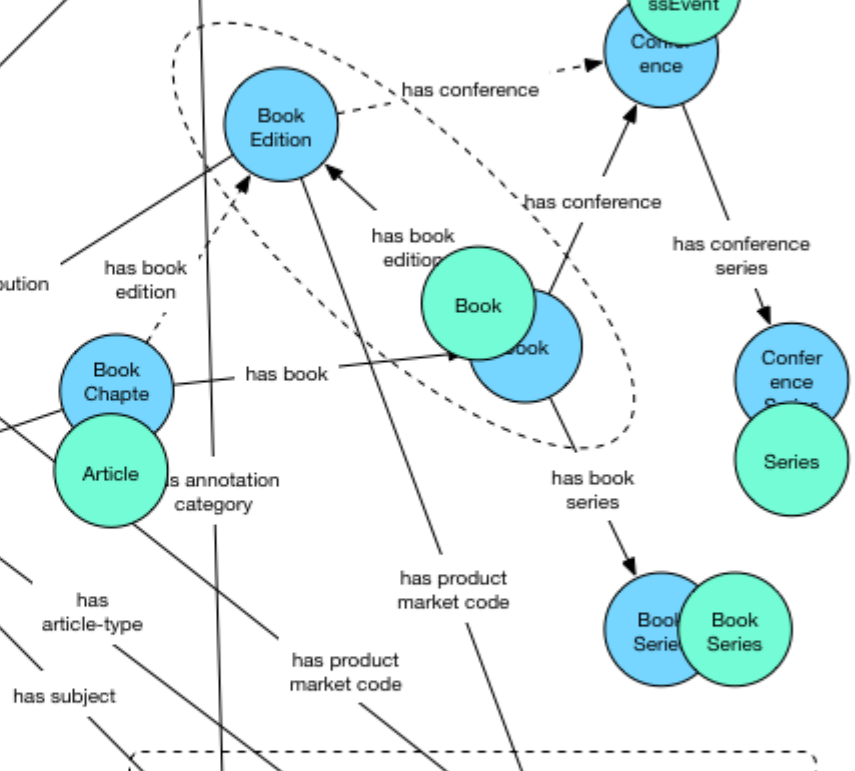
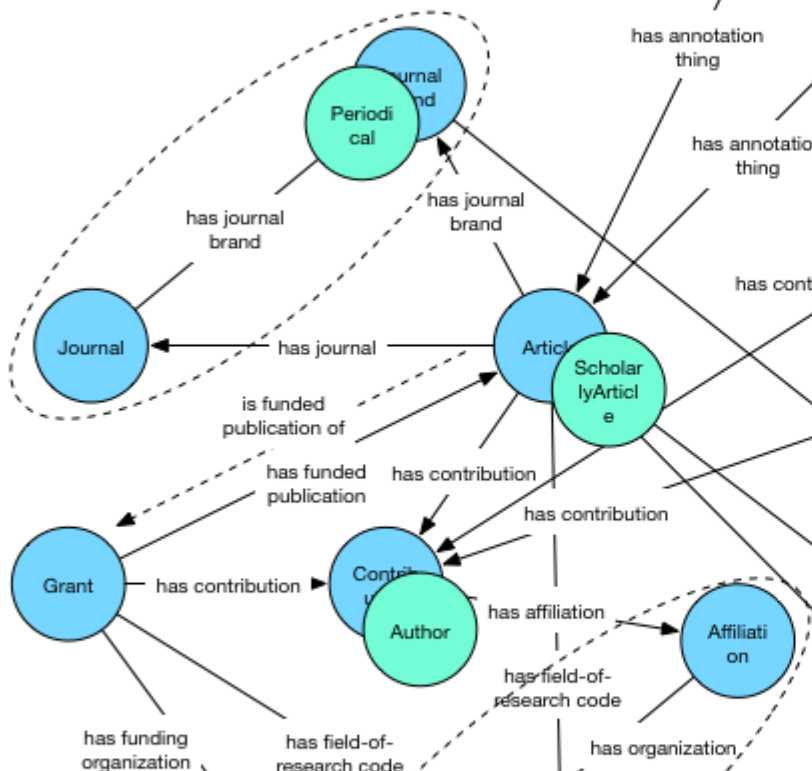
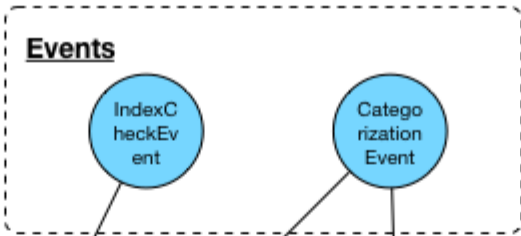
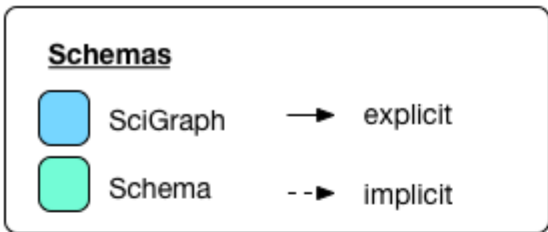
Three areas of knowledge we care about





Springer Nature SciGraph Data Landscape





Springer Nature SciGraph Applications

Springer Nature Scigraph: Applications



Linked Open Data Publishing

- **Researchers** can analyze/build upon our data
- Contributing to Open Research

Business Intelligence and Analytics

- Dashboards for understanding the research landscape
- **Editors, Sales, Marketing** etc.

Content Discoverability

- SN SciGraph Data Explorer
- APIs for better **end user** applications

Applications

Analytics Dashboards

Springer Nature SciGraph Analytics Dashboards

Springer Nature SciGraph Analytics Dashboards Journals Institutions Countries Subject Areas

BMC Cell Biology

Journal ID: 12860

Note: In order to obtain the raw data for this dashboard please contact the [Knowledge Graph team](#)

PUBLICATION VOLUME JOURNAL METRICS AUTHORS COUNTRIES & INSTITUTIONS FIELD OF RESEARCH RESEARCH FUNDING DATA QUALITY

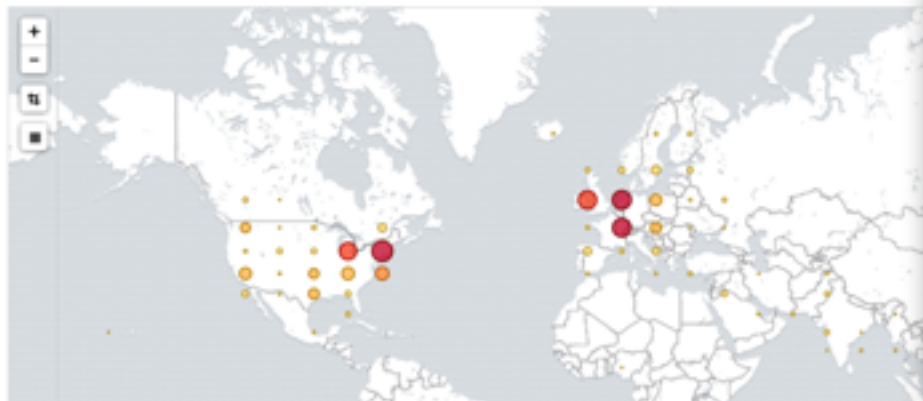
Section - Countries and Institutions

Countries and Institutions

Use this section to find out which are the top countries and institutions contributing to a publication.

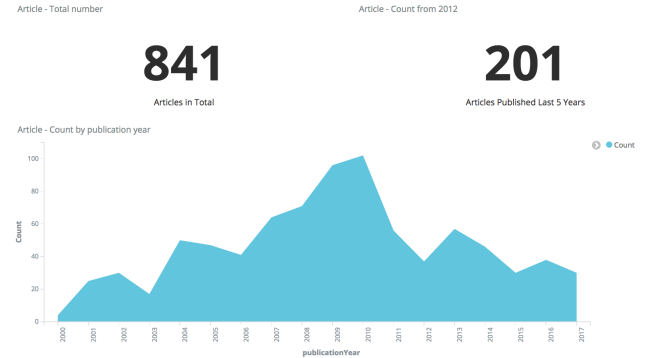
Note: this information comes from the GRID database (<https://www.grid.ac/>).

Article - map view



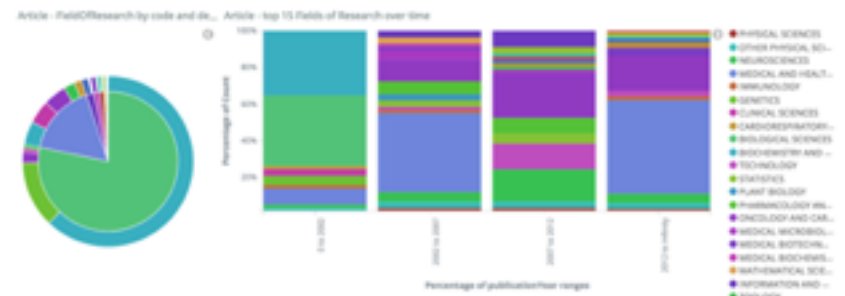
Publication Volume

This section provides statistics useful to understand the type and volume of content linked to a publication. For example, how many articles have been published over the years, which are the most frequently used article types and how much of this content has been indexed in external databases.

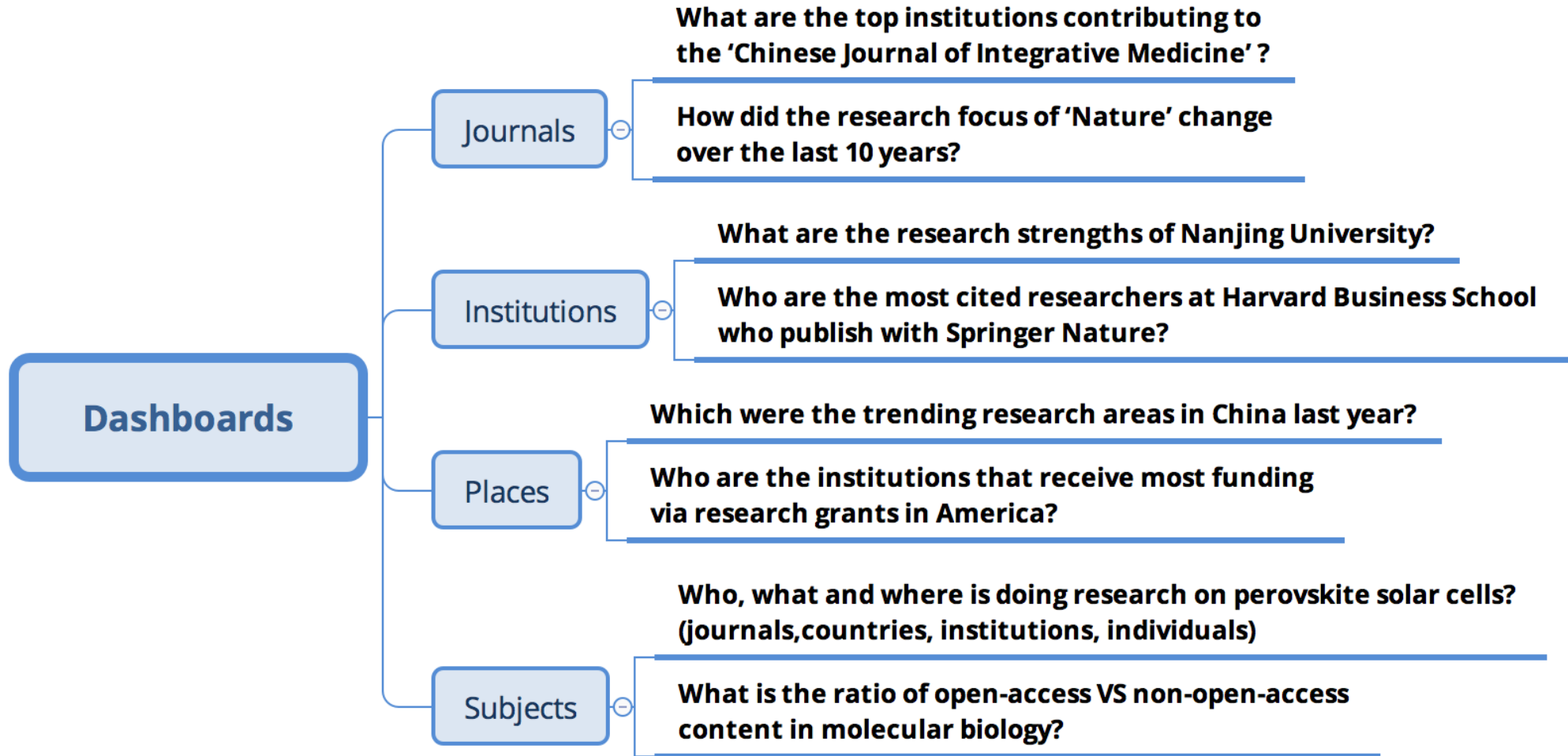


Fields of Research

This section provides a breakdown of publication content based on subject areas. The subject areas are derived from the Australian and New Zealand Standard Research Classification (ANZSRC) <http://www.anzsrc.gov.au/australian%20and%20new%20zealand%20standard%20research%20classification>



Springer Nature SciGraph Analytics: Supporting Data Driven Decisions



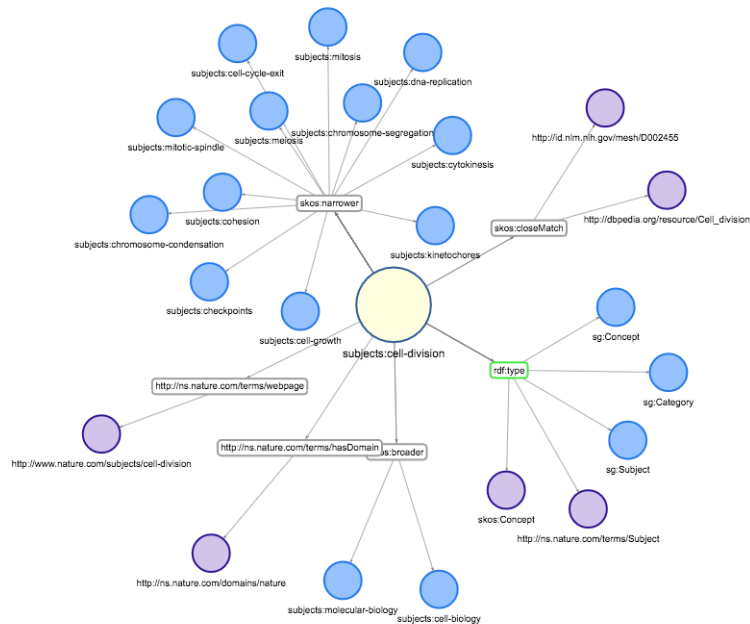
Applications

Linked Data Explorer

Linked Data Explorer

Cell division

Cell division is the process by which a parental cell gives rise to two daughter cells. The process involves both nuclear division and cytokinesis and can either produce two equal cells (symmetric cell division) or two cells with different cellular fates (asymmetric cell division).



Purpose

- simple UI for exploring graph contents interactively
- internal version: unrestricted access to data and statistics
- external version: only public data, allows linked data 'dereferencing'

Features

- text & graphical UI on top of graph database
- one page per URI, permits to download machine readable descriptions of data

Summary, Next Steps and Public Outreach

Looking Ahead

- **Summary**

- Springer Nature Scigraph is our LOD platform: Focus on data re-use, integration and discoverability
- Collaboration between Springer Nature and Digital Science (and other partners)
- Data publishing: ~150M triples released, 1B+ in our graph
- Internal use cases: Ontology management, analytics dashboards, semantic publishing

Looking Ahead

- **Next Steps**

- Data publishing: New release including complete archive, updated CC license
- Visualization, interactive exploration of the graph (Linked Data Browser), analytics dashboards in particular to support *Sales, Marketing and Editorial*
- Exploring metadata deliveries to third parties (Google, libraries)
- SN Taxonomy Manager: Beta going live this month
- Going Pan-publisher

- **Public Outreach**

- Collaboration with DBpedia: Internship in London/Leipzig
- Hack Day: Research data Publishing hack day (London, November 17, 2017)

More Discovery Tools from Springer Nature

<http://recommended.springernature.com>

recommended@springernature.com

Recommended

Keeping up-to-date with published research just got easier

With 4000* new papers published every day, Recommended helps you find the right one to read next.

Start receiving recommendations

*Source: The STM Report, March 2015

SPRINGER NATURE

Recommended is a personalised service that suggests relevant papers to you, based on what you've previously read, from all publishers.

How we deliver recommendations

The popup:

Only shown on article pages across Nature.com, BioMed Central and SpringerOpen.

Only appears after scrolling the page.

Only shown *if* we have at least one recommendation to show.

Up to 5 recommendations shown, with an email sign up onward journey.

E.g. <https://www.nature.com/ncomms/>

The screenshot displays a web page from Nature.com. The main article is titled "Genomic prediction contributing to a promising global strategy to turbocharge gene banks". A "Recommended" popup is overlaid on the page, featuring a blue header with a yellow 'R' icon and a close button. The popup content includes the title "Coseismic and early postseismic deformation due to the 25 April 2015, Mw7.8 Gorkha, Nepal, earthquake from InS...", the authors "Sreejith, K. M... Rajawat, A. S.", and the journal "Geophysical Research Letters (2016)". Below the title is a "TRY NOW FOR FREE" button and a "Tell me more" link. The popup also shows a "1 of 6" indicator and a "Next" button. The background article text is partially visible, showing the abstract and author information.

How we deliver recommendations

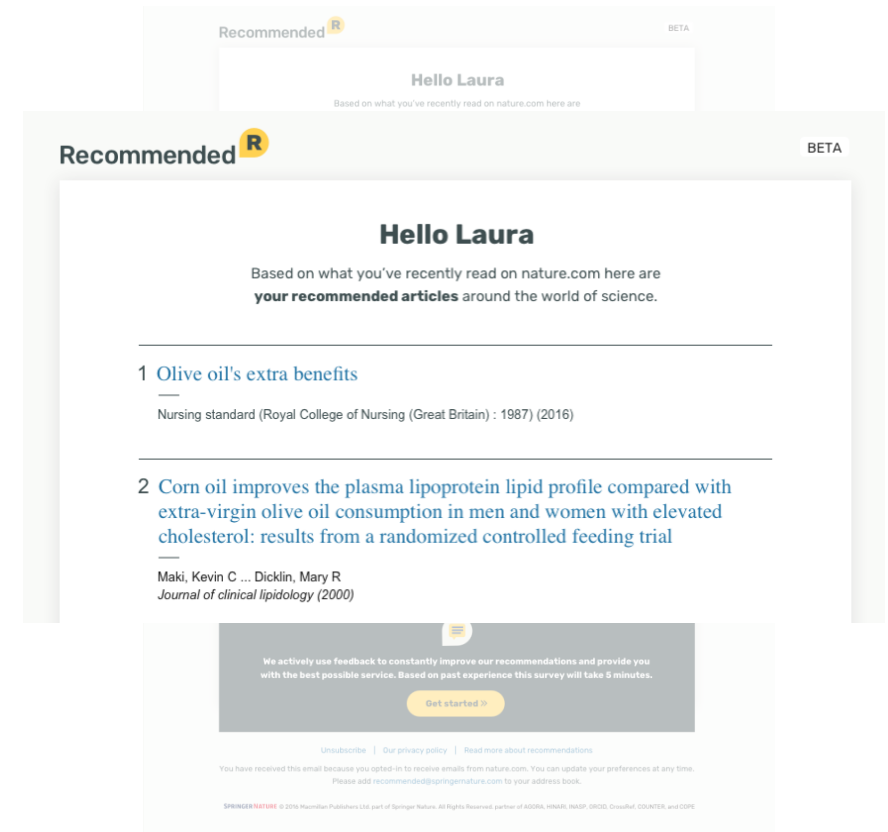
Email:

Only sent to people who have signed up from the popup or home page.

Up to 10 recommendations in an email.

Sent weekly if we have *new* recommendations to show.

Users can change the frequency through a link in the email.



<http://www.springernature.com/sharedit>



SN SharedIt

SharedIt allows Springer Nature authors and subscribers to share groundbreaking research for free on a global level

ReadCube technology makes sharing Springer Nature SharedIt article links quick and easy

Select from 2,300 journals from world leaders in research, including Nature Research, Springer, BioMed Central, and Palgrave Macmillan

SPRINGER NATURE

sharedit@springernature.com

① Users with access rights share with colleagues and collaborators

With the content sharing function ...

The screenshot shows the Nature journal website. The article title is "The Sagittarius impact as an architect of spirality and outer rings in the Milky Way". A popup window displays a "Shareable Link" with the URL <http://rdcu.be/bHrz>. Below the link are social media sharing options: CiteULike, Facebook, Twitter, Delicious, Digg, Google+, LinkedIn, and StumbleUpon. The article text below the popup reads: "Like many galaxies of its size, the Milky Way is a disk with a central bar¹, although our knowledge of its structure and origin is incomplete. Traditional attempts to understand our Galaxy's morphology assume that it has been unperturbed by major external forces. Here we report simulations of the response of the Milky Way to the infall of the Sagittarius² dwarf galaxy (Sgr), which results in the formation of spiral arms, influences the central bar and produces a flared outer disk. Two ring-like wrappings emerge towards the Galactic anti-Centre in our model that are reminiscent of the low-

The screenshot shows a "Compose new Tweet" window. The text in the tweet is "Minor collisions affect galaxy morphology" followed by the shareable link <http://rdcu.be/bG6K>. Below the text are options for "Add photo", "Location disabled", and a "Tweet" button with a retweet count of 76.



The screenshot shows the full article page on the Nature website. The article title is "The Sagittarius impact as an architect of spirality and outer rings in the Milky Way". The authors listed are Chris W. Purcell¹, James S. Bullock², Erik J. Tollerud², Miguel Rocha³ & Sukanya Chakrabarti¹. The article text begins: "Like many galaxies of its size, the Milky Way is a disk with prominent spiral arms nested in a central bar¹, although our knowledge of its structure and origin is incomplete. Traditional attempts to understand our Galaxy's morphology assume that it has been unperturbed by major external forces. Here we report simulations of the response of the Milky Way to the infall of the Sagittarius² dwarf galaxy (Sgr), which results in the formation of spiral arms, influences the central bar and produces a flared outer disk. Two ring-like wrappings emerge towards the Galactic anti-Centre in our model that are reminiscent of the low-latitude arcs observed in the same area of the Milky Way. Previous models have focused on Sgr itself^{3,4} to reproduce the dwarf's orbital history and place associated constraints on the shape of the Milky Way gravitational potential, treating the Sgr impact event as a trivial influence on the Galactic disk. Our results show that the Milky Way's morphology is not purely secular in origin and that low-mass minor mergers predicted to be common throughout the Universe⁵ probably have a similarly important role in shaping galactic structure. To discover the specific effects of the Sgr impact on the Galactic disk, we need to simulate directly the dark matter and stellar components in both the Milky Way and the Sgr progenitor and to assess that Sgr has a realistic dark-to-baryonic mass ratio, given the Λ CDM (where Λ represents the accelerating expansion of our Universe, which has a matter weaker ancillary arm with a pitch angle different from that of the primary mode and begins to liberate stellar material from Sgr. These repeated polar encounters produce flaring, asymmetric, clashing in the disk plane, and vertical oscillations above and below the plane of the ascending spiral strips. The evolution of the central bar can also be affected by perturbing impacts. Although bar formation is sensitive to initial conditions, it is interesting to compare results from run to run, which rely on identical

Just by attaching share URLs to email or social media, anyone can access

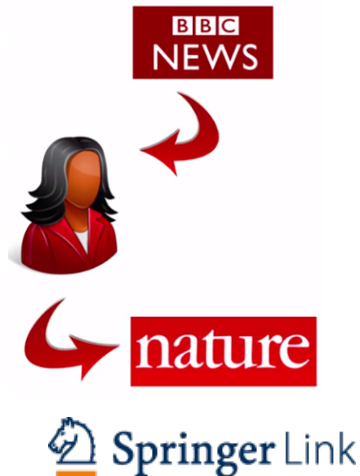
① Users with access rights share with colleagues and collaborators

If you do not have access to the paper, you can only browse the article (Printing, PDF saving is not possible)

The screenshot shows a web browser displaying a Nature article. The browser tabs include 'Edward Xample (@EXample)' and 'The Sagittarius impact as...'. The URL is www.nature.com/articles/nature10417.epdf?shared_access_token=y042c9QbHnvkkjg5_J5auNRgN0JAjWel9Jnr3ZoTv0PWID4iLFHnF2oFIR8NNuxte9D8b3AGbky_Zhn6FQ6aiuqAR5olJA. The article title is 'The Sagittarius impact as an architect of spirality and outer rings in the Milky Way' by Chris W. Purcell^{1,2}, James S. Bullock², Erik J. Tollerud², Miguel Rocha² & Sukanya Chakrabarti¹. The article is categorized as a 'LETTER' with a DOI of 10.1038/nature10417. The abstract states: 'Like many galaxies of its size, the Milky Way is a disk with prominent spiral arms rooted in a central bar', although our knowledge of its structure and origin is incomplete. Traditional attempts to understand our Galaxy's morphology assume that it has been unperturbed by major external forces. Here we report simulations of the response of the Milky Way to the infall of the Sagittarius dwarf galaxy (Sgr), which results in the formation of spiral arms, influences the central bar and produces a flared outer disk. Two ring-like wrappings emerge towards the Galactic anti-Centre in our model that are reminiscent of the low-latitude arcs observed in the same area of the Milky Way. Previous models have focused on Sgr itself⁴ to reproduce the dwarf's orbital history and place associated constraints on the shape of the Milky Way gravitational potential, treating the Sgr impact event as a trivial influence on the Galactic disk. Our results show that the Milky Way's morphology is not purely secular in origin and that low-mass minor mergers predicted to be common throughout the Universe⁵ probably have a similarly important role in shaping galactic structure. To discern the specific effect of the Sgr impact on the Galactic disk, we need to simulate directly the dark matter and stellar components in both the Milky Way and the Sgr progenitor and to ensure that Sgr has a realistic dark-to-baryonic mass ratio, given the Λ CDM (where Λ represents the accelerating expansion of our Universe, which has a matter weaker ancillary arm with a pitch angle different from that of the primary mode and begins to liberate stellar material from Sgr. These repeated polar encounters produce flaring, asymmetric sloshing in the disk plane, and vertical oscillations above and below the plane of the forming spiral wraps. The evolution of the central bar can also be affected by perturbing impacts. Although bar formation is sensitive to initial conditions, it is interesting to compare results from run to run, which rely on identical

The article includes several figures: Figure 1 shows evolved disk end states; Figure 2 shows face-on surface density visualizations of the Milky Way at four important moments; Figure 3 shows the observed Sgr tidal debris stream and remnant core; Figure 4 shows endstate disk overdensities. A sidebar on the left lists these figures with small thumbnail images. The bottom of the page features a navigation bar with options for 'Download PDF', 'Supplements', and 'Related Articles', along with a search icon and a user profile for 'Phill Jones'.

② Access to articles from about 200 news organizations in the world and science blog sites



Providing access to full text articles to readers of about 100 news organizations and science blog sites around the world aiming to convey deeper information on important research widely and in general

The screenshot shows a web browser displaying an article on the Scientific American website. The URL is www.scientificamerican.com/article/sagittarius-satellite-spiral/. The article title is "Star-Crossed: Milky Way's Sagittarius Dwarf Galaxy". The author is Erik Tollerud. The article text discusses the Sagittarius Dwarf Galaxy and its interaction with the Milky Way. A red box highlights a sentence in the text: "The effects, as it turned out, were strong. In the simulations, described in a study published in the September 15 issue of *Nature*, Sagittarius stirred up enough ripples to make a smooth, circular, spinning galactic disk evolve into a spiral much like the Milky Way." The article includes a colorful image of the galaxy and a sidebar with "Latest News" and "Most Read" sections. There is also a large advertisement for XARELTO on the right side of the page.

Thank you

How to get in touch:

- E-mail
michele.pasin@springernature.com
- Portal
<http://www.springernature.com/scigraph>
- E-Mail
scigraph@springernature.com
- Twitter
[#scigraph](https://twitter.com/scigraph)