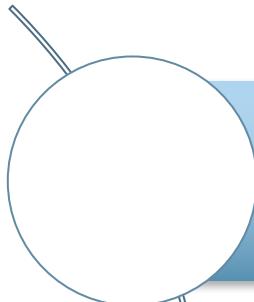


Building a Linked Open Data Knowledge Graph

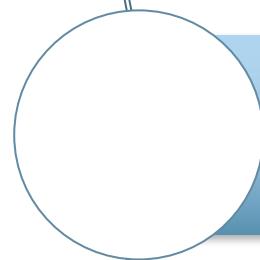
Michele Pasin

Library Fair Forum 2017
November 2017

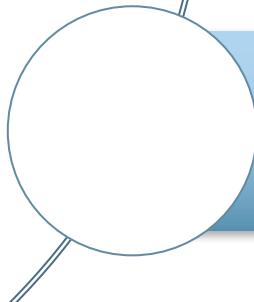
Springer Nature's Metadata Mission Statement



We understand metadata as the gateway to our content.



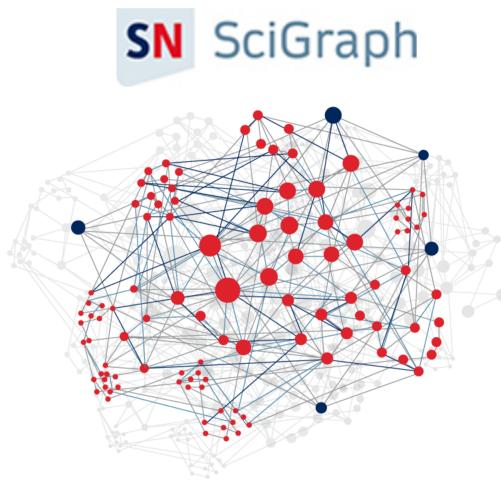
We provide best quality metadata with state-of-the-art enrichment in all key formats and flavors, available in all relevant delivery models.



Our bibliographic metadata is free, open and reusable.

Springer Nature SciGraph

A Linked Open Data platform for the scholarly domain



- > 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

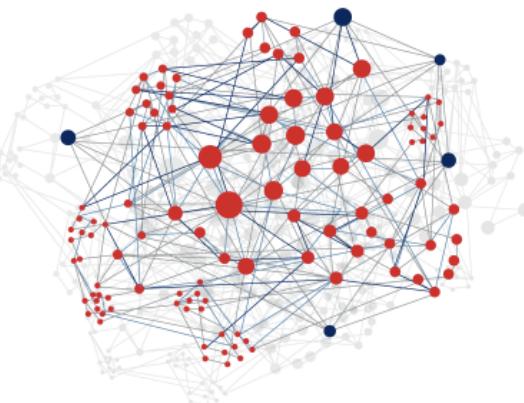
SPRINGER NATURE

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

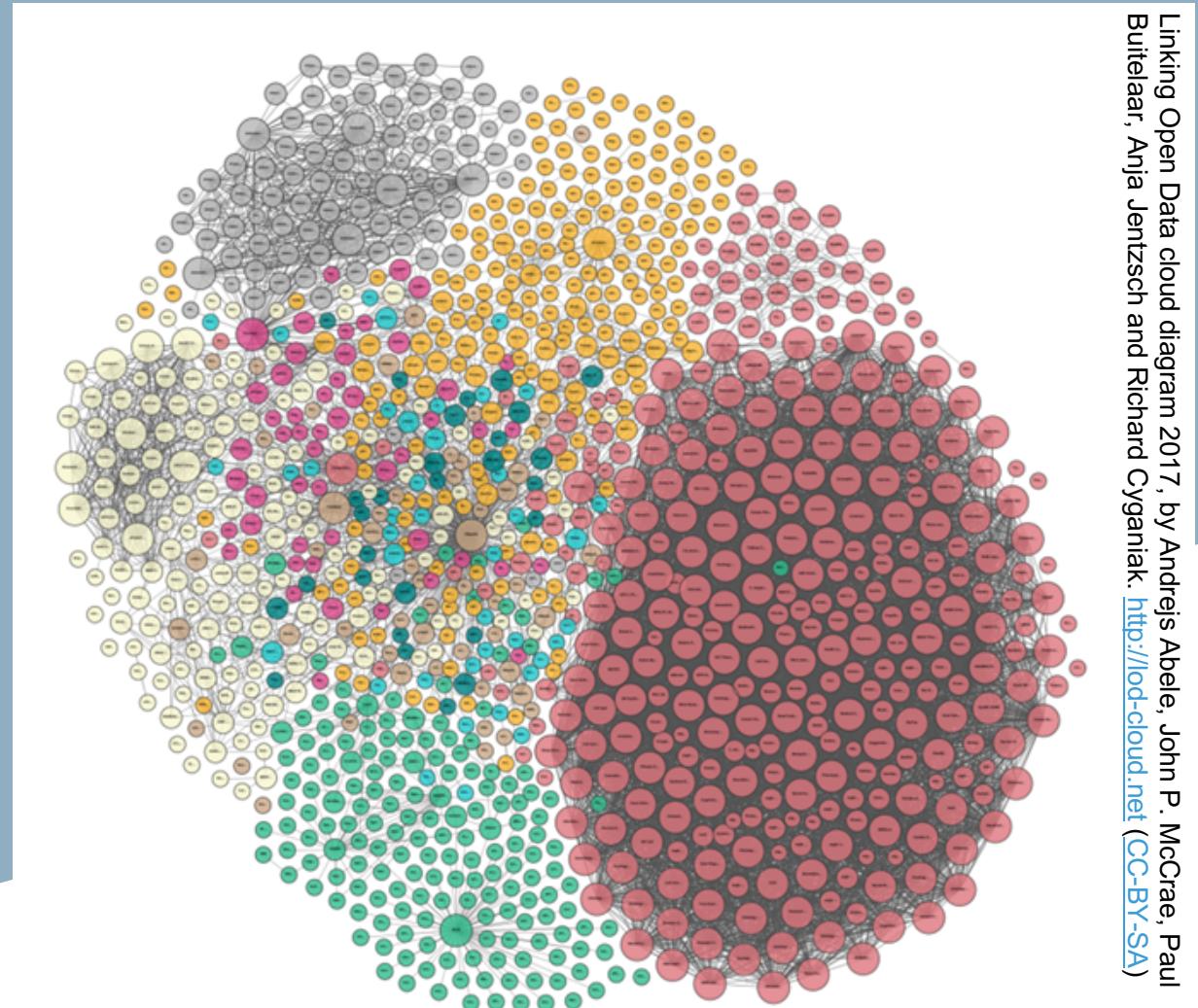
Aims and Scope

- 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 Andrejs Abele, John P. McCrae, Paul Buitelaar, Anja Jentzsch and Richard Cyganiak. <http://lod-cloud.net> (CC-BY-SA)

Libraries using Linked Data

LIBRARY OF CONGRESS

ASK A LIBRARIAN DIGITAL COLLECTIONS LIBRARY CATALOGS

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

Search

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	> MARC Relators
> LC Name Authority File	> MARC Countries
> LC Classification	> MARC Geographic Areas
> LC Children's Subject Headings	> MARC Languages
> LC Genre/Form Terms	> MARC Genre Terms
> LC Medium of Performance Thesaurus for Music	> ISO639-1 Languages
	> ISO639-2 Languages

Schemes

- > [Identifiers](#)
- > [Carriers](#)
- > [Content Types](#)
- > [Media Types](#)
- > [Resource Types](#)
- > [Description Conventions](#)

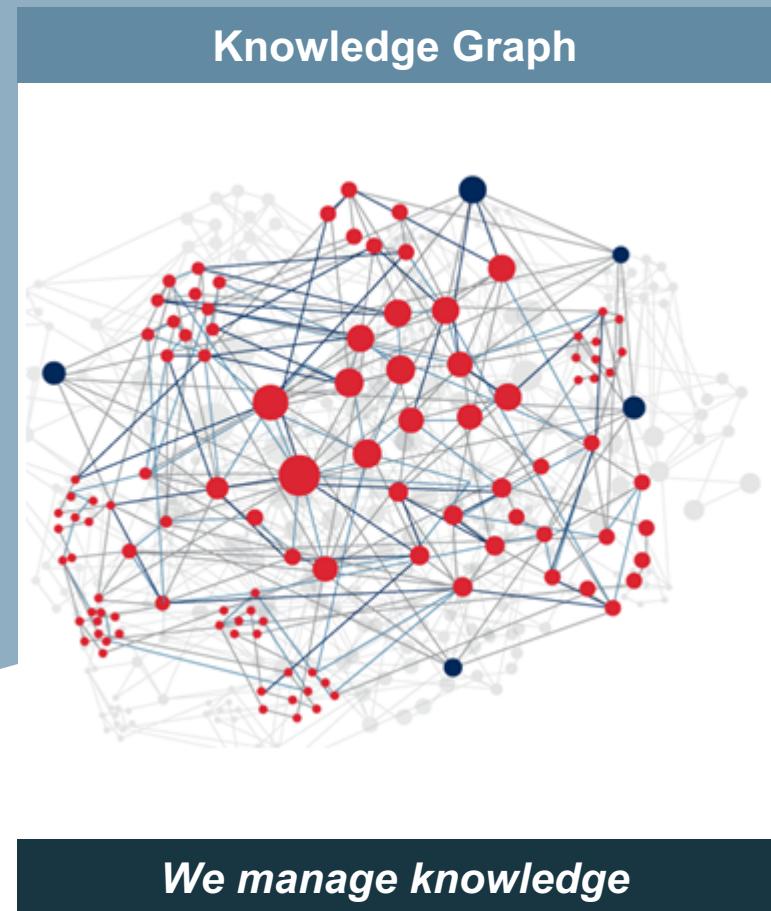
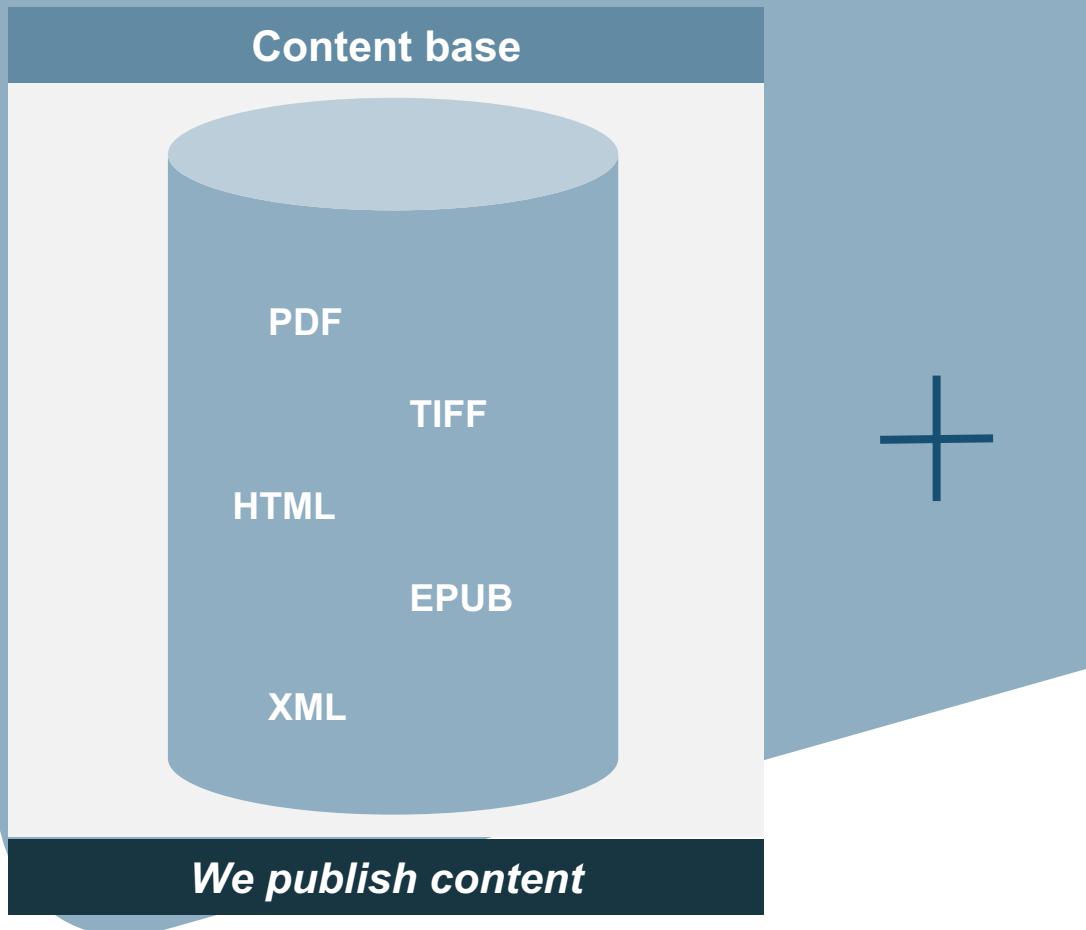
Library of Congress Linked Data Service (2009)

- 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



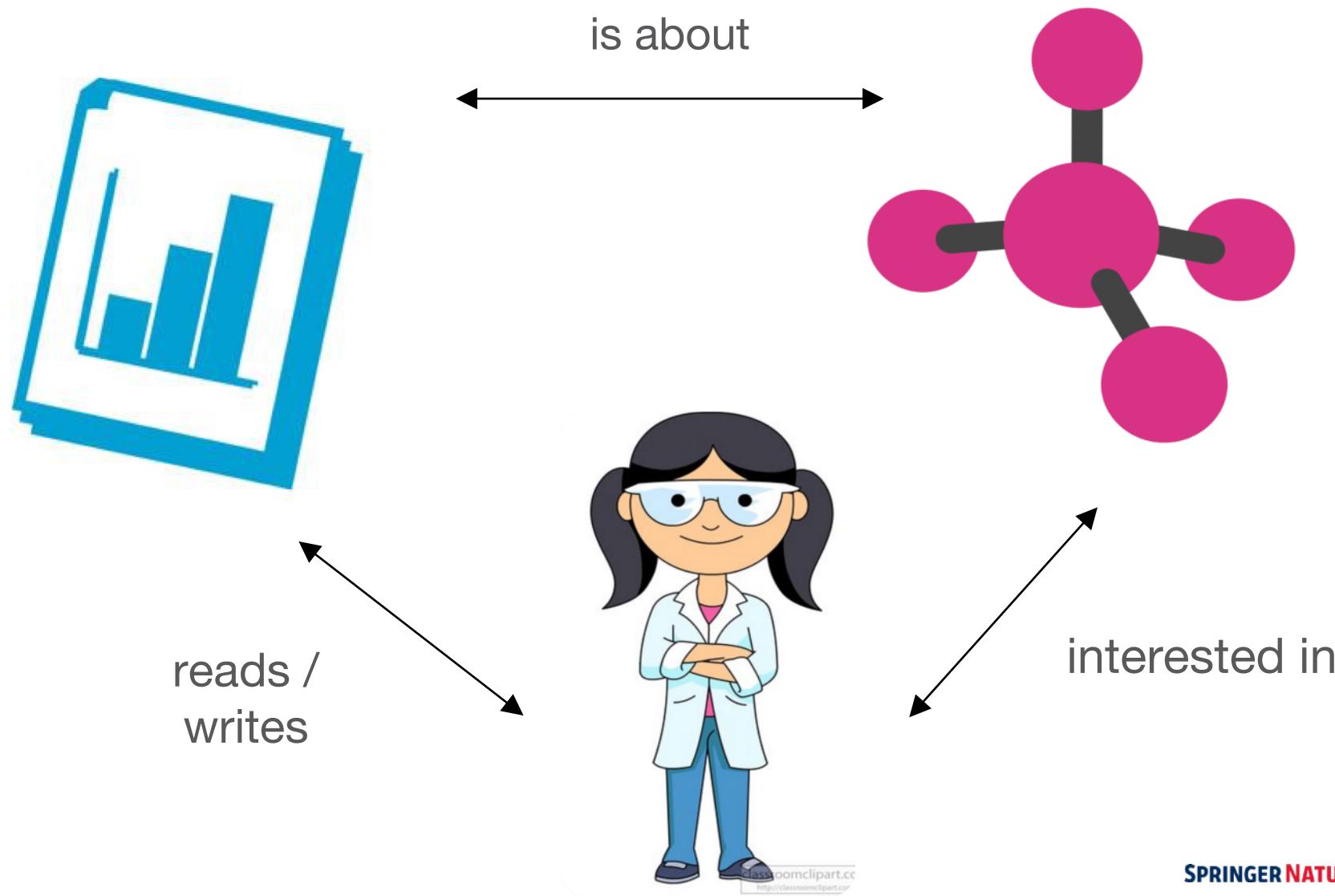
Vision: From Content to Data

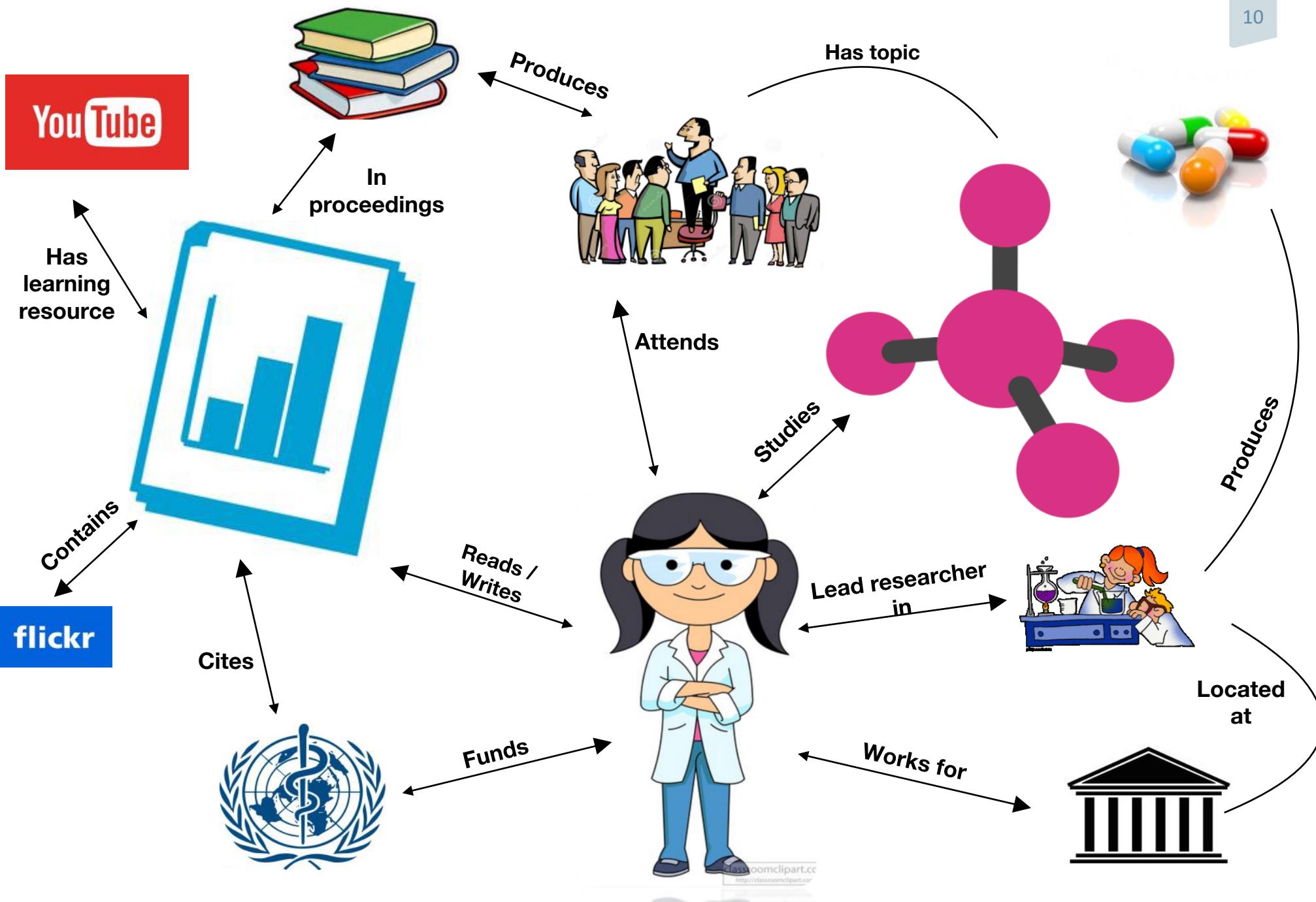
- We create the largest state-of-the-art linked open data aggregation platform for the scholarly domain.
- In doing so, we increase content discoverability and provide data tools and services for researchers, authors, editors, librarians, data scientists, funders, conference organizers, and many others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

We publish content

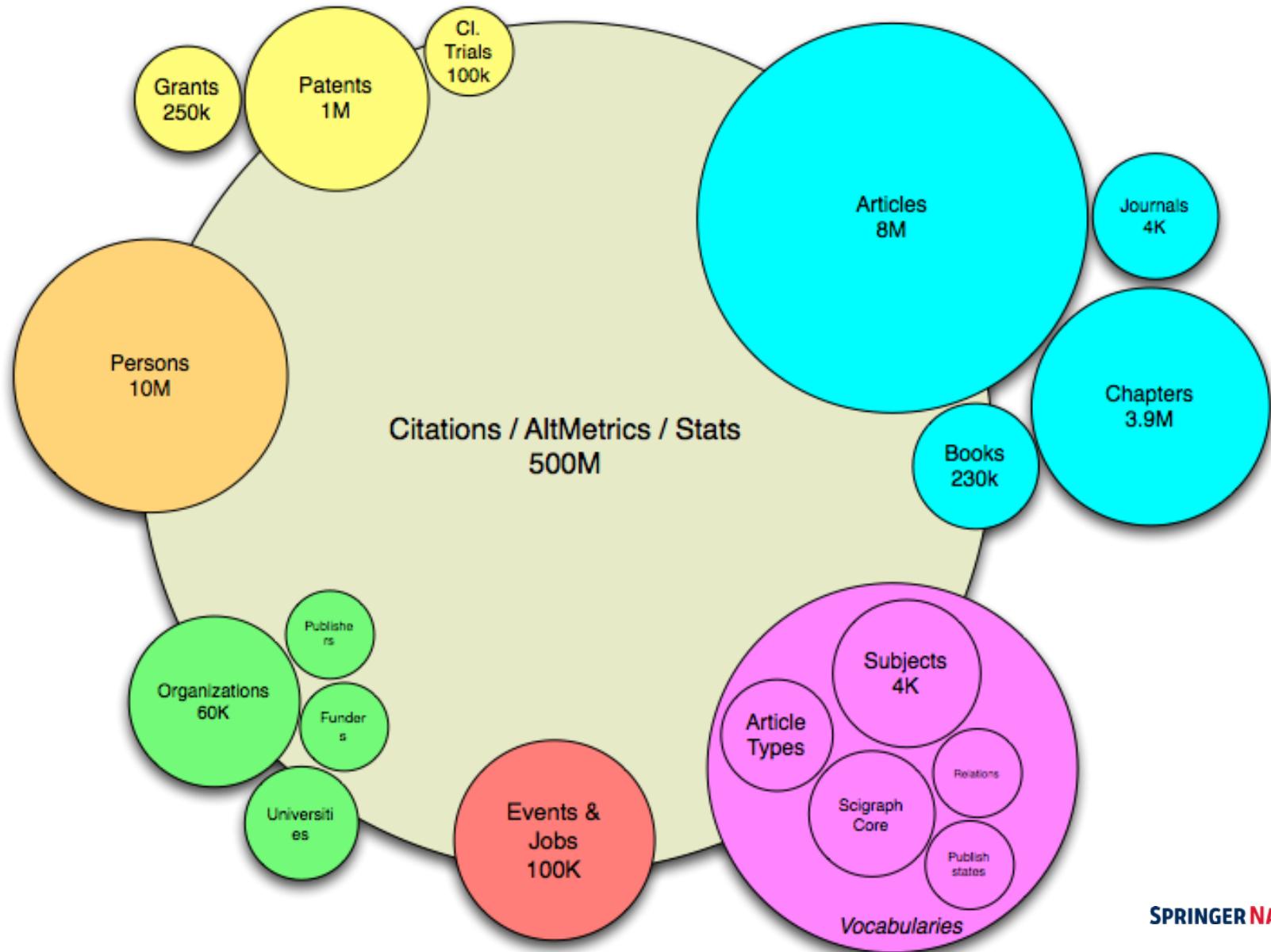
We manage knowledge

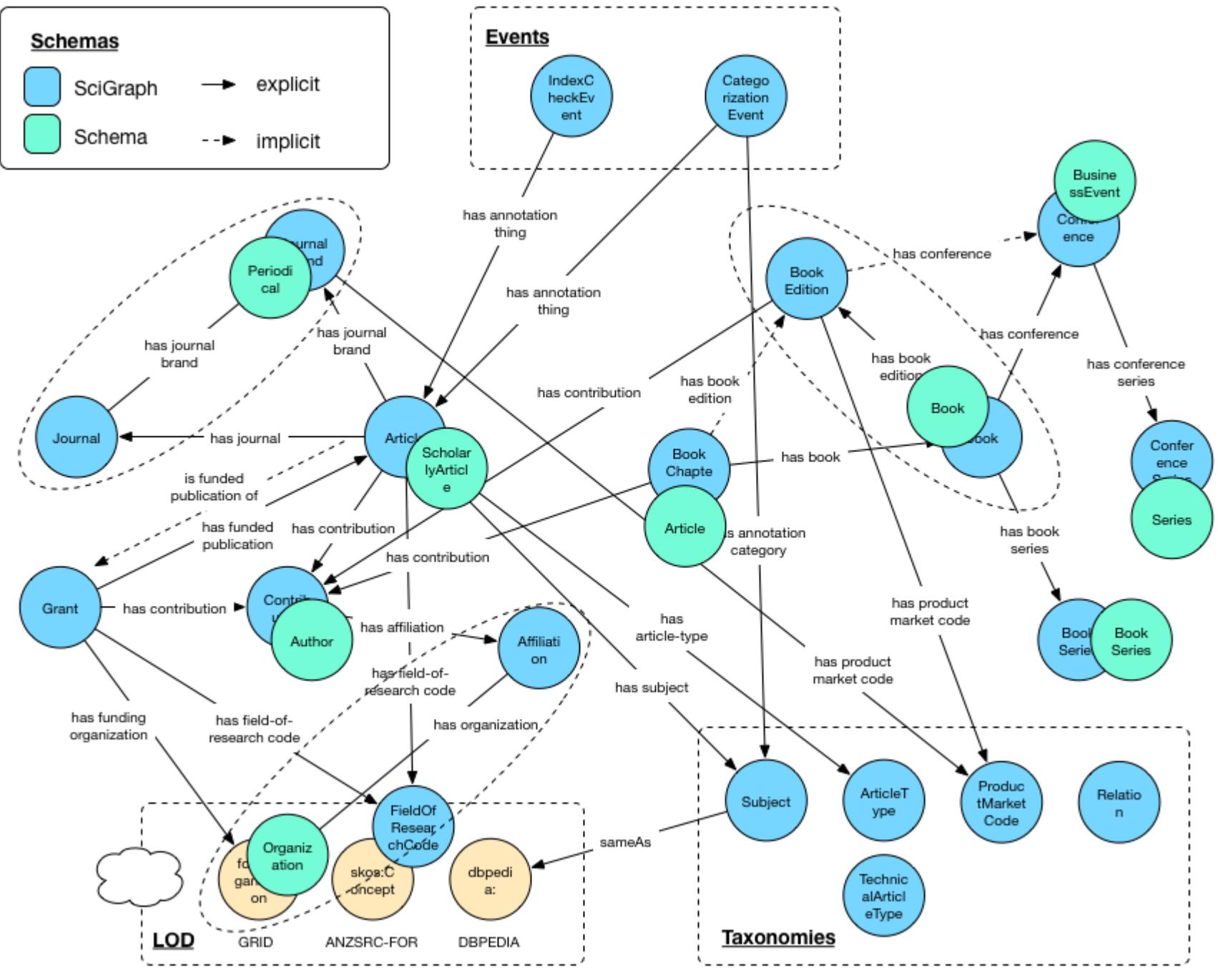
Three areas of knowledge we care about





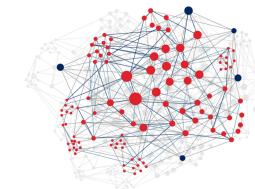
Springer Nature SciGraph Data Landscape





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.

Article - Total number Article - Count from 2012

841 **201**

Articles in Total Articles Published Last 5 Years

Article - Count by publication year

Count publicationYear

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.abs.gov.au/ausstats/abs@.nsf/088B427A8988C205CA27471800044E3E>

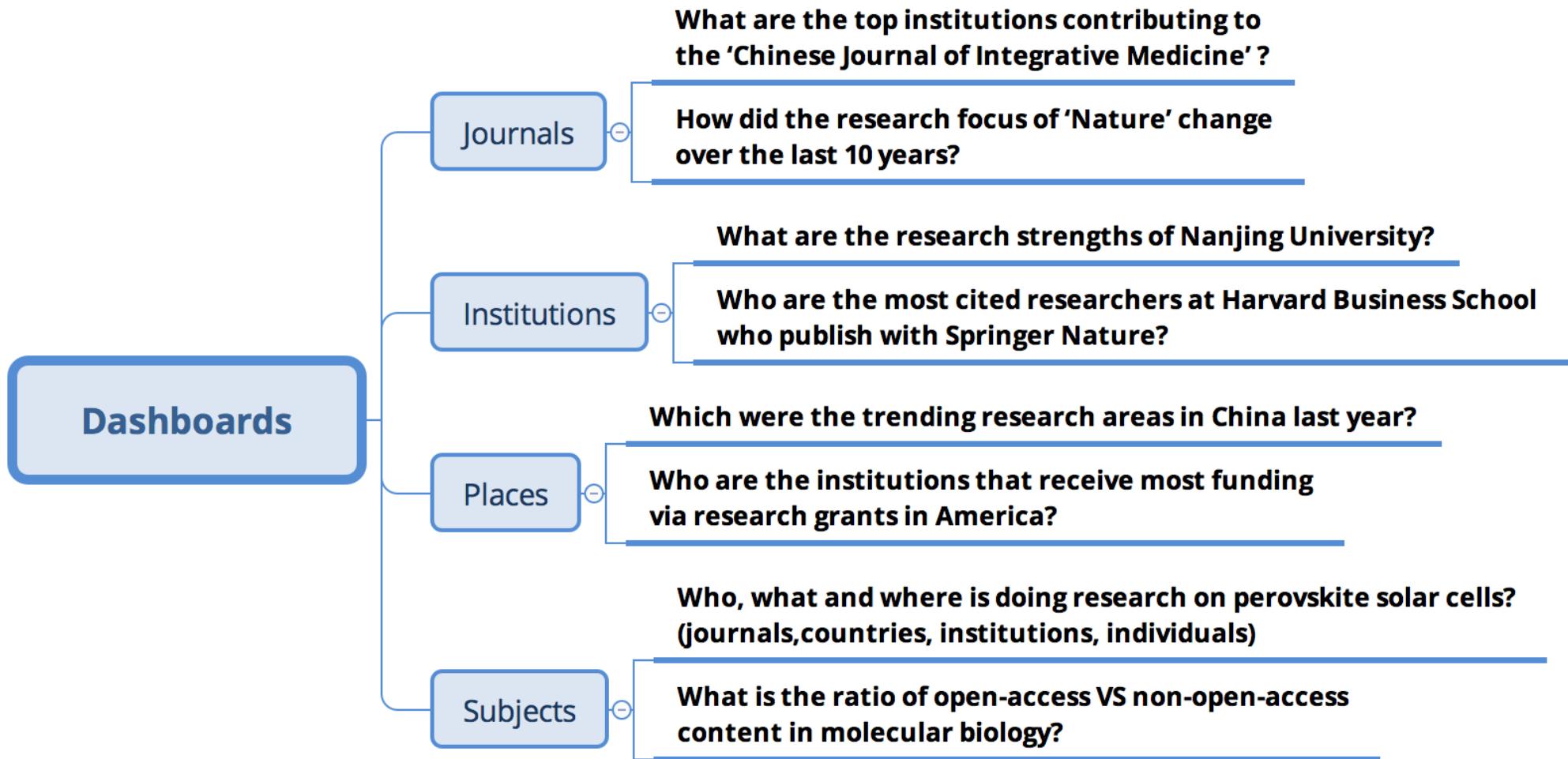
Article - FieldOfResearch by code and description Article - top 15 Fields of Research over time

Percentage of publications

Percentage of publication from range

- PHYSICAL SCIENCES
- OTHER PHYSICAL SCIENCES
- NEUROSCIENCES
- MEDICAL AND HEALTH SCIENCES
- IMMUNOLOGY
- GENETICS
- PHYSICAL SCIENCES
- CAMPUS/PROGRAMME
- BIOLOGICAL SCIENCES
- BIOTECHNOLOGY AND BIOPROCESS
- TECHNOLOGY
- STATISTICS
- PLANT BIOLOGY
- PHYSIOLOGY/MOLECULAR MEDICINE
- ONCOLOGY AND CARCINOGENESIS
- MEDICAL MICROBIOLOGY
- MEDICAL BIOTECHNOLOGY
- MEDICAL BIOCHEMISTRY
- MATHEMATICAL SCIENCES
- INFORMATION AND COMPUTER SCIENCE
- ANIMALS/FAUNA

Springer Nature SciGraph Analytics: Supporting Data Driven Decisions

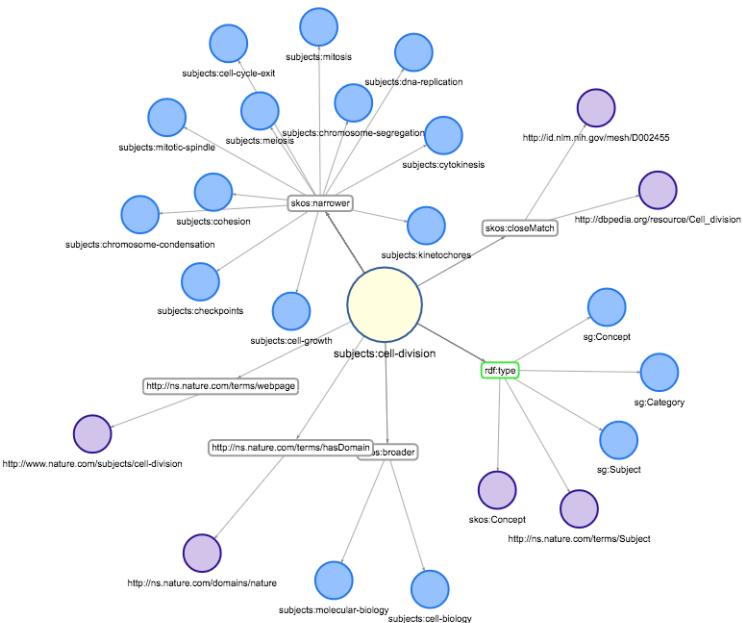


Applications Linked Data Explorer

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

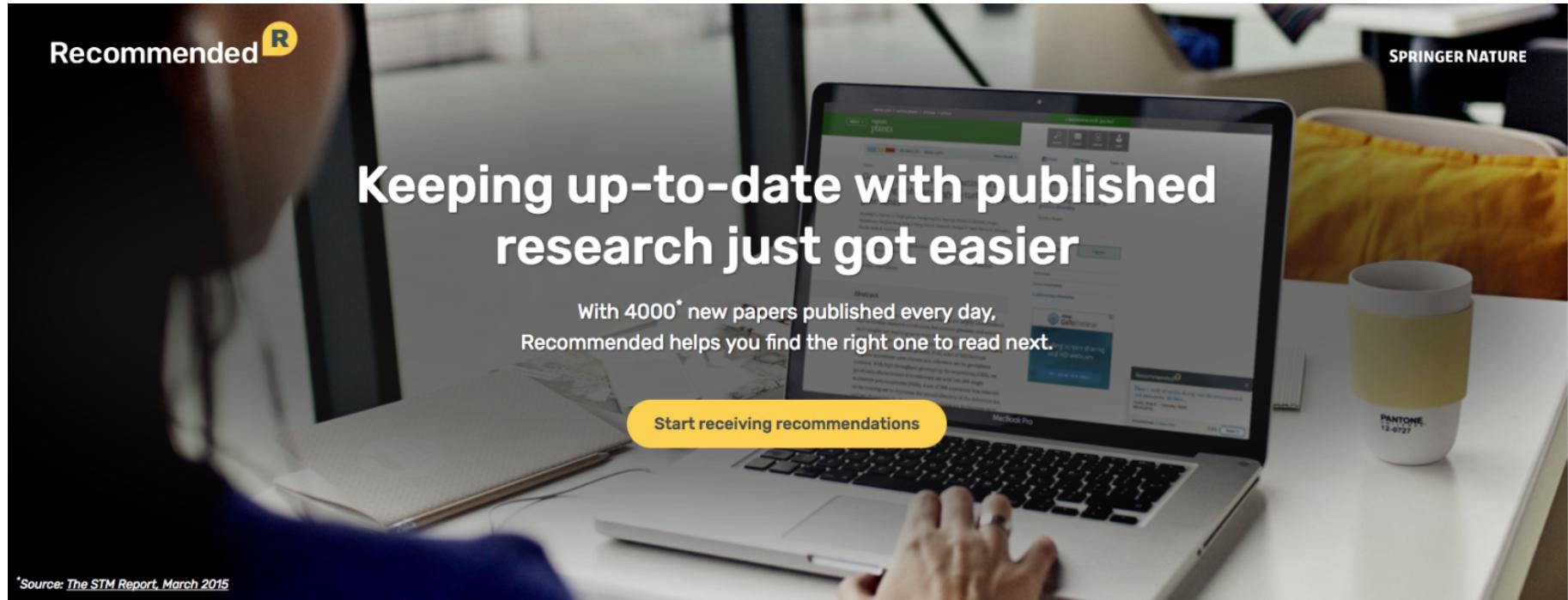
Looking Ahead

- **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



*Source: *The STM Report*, March 2015

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:

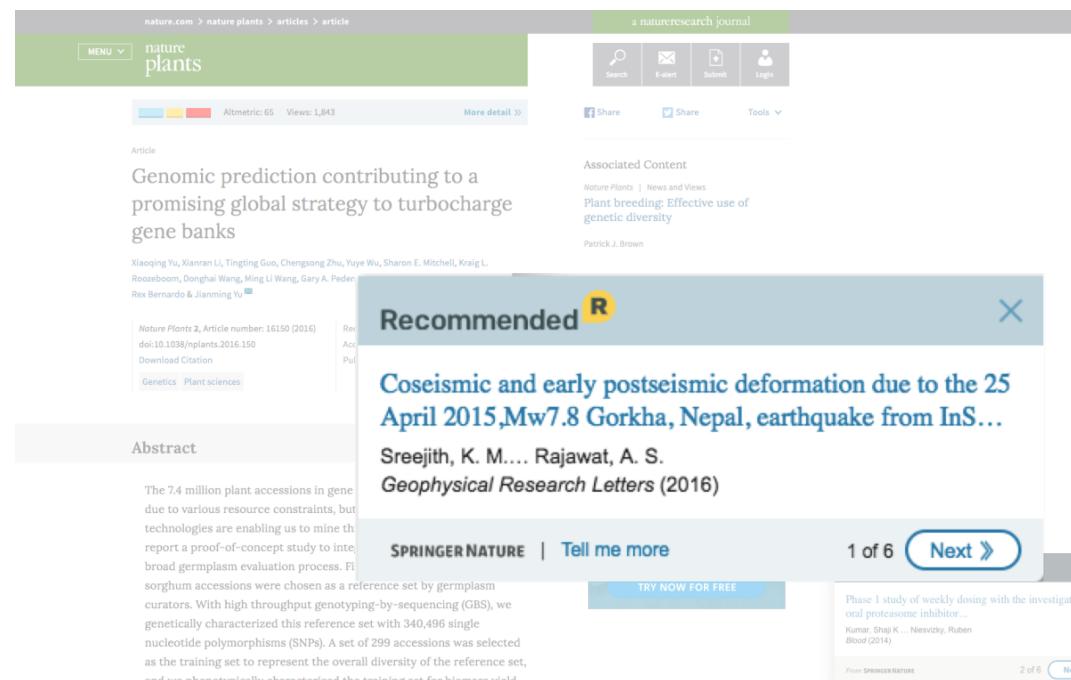
Only shown on article pages across Nature.com, BioMed Central and SpringerOpen.

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/>



How we deliver recommendations

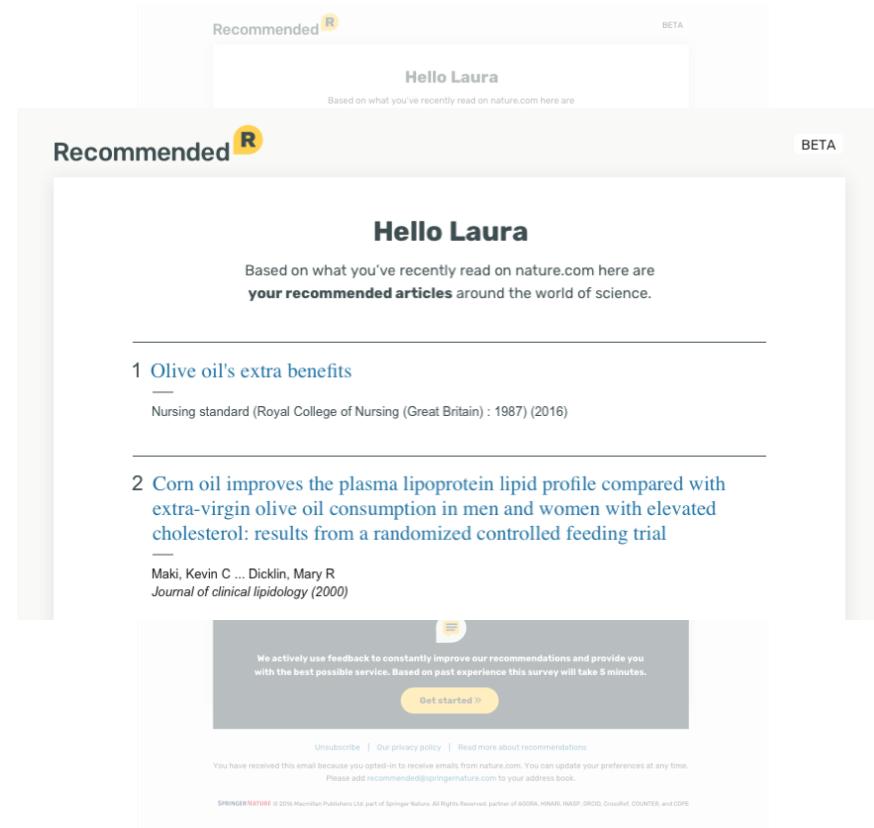
Email:

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>



The image shows the SharedIt landing page. At the top left is the Springer Nature logo (SN). To its right is the word "SharedIt". The background features a light blue gradient with white dandelion seed heads. Three dark blue rectangular boxes contain text: the first box on the left says "SharedIt allows Springer Nature authors and subscribers to share groundbreaking research for free on a global level"; the middle box says "ReadCube technology makes sharing Springer Nature SharedIt article links quick and easy"; and the box on the right says "Select from 2,300 journals from world leaders in research, including Nature Research, Springer, BioMed Central, and Palgrave Macmillan". At the bottom right is the Springer Nature logo.

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 a Nature journal article titled "The Sagittarius impact as an architect of spirality and outer rings in the Milky Way". The article is by Chris W. Purcell, James S. Bullock, Erik J. Tollerud, Miguel Rocha, and Sukanya Chakrabarti. It is from Volume 477, Issue 7364, published on 15 September 2011. A callout box highlights the "Shareable Link" button, which generates the URL <http://rdcu.be/bHzr>. Below the link are icons for various social media platforms: CiteULike, Facebook, Twitter, Delicious, Digg, Google+, LinkedIn, Reddit, and StumbleUpon.

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

The first screenshot shows a "Compose new Tweet" window with the text "Minor collisions affect galaxy morphology" and the shareable link <http://rdcu.be/bG6K>. The second screenshot shows the ReadCube app interface with the same article and shareable link. The third screenshot shows the final view of the Nature article page with the shareable link <http://rdcu.be/bG6K> highlighted.

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 Sagittarius impact as an architect of spirality and outer rings in the Milky Way

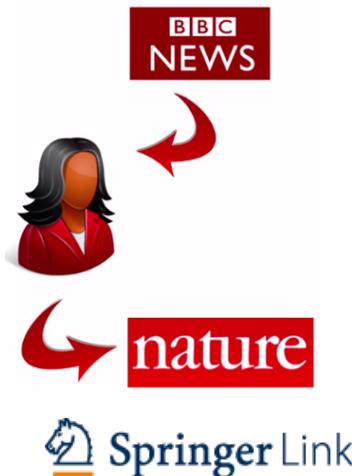
Chris W. Purcell^{1,2}, James S. Bullock², Erik J. Tollerud², Miguel Rocha² & Sukanya Chakrabarti³

Like many galaxies of its size, the Milky Way is a disk with prominent spiral arms rooted 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 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 the model that are reminiscent of the low-latitude arcs observed in the outer disk of the Milky Way. Previous models have focused on Sgr itself^{2,3} to reproduce the dwarf's orbital history and place associated constraints on the shape of the Milky Way's gravitational potential, treating the Sgr impact 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 Universe probably have a similarly important role in shaping galactic structure.

To discuss 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



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 window with the URL www.scientificamerican.com/article/sagittarius-satellite-spiral/. The page content discusses the Sagittarius Dwarf Galaxy and its interaction with the Milky Way. A central image shows a simulation of the galaxy's spiral arms. The text includes a quote from Erik Tollerud and a link to a video simulation. On the right side of the page, there is a sidebar with an advertisement for XARELTO and links to 'Latest News' and 'Most Read' articles.

Star-Crossed: Milky Way's Sagittarius Satellite Spiral

with its symmetric spiral arms winding outward from a central bulge, may be scars from a smaller galaxy punching above its weight. A new computer re-enactment of billions of years of galactic evolution suggests that the Milky Way owes much of its current shape to interactions with a nearby dwarf galaxy.

The Sagittarius Dwarf Galaxy, first discovered in 1994, is a satellite galaxy that is slowly being torn apart and ingested into the larger Milky Way. In the process, however, Sagittarius seems to have been making its presence felt. A group of astrophysicists at the University of Pittsburgh, the University of California, Irvine, and Florida Atlantic University simulated the gravitational infall of Sagittarius over the past few billion years to uncover what effects the dwarf galaxy may have had on the Milky Way. [Read more about the structure of the Milky Way.]

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. (Scientific American is part of Nature Publishing Group.) The resulting galactic perturbation also resulted in the development of loose strands of stars at its periphery that resemble an outer Milky Way feature known as the Monoceros ring. [See a video of the simulation below.]

Have you suffered severe side effects such as INTERNAL BLEEDING from XARELTO?

Click On This Ad To File Your Claim For Compensation

THIS MESSAGE BROUGHT TO YOU BY WENTON LEGAL SERVICES, Poughkeepsie, NY (866) www.wentonlaw.com

Latest News **Most Read**

Physicists and Philosophers Unite to Study Time's Arrow

First Clues Found in Mysterious Sea Star Die-Off

Life on the Volcano Is Increasingly Tough for These Hawaiian Birds

Better Barley Let People Settle Tibetan Plateau

Amazon and Google Change Places on Going Green

Follow Us: [Facebook](#) [Twitter](#) [Google+](#) [YouTube](#) [RSS](#)

See what we're tweeting about [Scientific American Contributors](#)

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](#)