如何通过信息学策略，跟上最新的研究进展?

杜建
北京大学医学部

概要

一、文献海洋中，如何开展精准查询法和引文跟踪法？

几个习惯都要有：

- 浏览 vs. 查询；
- 被动接收 vs. 主动检索；
- 文档查询 vs. 文本挖掘

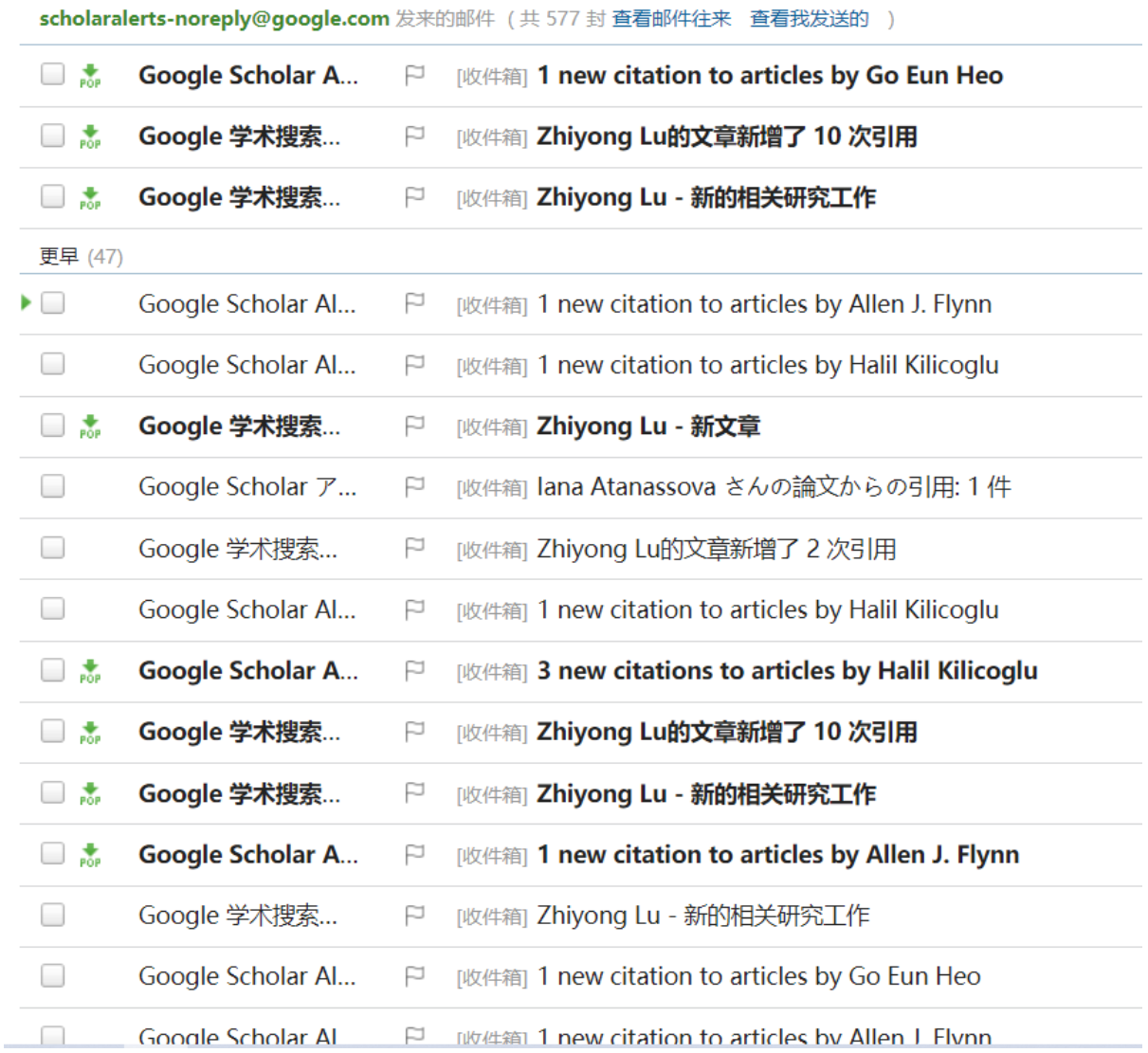
二、大数据时代，如何开展全域信息获取？

- 超越文献，定位前端的基金资助、后端的专利技术等。

- 精准查询法
  - site: gov.cn
  - intitle:
  - filetype:

- 浏览法
  - 定制领域权威期刊
  - 每期自动发送
  - 交叉性启发

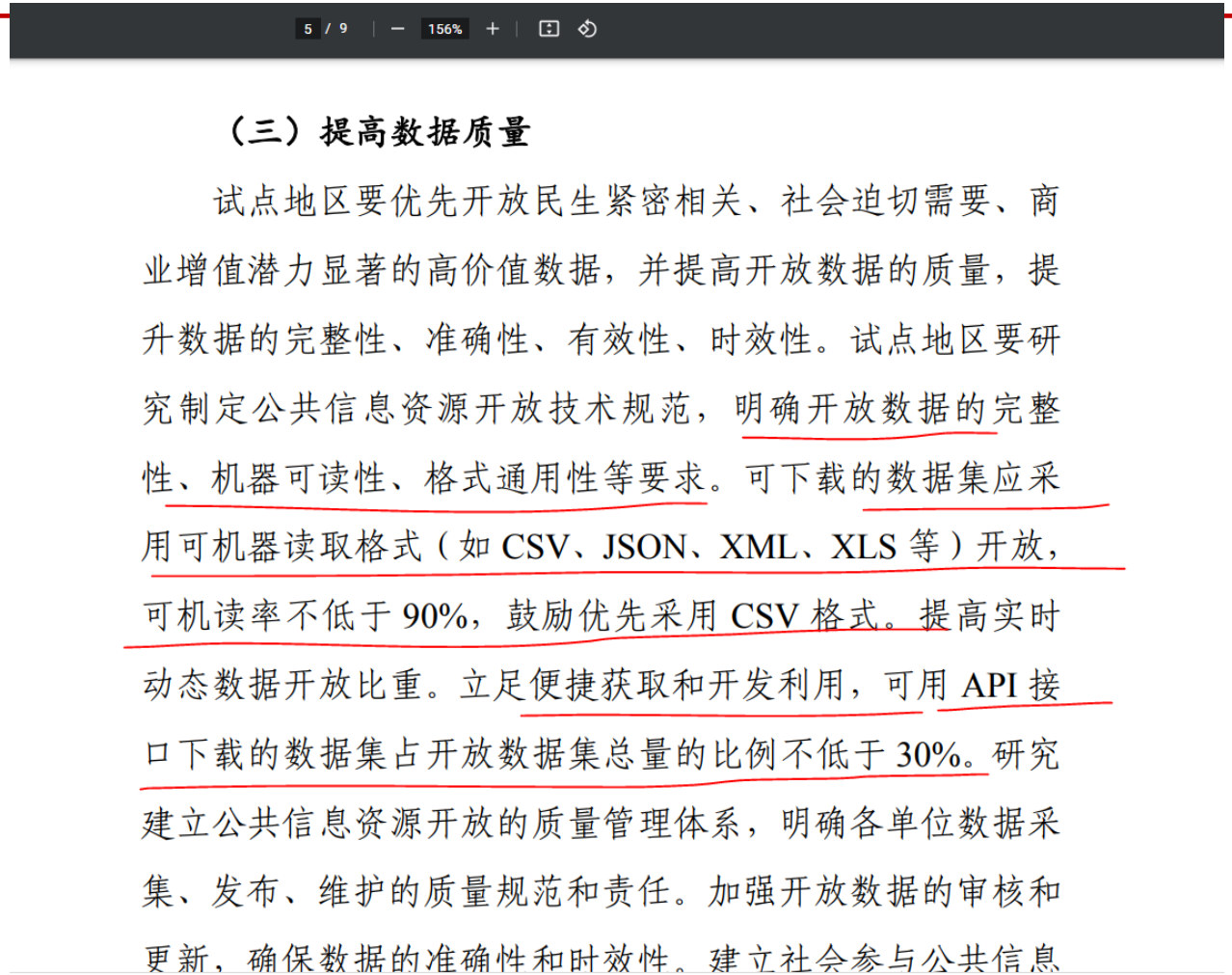
- 引文跟踪法



Google Scholar Alerts

碎片化时间读文献，搜索引擎与数据库Alerts

中央网信办、发展改革委、工业和信息化部联合印发《公共信息资源开放试点工作方案》（2018）



（三）提高数据质量

试点地区要优先开放民生紧密相关、社会迫切需要、商业增值潜力显著的高价值数据，并提高开放数据的质量，提升数据的完整性、准确性、有效性、时效性。试点地区要研究制定公共信息资源开放技术规范，明确开放数据的完整性、机器可读性、格式通用性等要求。可下载的数据集应采用可机器读取格式（如 CSV、JSON、XML、XLS 等）开放，可机读率不低于 90%，鼓励优先采用 CSV 格式。提高实时动态数据开放比重。立足便捷获取和开发利用，可用 API 接口下载的数据集占开放数据集总量的比例不低于 30%。研究建立公共信息资源开放的质量管理体系，明确各单位数据采集、发布、维护的质量规范和责任。加强开放数据的审核和更新，确保数据的准确性和时效性。建立社会参与公共信息

http://www.cac.gov.cn/2018-01/05/c_1122215495.htm

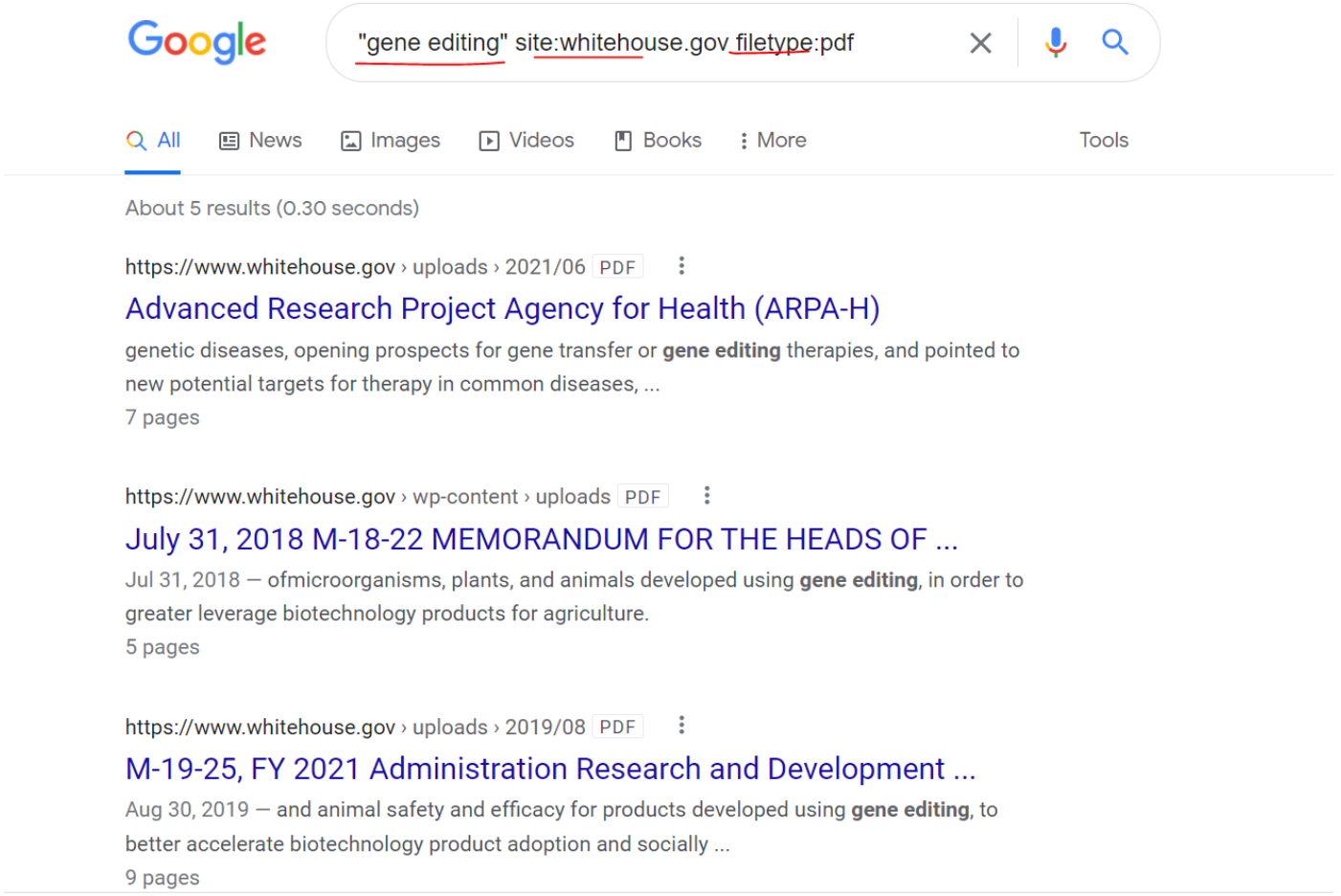
intitle:

site:

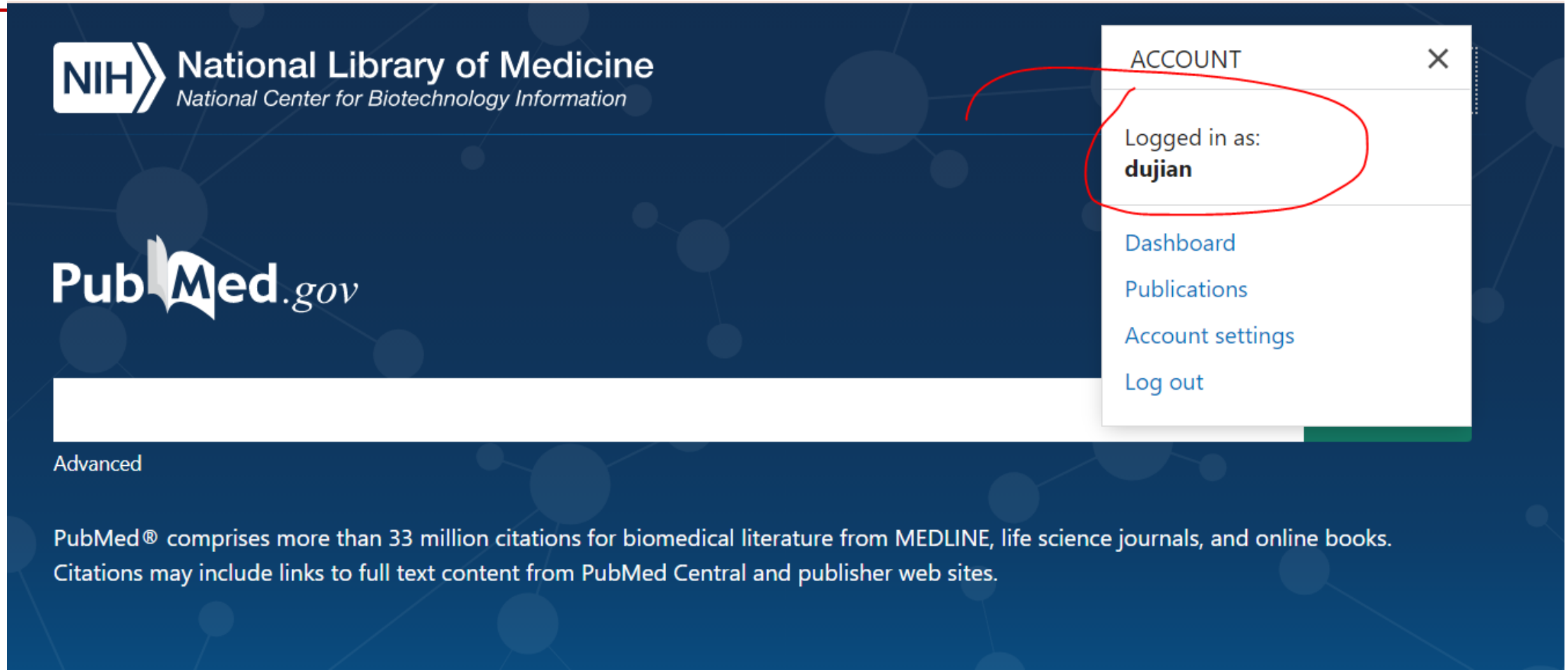
filetype:

搜索引擎精准获取：权威机构发布的报告

- "gene editing" site:whitehouse.gov filetype:pdf



注册 My NCBI


NIH
National Library of Medicine
National Center for Biotechnology Information
ACCOUNT
Logged in as:
dujian
Dashboard
Publications
Account settings
Log out
PubMed.gov
Advanced
PubMed® comprises more than 33 million citations for biomedical literature from MEDLINE, life science journals, and online books.
Citations may include links to full text content from PubMed Central and publisher web sites.

My NCBI

Search NCBI databases

Search : PubMed

Hint: clicking the "Search" button without any ter
that database's homepage.

Your PubMed search

Name of saved search: semrep[tiab] OR SemMedDB[tiab]

Search terms:

```
semrep[tiab] OR SemMedDB[tiab] OR "Semantic Medline"[tiab] OR "Knowledge Graph"[ti]
```

Test search terms

	What's New	Last Searched
⚙	0	3 days ago

Manage Saved Searches »

Would you like e-mail updates of new search results?

- ○ No, thanks.
- ◉ Yes, please.

E-mail: windowsdujian@163.com (change)

Schedule:

Frequency: Daily

Formats:

Report format: Abstract

Number of items:

Send at most: 20 items ☐ Send even when there aren't any new results

Any text you want to be added at the top of your e-mail (optional):

Save Cancel Delete

My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		9月25日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		9月21日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		9月14日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		9月8日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		9月4日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		8月20日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		7月31日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		7月24日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		7月21日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		7月20日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		7月19日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		7月13日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		7月9日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		7月3日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		7月2日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		6月30日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		5月8日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		3月20日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		3月6日
My NCBI	[收件箱] What's new for 'semrep[tiab] OR SemMedDB[tiab]' in PubMed		2月27日

例：青蒿素获诺奖的若干早期线索

Nobel Prizes and Laureates

Medicine Prizes | 2015

▼ About the Nobel Prize in Physiology or Medicine 2015
- Summary
- Prize Announcement
- Press Release
- Advanced Information
- Greetings
- Award Ceremony Video
- Award Ceremony Speech

▶ William C. Campbell
▶ Satoshi Ōmura
▶ Youyou Tu

All Nobel Prizes in Physiology or Medicine
All Nobel Prizes in 2015

The Nobel Prize in Physiology or Medicine 2015
William C. Campbell, Satoshi Ōmura, Youyou Tu

Share this: 2.6K

The Nobel Prize in Physiology or Medicine 2015

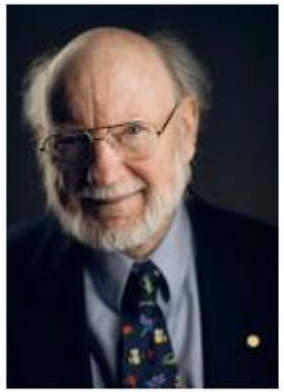


Photo: A. Mahmoud
William C. Campbell
Prize share: 1/4

Photo: A. Mahmoud
Satoshi Ōmura
Prize share: 1/4

Photo: A. Mahmoud
Youyou Tu
Prize share: 1/2

The Nobel Prize in Physiology or Medicine 2015 was divided, one half jointly to William C. Campbell and Satoshi Ōmura *"for their discoveries concerning a novel therapy against infections caused by roundworm parasites"* and the other half to Youyou Tu *"for her discoveries concerning a novel therapy against Malaria"*.

▼ About the Nobel Prize in Physiology or Medicine 2015
Summary
Prize Announcement
Press Release
Advanced information
Greetings
Award Ceremony Video
Award Ceremony Speech

▶ William C. Campbell
▶ Satoshi Ōmura
▶ Youyou Tu

All Nobel Prizes in Physiology or Medicine
All Nobel Prizes in 2015

English
English (pdf)
Swedish
Swedish (pdf)

Nobelförsamlingen
The Nobel Assembly at Karolinska Institutet

Press Release

2015-10-05

The Nobel Assembly at Karolinska Institutet has today decided to award

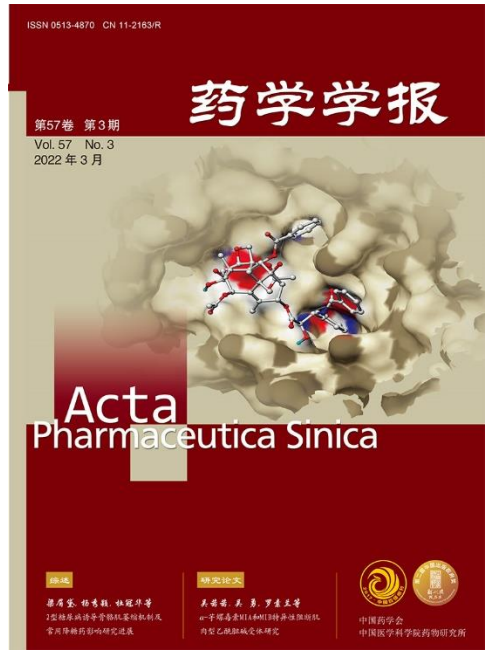
the 2015 Nobel Prize in Physiology or Medicine

Key publications:

Burg et al., *Antimicrobial Agents and Chemotherapy* (1979) 15:361-367.

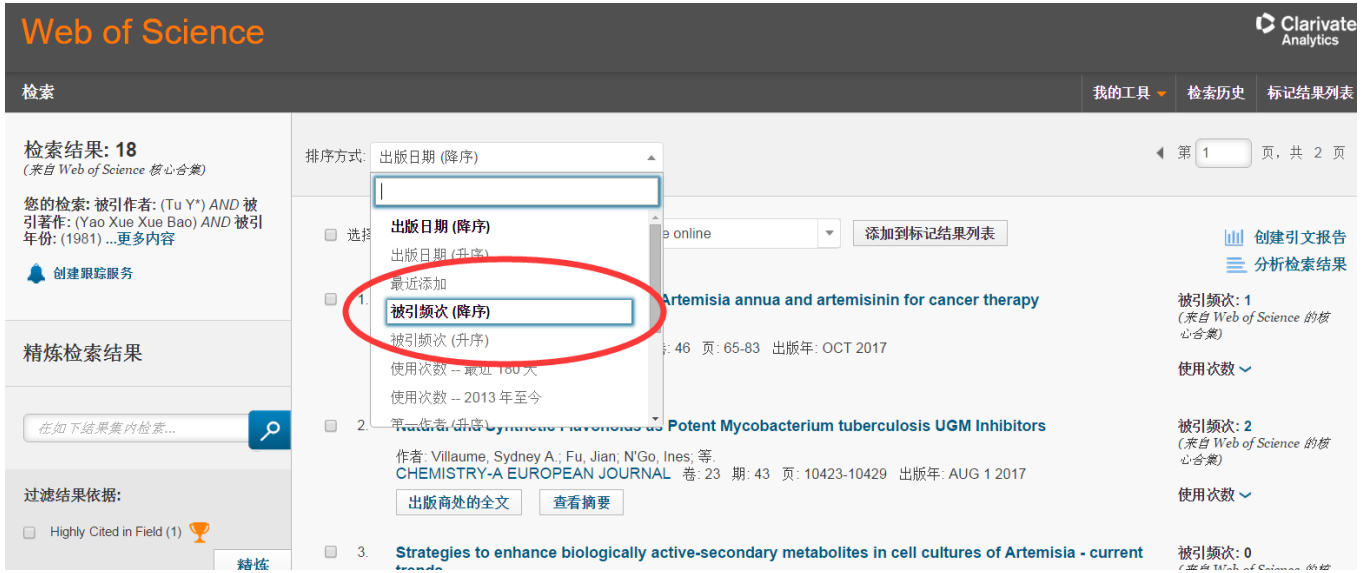
Egerton et al., *Antimicrobial Agents and Chemotherapy* (1979) 15:372-378.

Tu et al., *Yao Xue Xue Bao* (1981) 16, 366-370 (Chinese)



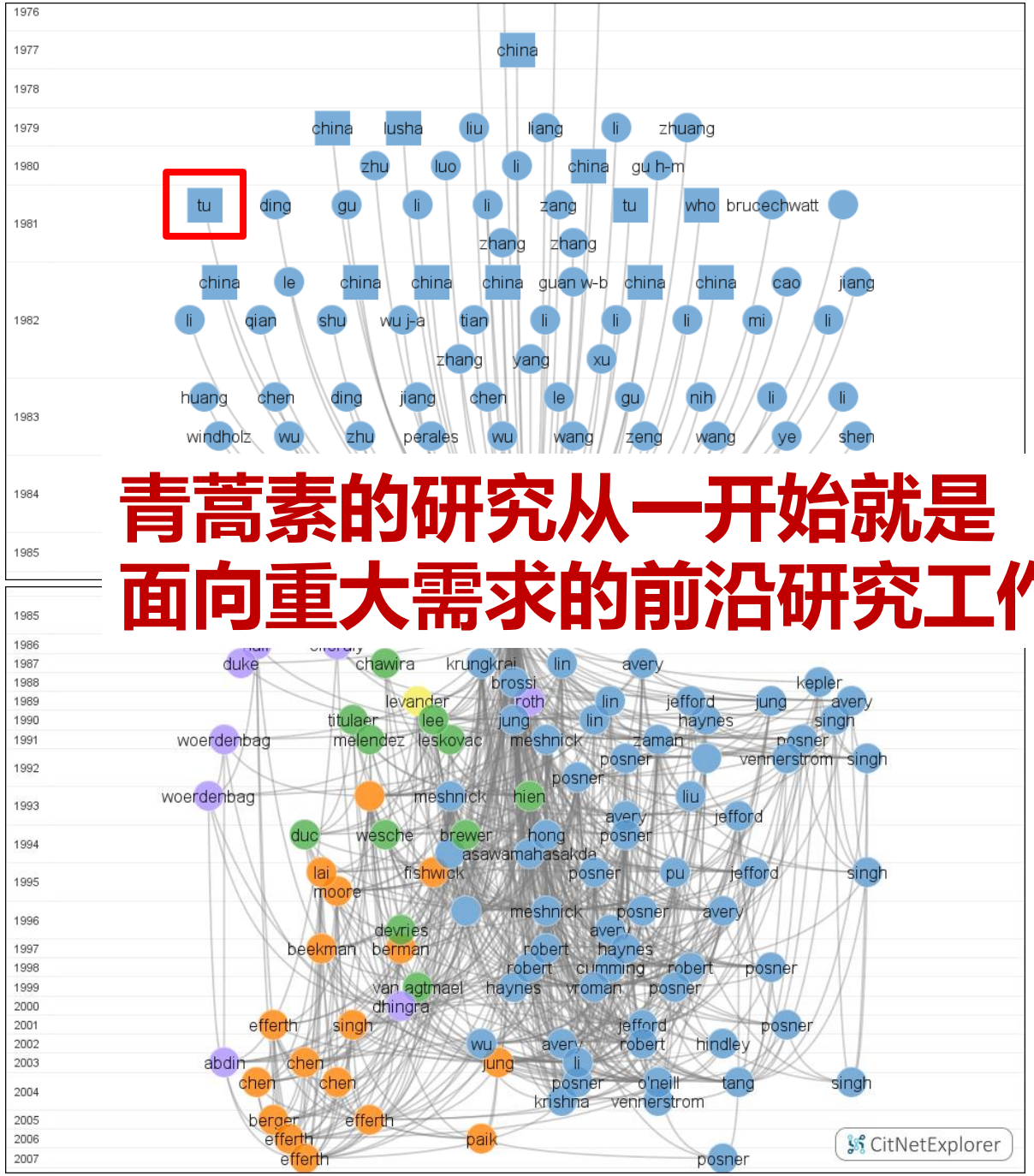
这篇中文论文在世界上的影响如何?

能查到在SCI中的被引情况吗?



青蒿素科学与技术的经典文献

论文	Cited by artemisinin related patents	Cited by artemisinin related publications
Klayman D. (1985)	19	1361

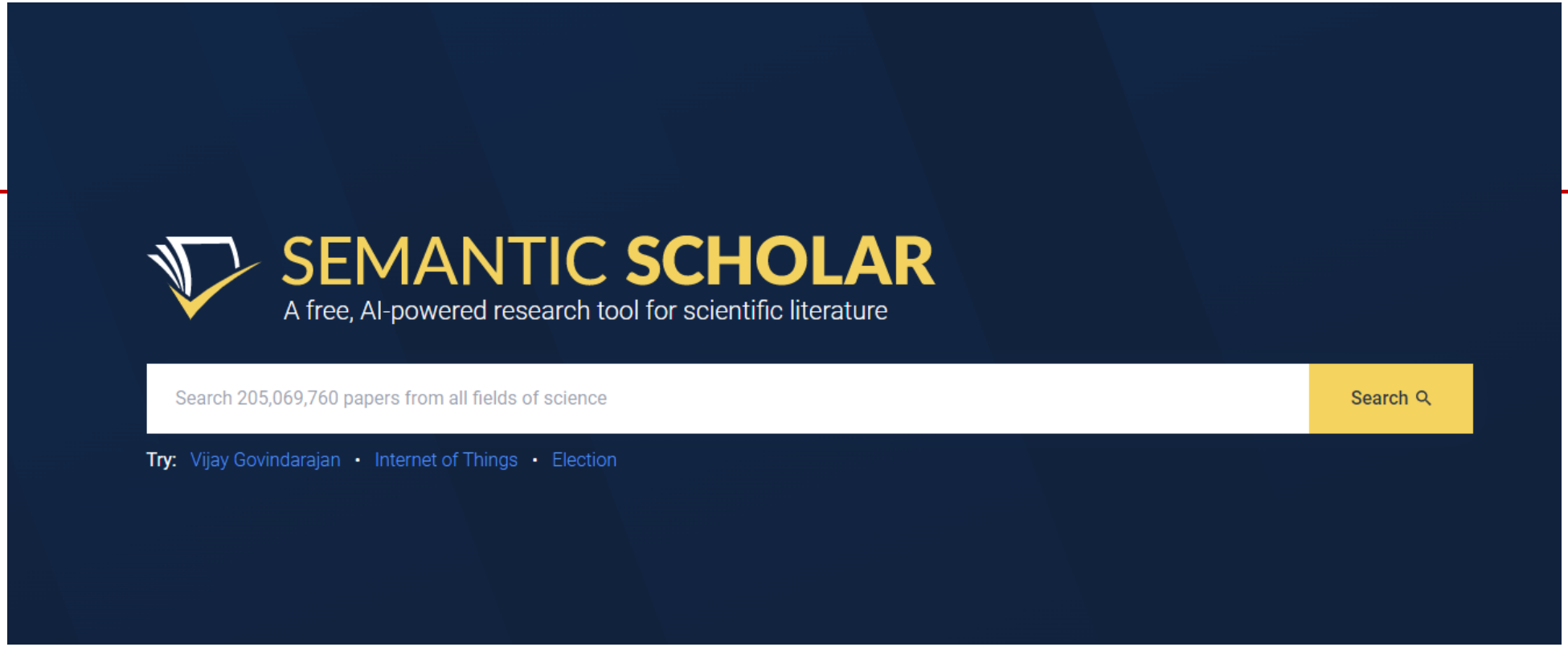


青蒿素的研究从一开始就是面向重大需求的前沿研究工作！

92 cited references，涉及43种期刊、著作或会议；
含中国学者以中文发表的所有著作（方框）

ıan, 1985 (*Science*)

截至2017年11月20日，被引用1558次：
70+ Subject Categories,
73 countries/territories，
500+ Journals



- 由微软联合创始人 Paul Allen运用机器学习技术创建的一款免费学术搜索引擎。
- 囊括了1.8亿多篇学术论文，涉及计算机科学、分子生物学、微生物学和神经科学等多学科领域。

https://www.semanticscholar.org/

DOI: 10.1038/nature.2016.20988 · Corpus ID: 4456184

CRISPR gene-editing tested in a person for the first time

David Cyranoski · Published in Nature 2016 · Biology, Medicine

Citations

Publications citing this paper.

CITATION TYPE

All Types

SORT BY

Influence

SHOWING 1-10 OF 120 CITATIONS, ESTIMATED 67% COVERAGE

First-in-human Phase 1 CRISPR Gene Editing Cancer Trials:Are We Ready?

Françoise Baylis, Marcus McLeod • Medicine, Biology • Current gene therapy • 2017

VIEW 19 EXCERPTS | CITES BACKGROUND | HIGHLY INFLUENCED

Genome Editing: A New Horizon for Oral and Craniofacial Research

Ning Yu, Jie-gang Yang, Yuji Mishina, William V Giannobile • Medicine • Journal of dental research • 2019

VIEW 4 EXCERPTS | CITES BACKGROUND | HIGHLY INFLUENCED

Advancements and Obstacles of CRISPR-Cas9 Technology in Translational Research

Liting You, Ruizhan Tong, +3 authors You Lu • Medicine, Computer Science • Molecular therapy. Methods & clinical

Application of CRISPR/Cas9-Based Gene Editing in HIV-1/AIDS Therapy

Qiaoqiao Xiao, Deyin Guo, Shuliang Chen • Biology, Medicine • Front. Cell. Infect. Microbiol. • 2019

≡ VIEW 1 EXCERPT | CITES BACKGROUND

CITES BACKGROUND FROM "CRISPR GENE-EDITING TESTED IN A PERSON FOR THE... "

The modified human T cells were reinfused back into an individual with metastatic nonsmall cell lung cancer, which is supposed to have a promising result (**Cyranoski, 2016**).

Attitudes of clinical geneticists and certified genetic counselors to genome editing and its clinical applications: A nation-wide questionnaire survey in Japan

Iku Taguchi, Takahiro Yamada, +10 authors Shinji Kosugi • Biology, Medicine • Journal of Human Genetics • 2019

≡ VIEW 1 EXCERPT | CITES BACKGROUND

CITES BACKGROUND FROM "CRISPR GENE-EDITING TESTED IN A PERSON FOR THE... "

This technology has been at research stage and has not matured enough for clinical applications yet, except for some clinical trials **[28, 29]**.

Paper Mentions

NEWS ARTICLE

Should we be Into Eugenics?

Medium US
27 February 2020

NEWS ARTICLE

Quest to Use CRISPR Against Disease Gains Ground

Scientific American
7 January 2020

NEWS ARTICLE

What CRISPR-baby prison sentences mean for research

Google News
3 January 2020

BLOG POST

CRISPR-modifizierte Immunzellen bei ersten Krebspatienten sicher eingesetzt

medONLINE
7 November 2019

BLOG POST

Inside China's Play to Become the World's CRISPR Superpower

Singularity Hub
18 August 2019

NEWS ARTICLE

The experiment that geneticists say went too far

CNN News
1 December 2018

NEWS ARTICLE

The FDA Has Just Put The Brakes on a Major CRISPR Trial in Humans

Science Alert
1 June 2018

NEWS ARTICLE

Gentechnik: Genschere verschnippelt sich doch nicht

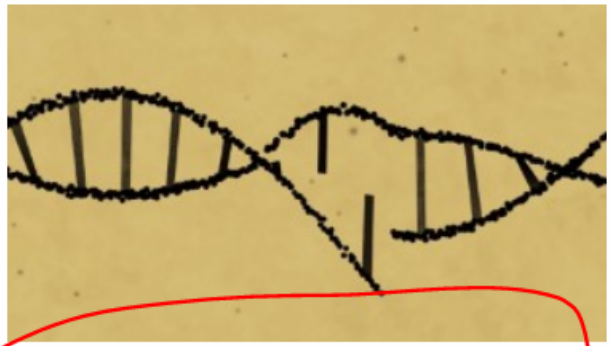
Der Tagesspiegel
3 April 2018

'A gold rush on new kind of knowledge'

That the announcement of the world's first genetically edited babies came out of China is perhaps not surprising.

Chinese scientists have accomplished many firsts in the genome world, including the first CRISPR-edited monkeys, the first use of the gene-editing tool CRISPR-Cas9 in humans, and the first reported use of gene editing technology to modify nonviable human embryos.

China has also pumped huge amounts of government money into gene-editing technology, using it to lure leading Chinese scientists living abroad back to the country, as well as foreigners who see the country as fertile ground for this kind of research.



Related Article: Chinese scientists use gene-editing techniques in humans for first time

"I just think right now China is a lot more driven, they incentivize their scientists to move faster and be bolde and it shows," said Victor J. Dzau, President of the Institute of Medicine at the US National Academy of Medicine.

Last year, China spent a record 1.76 trillion yuan ($254 billion) on research and development, and the country is catching up with the US' investment in the same area, spurring a genetic arms race that has been labeled "Sputnik 2.0" by Dr. Carl June, an

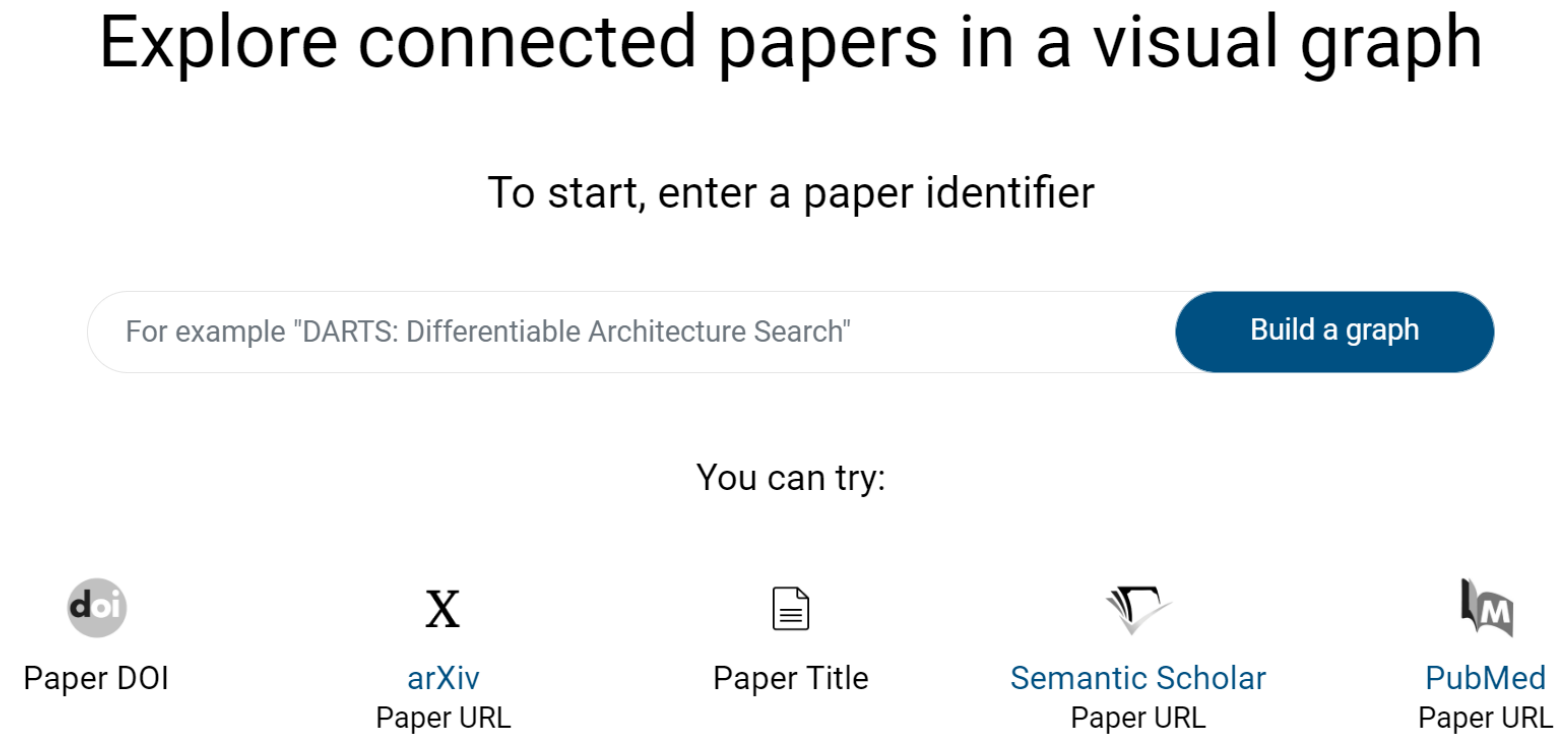
have been testing CRISPR's much less controversial ability to disable or correct problematic genes in other cells in order to treat a host of diseases.

In 2016, Chinese researchers announced that they had treated the first person with a CRISPR–Cas9 therapy designed to fight cancer. In cells extracted from a participant's blood, the researchers disabled the gene that codes for a protein called PD-1, which holds the immune system in check but can shield cancer cells in the process. The scientists then reinjected the cells.

By 2019, the US government's clinicaltrials.gov database listed more than a dozen active studies that are testing CRISPR–Cas9 as a treatment for a range of diseases from cancer to HIV and blood disorders.

Connected Papers：https://www.connectedpapers.com/

- 将相关领域已发表论文进行关联分析
- Prior Works vs. Derivative Works



CONNECTED PAPERS | CHARLS | Share | Follow | About | Sponsors | Feedback

Health Insurance and Long-Term Care Services for the Disabled Elderly in China: ... | Prior works | Derivative works | Sponsored by DagsHub

Search... | Expand

Origin paper

Health Insurance and Long-Term Care Services for the Disabled Elderly in...

Linhong Chen, Xiaolu Zhang, Xiaocang... 2020

The availability and affordability of long-term care for disabled older...

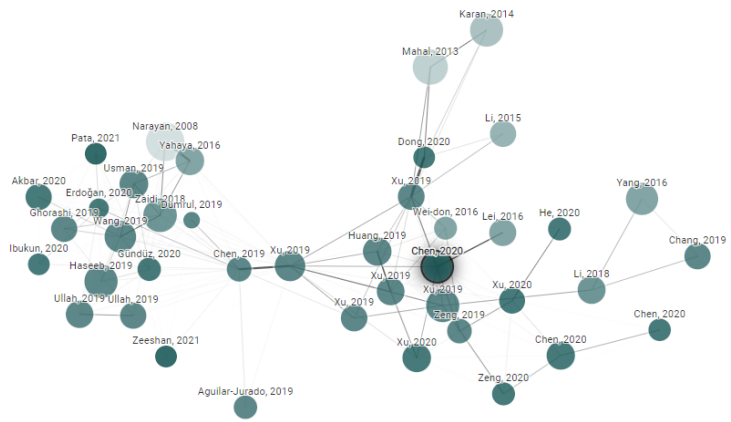
P. Lei, Zhixin Feng, Zhuo-chun Wu 2016

Family Economic Burden of Elderly Chronic Diseases: Evidence from China

Xiaocang Xu, Xiuquan Huang, Xiaolu... 2019

An assessment of the economic burden of senile chronic diseases in...

Xiayan Dong, Linhong Chen, Zhiming... 2020



Show updated graph

2008 2021

Health Insurance and Long-Term Care Services for the Disabled Elderly in China: Based on CHARLS Data

Linhong Chen, Xiaolu Zhang, Xiaocang Xu

2020, Risk management and healthcare pol...

28 Citations, 47 References

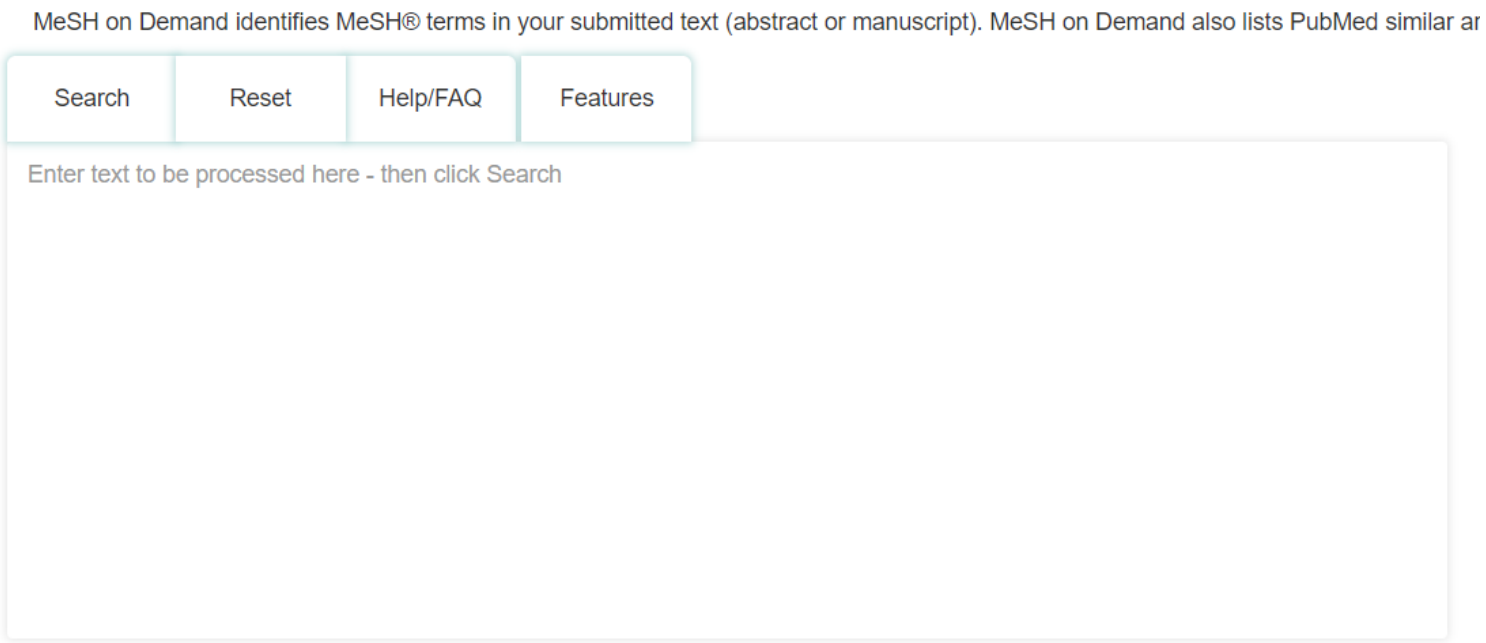
Open in:

Purpose This paper aimed to explore the relationship between the different factors, especially health insurance, and the availability of long-term care (LTC) services, among the disabled elderly. Methods Based on the data of China Health and Retirement Longitudinal Study (CHARLS), the logistic regression model was utilized

https://www.connectedpapers.com/

MeSH on Demand

- MeSH on Demand
  - identifies MeSH® terms in your submitted text (abstract or manuscript).
  - also lists PubMed similar articles relevant to your submitted text.



https://meshb.nlm.nih.gov/MeSHonDemand

MeSH on Demand identifies MeSH® terms in your submitted text (abstract or manuscript). MeSH on Demand also lists PubMed similar articles

Search | Reset | Help/FAQ | Features

Background: In China, there are two major government incentives in the development of medical informatics, yet there have been limited systematic investigations on the pattern of interactions between informatics concepts/techniques and the health/medical needs.

Objective: We propose an approach to mapping the interplay between different knowledge entities by using the tree structure of Medical Subject Headings (MeSH) to gain insights into the interactions between informatics-supply and health-demand in China.

Methods: All terms under the MeSH tree parent node “Diseases [C]” or node “Health [N01.400]” or “Public Health [N06.850]” were labelled as H. All terms under the node “Information Science [L]” as I and all terms under node “Analytical, Diagnostic and Therapeutic Techniques, and Equipment [E]” as T. The H-I and H-I-T interactions can be measured by using their co-occurrences in a given publication.

Results: In China, Medical informatics publications with both H- and I-related MeSH terms increased rapidly. Still, the overall H-I

MeSH on Demand identifies MeSH® terms in your submitted text (abstract or manuscript). MeSH on Demand also lists PubMed similar articles relevant to your submitted text.

Search | Reset | Help/FAQ | Features

Background In China there are two major government incentives in the development of medical informatics yet there have been limited systematic investigations on the pattern of interactions between informatics concepts techniques and the health medical needs Objective We propose an approach to mapping the interplay between different knowledge entities by using the tree structure of Medical Subject Headings MeSH to gain insights into the interactions between informatics supply and health demand in China Methods All terms under the MeSH tree parent node Diseases [C] or node Health [N01 400] or Public Health [N06 850] were labelled as H All terms under the node Information Science [L] as I and all terms under node Analytical Diagnostic and Therapeutic Techniques and Equipment [E] as T The H I and H I T interactions can be measured by using their co occurrences in a given publication Results In China Medical informatics publications with both H and I related MeSH terms increased rapidly Still the overall H I interactions are sparse if compared to the United States The most obvious interactions were only focused on several diseases with two primary informatics techniques In the United States almost all diseases have a strong interaction with informatics techniques especially using computing methodologies By establishing the H I T connection it was found that the H I T link in Chinese informatics research is weaker than in US research We recommend enhancing the technology transfer for medical informatics in China There is a positive correlation between the burden and the informatics research efforts for diseases in China Artificial Intelligence AI is a competing field of medical informatics research both China and the United States focusing on deep neural networks Besides AI technology the US has more emphasis on natural language processing Conclusions Research on evidence based medical informatics and electronic health records EHRs should be strengthened to improve the real world applications of health information technologies and big data in health and medicine in the future

Start PubMed Search

Export Data

MeSH Terms

- United States
- Medical Subject Headings
- Natural Language Processing
- Electronic Health Records
- Artificial Intelligence
- Big Data
- Technology Transfer
- Public Health
- Motivation
- Medical Informatics
- Information Science
- Computing Methodologies
- Neural Networks, Computer
- Government
- Asian Continental Ancestry Group
- China

PubMed/MEDLINE Similar Articles

The following articles are 10 similar PubMed Related Citations that were also used in computing these MeSH recommendations. The order is from most to least relevant. Selecting any of the titles opens a new window or tab with that related citation in PubMed's Abstract view.

1. *informatics: Identifying and Tracking Informatics Sub-Discipline Terms in the Literature. PMID: 25998007
2. The evolution of medical informatics in China: A retrospective study and lessons learned. PMID: 27318067
3. The Gap in Medical Informatics and Continuing Education Between the United States and China: A Comparison of Conferences in 2016. PMID: 28637638
4. Artificial Intelligence in Health in 2018: New Opportunities, Challenges, and Practical Implications. PMID: 31419815
5. Public Health, Population Health, and Epidemiology Informatics: Recent Research and Trends in the United States. PMID: 29063572
6. State of the art in clinical informatics: evidence and examples. PMID: 23974543
7. On the foundation and structure of medical informatics. PMID: 8591423
8. An overview of medical informatics education in China. PMID: 22704233
9. Publication trends in the medical informatics literature: 20 years of "Medical Informatics" in MeSH. PMID: 19159472
10. Trends and characteristics of global medical informatics conferences from 2007 to 2017: A bibliometric comparison of conference publications from Chinese, American, European and the Global Conferences. PMID: 30415715

Additional Terms

Humans

MeSH on Demand

MeSH on Demand identifies MeSH® terms in your submitted text (abstract or manuscript). MeSH on Demand also lists PubMed similar articles relevant to your submitted text.

Search | Reset | Help/FAQ | Features

Is it correct to argue for health informatics research to be focussed on greatest burden of disease or years of life lost A counter argument would be for it to be focussed on where it can have most impact The disease based approach cuts across key issues such as a patient focus or support of preventive health and salutogenic lifestyle applications Integrated care across multi morbidity is increasingly recognised as important but would not score highly in the approach taken in this paper Disease based linkage will also tend to bias towards biophysical and medical topics Means of treatment delivery and the people based contributions of professions such as nursing and therapists will not be identified strongly but are important for holistic person centred care and good patient outcomes to a large extent they are disease independent

Start PubMed Search

Export Data

MeSH Terms

- Humans
- Cost of Illness
- Medical Informatics
- Self Care
- Preventive Health Services
- Allied Health Personnel
- Morbidity
- Life Style
- Delivery of Health Care, Integrated

PubMed/MEDLINE Similar Articles

The following articles are 10 similar PubMed Related Citations that were also used in computing these MeSH recommendations. The order is from most to least relevant. Selecting any of the titles opens a new window or tab with that related citation in PubMed's Abstract view.

1. Instantiating informatics in nursing practice for integrated patient centred holistic models of care: a discussion paper. PMID: 26890201
2. Sliding doors: Did drama-based inter-professional education improve the tensions round person-centred nursing and social care delivery for people with dementia: A mixed method exploratory study. PMID: 28088047
3. Development of a framework for person-centred nursing. PMID: 17078823
4. Person-centred web-based support--development through a Swedish multi-case study. PMID: 24139057

Pamela A. Hussey PhD MSc MEd RN RCN, Margaret Ann Kennedy PhD RN CPHIMS-CA

First published: 18 February 2016 | https://doi.org/10.1111/jan.12927 | Citations: 6

Abstract

Aim

A discussion on how informatics knowledge and competencies can enable nursing to instantiate transition to integrated models of care.

Background

Costs of traditional models of care are no longer sustainable consequent to the spiralling incidence and costs of chronic illness. The international community looks towards technology-enabled solutions to support a shift towards integrated patient-centred models of care.

译

Design

Discussion paper.

Data sources

A search of the literature was performed dating from 2000–2015 and a purposeful data sample based on relevance to building the discussion was included.

概要

一、文献海洋中，如何开展精准查询法和引文跟踪法？

几个习惯都要有：

- 浏览 vs. 查询；
- 被动接收 vs. 主动检索；
- 文档查询 vs. 文本挖掘

二、大数据时代，如何开展全域信息获取？

- 超越文献，定位前端的基金资助、后端的专利技术等。

J Am Med Inform Assoc. 2020 Jun 1;27(6):972-975. dc

Is it time for computable evid

Adam G Dunn [1 2 3], Florence T Bourgeois [3 4]

Affiliations + expand

PMID: 32337600 PMCID: PMC7309243 DOI: 10.1093/

Free PMC article

Abstract

Efforts aimed at increasing the pace of evidence synthes
published articles, but these are a relatively delayed, inc
results data. Compared to those in bibliographic databa
registries may be more timely, complete, and accessible
advantages of using structured results data include the
accumulation of relevant evidence and use it to signal w
well as to prospectively assign trials to already publishe
of structured trial data may provide the impetus to build
continuous evidence surveillance.

Cited by 2 articles

The Evolution of Clinical Knowledge During COVID-19: Towards a Global Learning Health System.

Verspoor K.

Yearb Med Inform. 2021 Aug;30(1):176-184. doi: 10.1055/s-0041-1726503. Epub 2021 Sep 3.

PMID: 34479389 Free PMC article.

It is time for computable evidence synthesis: The COVID-19 Knowledge Accelerator initiative.

Alper BS, Richardson JE, Lehmann HP, Subbian V.

J Am Med Inform Assoc. 2020 Aug 1;27(8):1338-1339. doi: 10.1093/jamia/ocaa114.

PMID: 32442263 Free PMC article. No abstract available.

Publication types

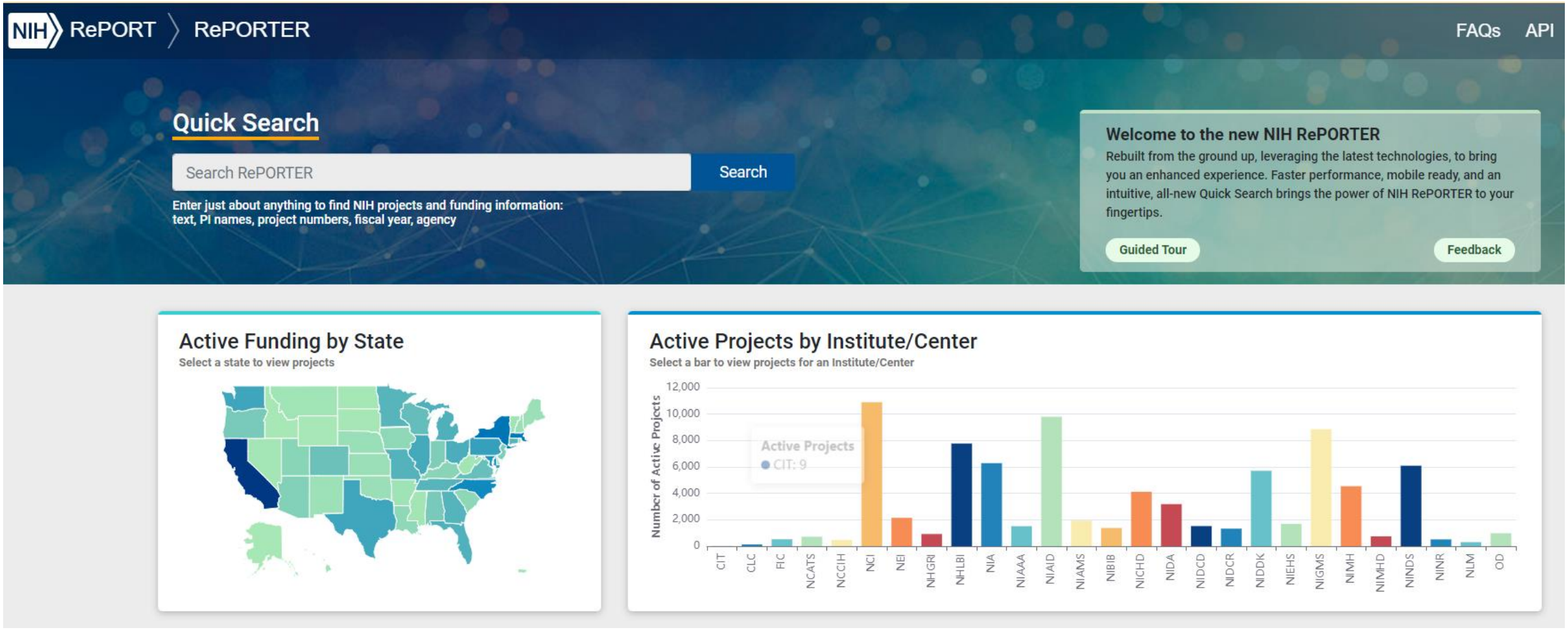
> Research Support, N.I.H., Extramural

MeSH terms

> Clinical Trials as Topic*
> Databases as Topic
> Information Storage and Retrieval / methods*
> Systematic Reviews as Topic*

Grant support

R01 LM012976/LM/NLM NIH HHS/United States



https://reporter.nih.gov/

Advanced Projects Search

Fiscal Year ?

Active Projects

Current FY is 2021

Principal Investigator (PI) ?

PI Names or Profile IDs, semicolon ";" separated

Organization ?

Enter at least 3 characters to search

Agency/Institute/Center ?

Admin Funding

Project Number/Application ID ? 1 ×

R01 LM012976

Format: 5R01CA012345-04/ 8515397, semicolon ";" separated

Reset Search

Looking for additional search fields? Click here to view All Search Fields

Matchmaker

Find potential Program Officials, ICs, and review panels for your research.

Get Started >

Publications Search

Find publications associated with extramural or intramural funded projects using PubMed IDs (PMID) or PubMed Central IDs (PMC ID).

Search Results

1 Project

Save... Edit Search Criteria Export Share

Search Criteria **Fiscal Year:** Active Projects **Project Number/Application ID:** R01 LM012976

Projects | Publications | Patents | Clinical Studies | News & More

Results are sorted by **Relevance**

List Charts Map

Filters

- Fiscal Years
- Org Names
- Agencies
- States
- Countries
- Principal Investigators

T	Act	Project	Year	Sub	Principal Investigator(s)/ Project Leader(s)	Organization	Fiscal Year	Admin IC	Funding IC	FY Total Cost by IC	Similar Projects
Coupling Results Data from ClinicalTrials.gov and Bibliographic Databases to Accelerate Evidence Synthesis											
5		R01LM012976-03			BOURGEOIS, FLORENCE	BOSTON CHILDREN'S HOSPITAL	2021	NLM	NLM	$328,000	View >

Back to Search Results

Description
Details
Sub-Projects
Publications
Patents
Outcomes
Clinical Studies
News and More
History
Similar Projects

Coupling Results Data from ClinicalTrials.gov and Bibliographic Databases to Accelerate Evidence Synthesis

Project Number
5R01LM012976-03

Contact PI/Project Leader
BOURGEOIS, FLORENCE

Awardee Organization
BOSTON CHILDREN'S HOSPITAL

Description

Abstract Text

Project Summary Clinical trials are foundational to evidence-based medicine, but results reporting from trials is incomplete and frequently delayed. It is estimated that as many as half of clinical trials are not published and as many as half of published trials underreport or misreport outcomes. This type of results reporting distorts the evidence available to clinicians—particularly when it comes to assessing the safety of interventions like drugs and devices—and may place patients at unnecessary risk. There is a critical need for novel methods to identify and monitor drug safety data. Through the infrastructure provided by ClinicalTrials.gov, structured trial results (including safety findings) are now becoming available for an increasing number of trials in a comprehensive and timely fashion. However, access and use of these data in evidence synthesis tasks remain limited. ClinicalTrials.gov is the largest single registry for clinical studies worldwide and includes more than 260,000 registered studies. Of the 108,941 completed trials registered with the site, 20% have uploaded results data for a total of 7.85 million participants. Results data reported on ClinicalTrials.gov have the potential to fill gaps created by delays and biases in published articles and provide an earlier and more complete overview of available trial evidence. We propose to develop novel informatics approaches based on combinations of information retrieval and machine learning methods to facilitate access and analysis of trial results reported in this registry. Focusing on trials testing drug interventions in type 2 diabetes, obesity, and oncology, we perform this work in three specific aims: 1) Develop semi-automated trial screening for identifying and aggregating trials relevant to a clinical intervention; 2) Extract adverse event and safety outcomes data from results reported in the registry; and 3) Perform validation studies to assess detection of adverse events and performance of semi- automated meta-analyses of safety outcomes. Methods developed in this project will facilitate timely, broad- scale use of trial results reported on ClinicalTrials.gov in order to augment the availability of comprehensive and timely drug safety data. All methods will be made publicly available in order to support adverse event monitoring and systematic reviews of drug interventions.

History

Total project funding amount for 3 projects is $1,003,250*

Export

** Only NIH, CDC and FDA funding data*

Project Number	Sub	Principal Investigator(s)/ Project Leader(s)	Organization	Fiscal Year	Admin IC	Funding IC	FY Total Cost by IC
Coupling Results Data from ClinicalTrials.gov and Bibliographic Databases to Accelerate Evidence Synthesis							
5R01LM012976-03		BOURGEOIS, FLORENCE	BOSTON CHILDREN'S HOSPITAL	2021	NLM	NLM	$328,000
Coupling Results Data from ClinicalTrials.gov and Bibliographic Databases to Accelerate Evidence Synthesis							
5R01LM012976-02		BOURGEOIS, FLORENCE	BOSTON CHILDREN'S HOSPITAL	2020	NLM	NLM	$328,000
Coupling Results Data from ClinicalTrials.gov and Bibliographic Databases to Accelerate Evidence Synthesis							
1R01LM012976-01		BOURGEOIS, FLORENCE	BOSTON CHILDREN'S HOSPITAL	2019	NLM	NLM	$347,250

Similar Projects

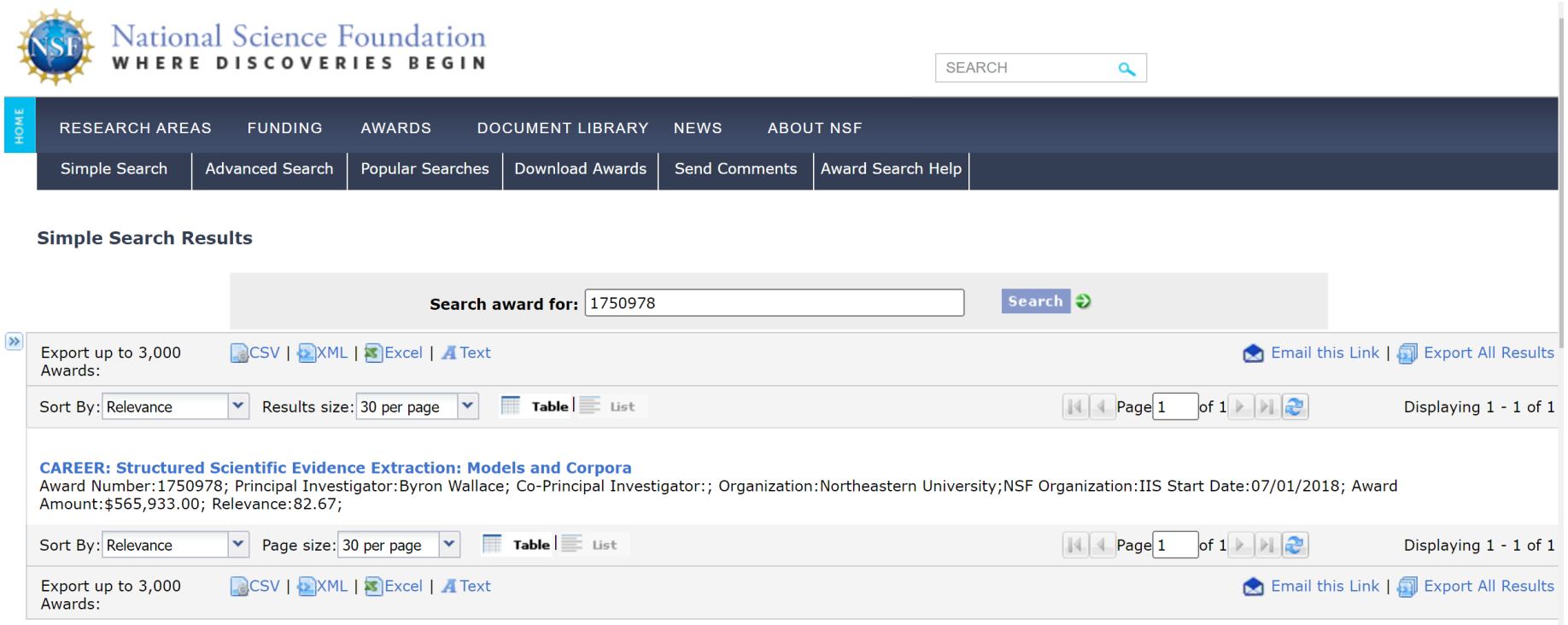
Match Score	Project Number	Sub	Principal Investigator(s)/ Project Leader(s)	Organization	Fiscal Year	Admin IC	Funding IC	FY Total Cost by IC
Coupling Technology with Mind-Body Exercise to Facilitate Physical Activity in Patients with Chronic Cardiopulmonary Disease								
196	5R34AT009354-03		YEH, GLORIA Y	BETH ISRAEL DEACONESS MEDICAL CENTER	2019	NCCIH	NCCIH	$237,717
Clinical safety and efficacy of pharmacogenetics in Veteran care								
215	5IK2CX001262-05		VASSY, JASON L	VA BOSTON HEALTH CARE SYSTEM	2021	VA		
Alpha-synculein/tau coupling in Alzheimer's disease								
190	1R21AG065693-01		LESNE, SYLVAIN E.	UNIVERSITY OF MINNESOTA	2020	NIA	NIA	$422,892
Sub-Clinical Vascular Dysfunction in American-Style Football Players: Temporal Trends, Mechanisms, and Effects on Ventriculo-Arterial Coupling								

基金资助机构

- 论文Acknowledgement信息

Acknowledgements

This work was funded in part by the National Institutes of Health (NIH) under the National Library of Medicine (NLM) grant 2R01LM012086, and by the National Science Foundation (NSF) CAREER award 1750978.



https://www.nsf.gov/

追踪某个PI的科研项目和产出

National Science Foundation
WHERE DISCOVERIES BEGIN

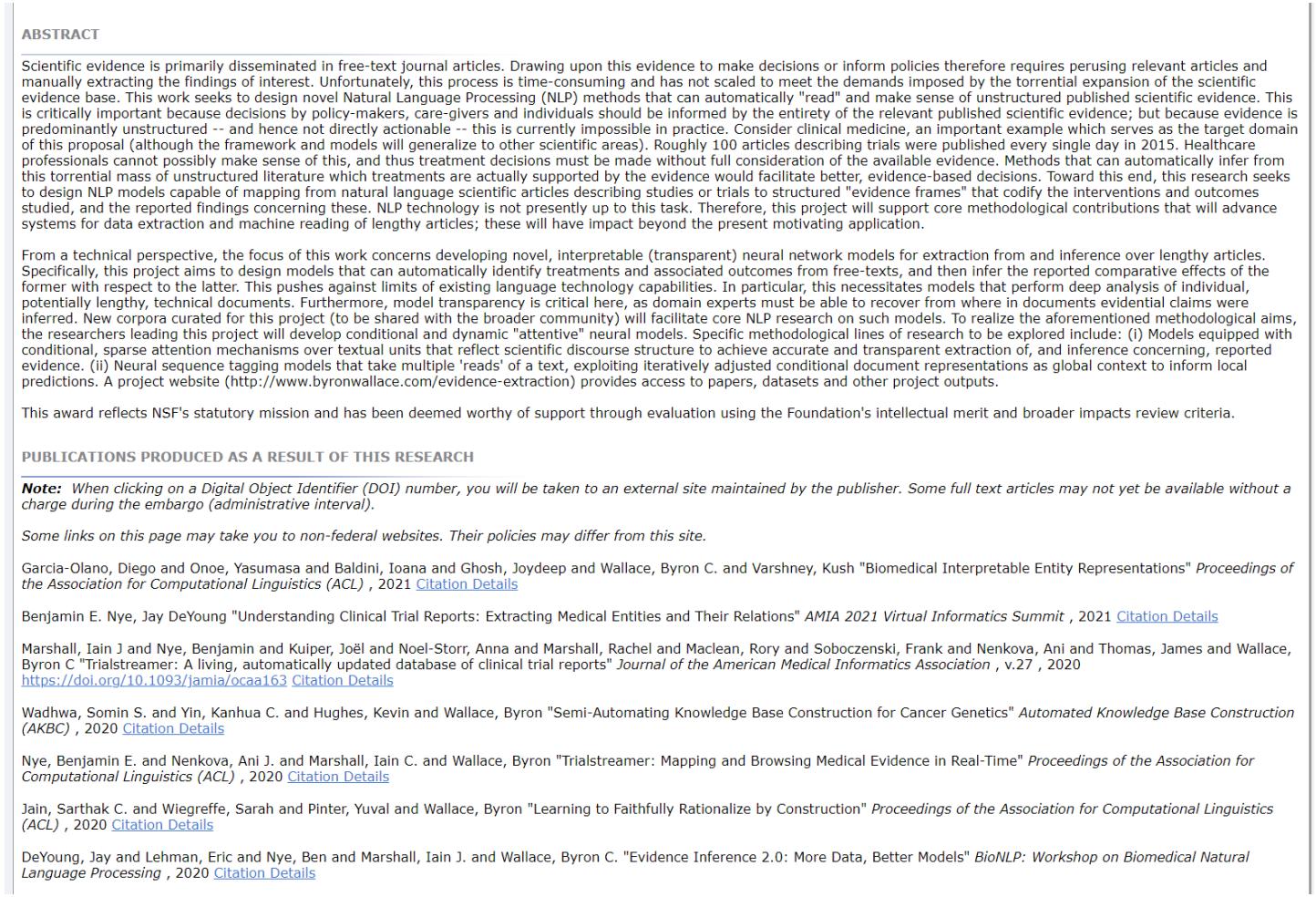
HOME | RESEARCH AREAS | FUNDING | AWARDS | DOCUMENT LIBRARY | NEWS | ABOUT NSF

Awards

- Search Awards
- Recent Awards
- Presidential and Honorary Awards
- About Awards
- How to Manage Your Award
- Grant General Conditions
- Cooperative Agreement Conditions
- Special Conditions
- Federal Demonstration Partnership
- Policy Office Website

Award Abstract # 1750978
CAREER: Structured Scientific Evidence Extraction: Models and Corpora

NSF Org:	IIS Div Of Information & Intelligent Systems
Awardee:	NORTHEASTERN UNIVERSITY
Initial Amendment Date:	February 5, 2018
Latest Amendment Date:	June 1, 2020
Award Number:	1750978
Award Instrument:	Continuing Grant
Program Manager:	Hector Munoz-Avila hmunoz@nsf.gov (703)292-4481 IIS Div Of Information & Intelligent Systems CSE Direct For Computer & Info Scie & Enginr
Start Date:	July 1, 2018
End Date:	June 30, 2023 (Estimated)
Total Intended Award Amount:	¥549,933.00
Total Awarded Amount to Date:	$565,933.00
Funds Obligated to Date:	FY 2018 = ¥115,343.00 FY 2019 = ¥132,659.00 FY 2020 = ¥317,931.00
History of Investigator:	Byron Wallace (Principal Investigator) b.wallace@northeastern.edu
	Northeastern University 360 HUNTINGTON AVE



ABSTRACT

Scientific evidence is primarily disseminated in free-text journal articles. Drawing upon this evidence to make decisions or inform policies therefore requires perusing relevant articles and manually extracting the findings of interest. Unfortunately, this process is time-consuming and has not scaled to meet the demands imposed by the torrential expansion of the scientific evidence base. This work seeks to design novel Natural Language Processing (NLP) methods that can automatically "read" and make sense of unstructured published scientific evidence. This is critically important because decisions by policy-makers, care-givers and individuals should be informed by the entirety of the relevant published scientific evidence; but because evidence is predominantly unstructured -- and hence not directly actionable -- this is currently impossible in practice. Consider clinical medicine, an important example which serves as the target domain of this proposal (although the framework and models will generalize to other scientific areas). Roughly 100 articles describing trials were published every single day in 2015. Healthcare professionals cannot possibly make sense of this, and thus treatment decisions must be made without full consideration of the available evidence. Methods that can automatically infer from this torrential mass of unstructured literature which treatments are actually supported by the evidence would facilitate better, evidence-based decisions. Toward this end, this research seeks to design NLP models capable of mapping from natural language scientific articles describing studies or trials to structured "evidence frames" that codify the interventions and outcomes studied, and the reported findings concerning these. NLP technology is not presently up to this task. Therefore, this project will support core methodological contributions that will advance systems for data extraction and machine reading of lengthy articles; these will have impact beyond the present motivating application.

From a technical perspective, the focus of this work concerns developing novel, interpretable (transparent) neural network models for extraction from and inference over lengthy articles. Specifically, this project aims to design models that can automatically identify treatments and associated outcomes from free-texts, and then infer the reported comparative effects of the former with respect to the latter. This pushes against limits of existing language technology capabilities. In particular, this necessitates models that perform deep analysis of individual, potentially lengthy, technical documents. Furthermore, model transparency is critical here, as domain experts must be able to recover from where in documents evidential claims were inferred. New corpora curated for this project (to be shared with the broader community) will facilitate core NLP research on such models. To realize the aforementioned methodological aims, the researchers leading this project will develop conditional and dynamic "attentive" neural models. Specific methodological lines of research to be explored include: (i) Models equipped with conditional, sparse attention mechanisms over textual units that reflect scientific discourse structure to achieve accurate and transparent extraction of, and inference concerning, reported evidence. (ii) Neural sequence tagging models that take multiple 'reads' of a text, exploiting iteratively adjusted conditional document representations as global context to inform local predictions. A project website (http://www.byronwallace.com/evidence-extraction) provides access to papers, datasets and other project outputs.

This award reflects NSF's statutory mission and has been deemed worthy of support through evaluation using the Foundation's intellectual merit and broader impacts review criteria.

PUBLICATIONS PRODUCED AS A RESULT OF THIS RESEARCH

Note: *When clicking on a Digital Object Identifier (DOI) number, you will be taken to an external site maintained by the publisher. Some full text articles may not yet be available without a charge during the embargo (administrative interval).*

Some links on this page may take you to non-federal websites. Their policies may differ from this site.

Garcia-Olano, Diego and Onoe, Yasumasa and Baldini, Ioana and Ghosh, Joydeep and Wallace, Byron C. and Varshney, Kush "Biomedical Interpretable Entity Representations" *Proceedings of the Association for Computational Linguistics (ACL)* , 2021 Citation Details

Benjamin E. Nye, Jay DeYoung "Understanding Clinical Trial Reports: Extracting Medical Entities and Their Relations" *AMIA 2021 Virtual Informatics Summit* , 2021 Citation Details

Marshall, Iain J and Nye, Benjamin and Kuiper, Joël and Noel-Storr, Anna and Marshall, Rachel and Maclean, Rory and Soboczenski, Frank and Nenkova, Ani and Thomas, James and Wallace, Byron C "Trialstreamer: A living, automatically updated database of clinical trial reports" *Journal of the American Medical Informatics Association* , v.27 , 2020 https://doi.org/10.1093/jamia/ocaa163 Citation Details

Wadhwa, Somin S. and Yin, Kanhua C. and Hughes, Kevin and Wallace, Byron "Semi-Automating Knowledge Base Construction for Cancer Genetics" *Automated Knowledge Base Construction (AKBC)* , 2020 Citation Details

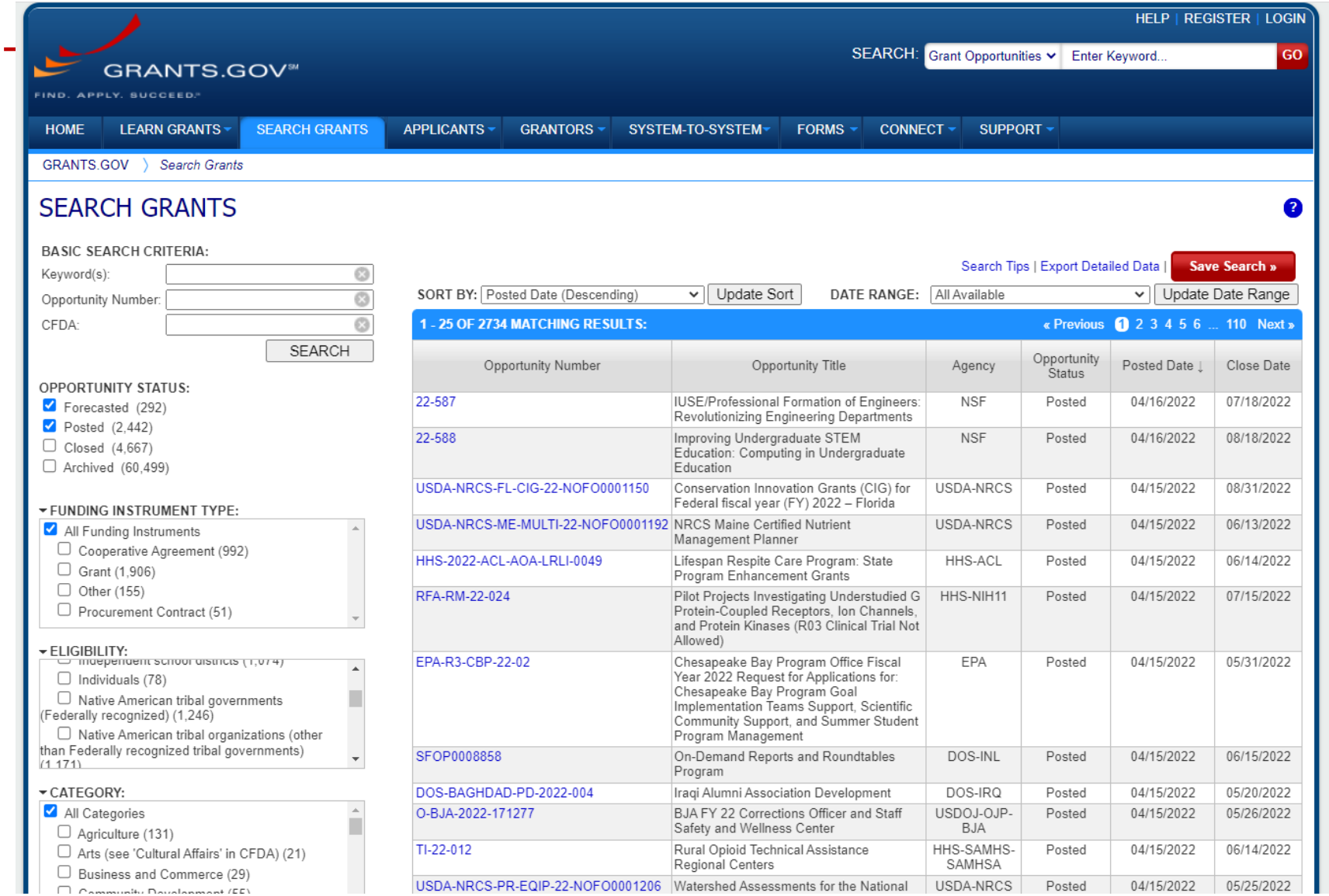
Nye, Benjamin E. and Nenkova, Ani J. and Marshall, Iain C. and Wallace, Byron "Trialstreamer: Mapping and Browsing Medical Evidence in Real-Time" *Proceedings of the Association for Computational Linguistics (ACL)* , 2020 Citation Details

Jain, Sarthak C. and Wiegreffe, Sarah and Pinter, Yuval and Wallace, Byron "Learning to Faithfully Rationalize by Construction" *Proceedings of the Association for Computational Linguistics (ACL)* , 2020 Citation Details

DeYoung, Jay and Lehman, Eric and Nye, Ben and Marshall, Iain J. and Wallace, Byron C. "Evidence Inference 2.0: More Data, Better Models" *BioNLP: Workshop on Biomedical Natural Language Processing* , 2020 Citation Details

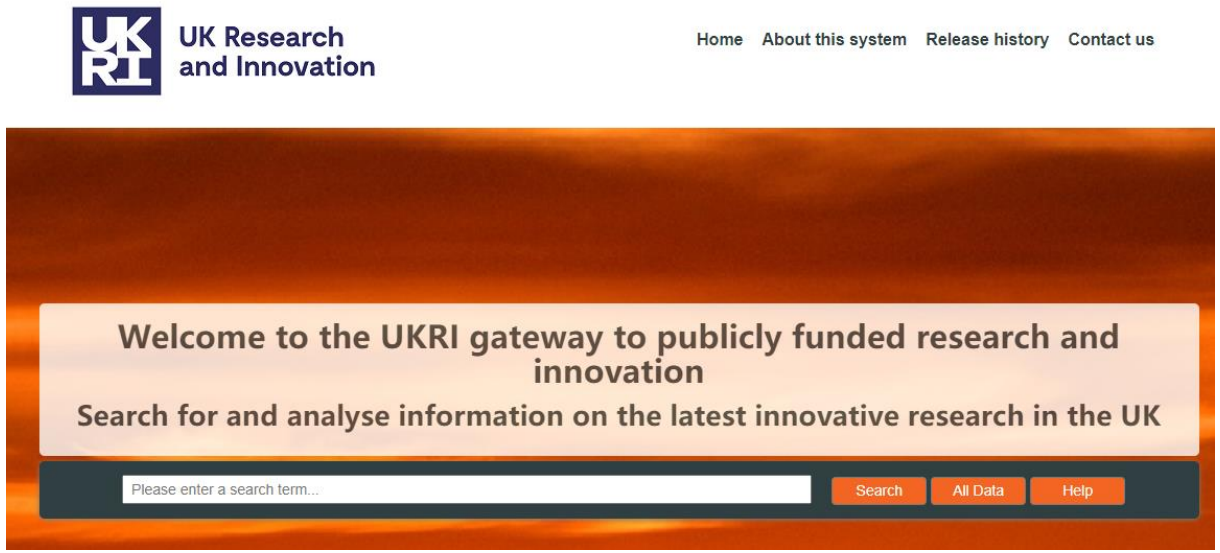
https://www.nsf.gov/awardsearch/showAward?AWD_ID=1750978&HistoricalAwards=false

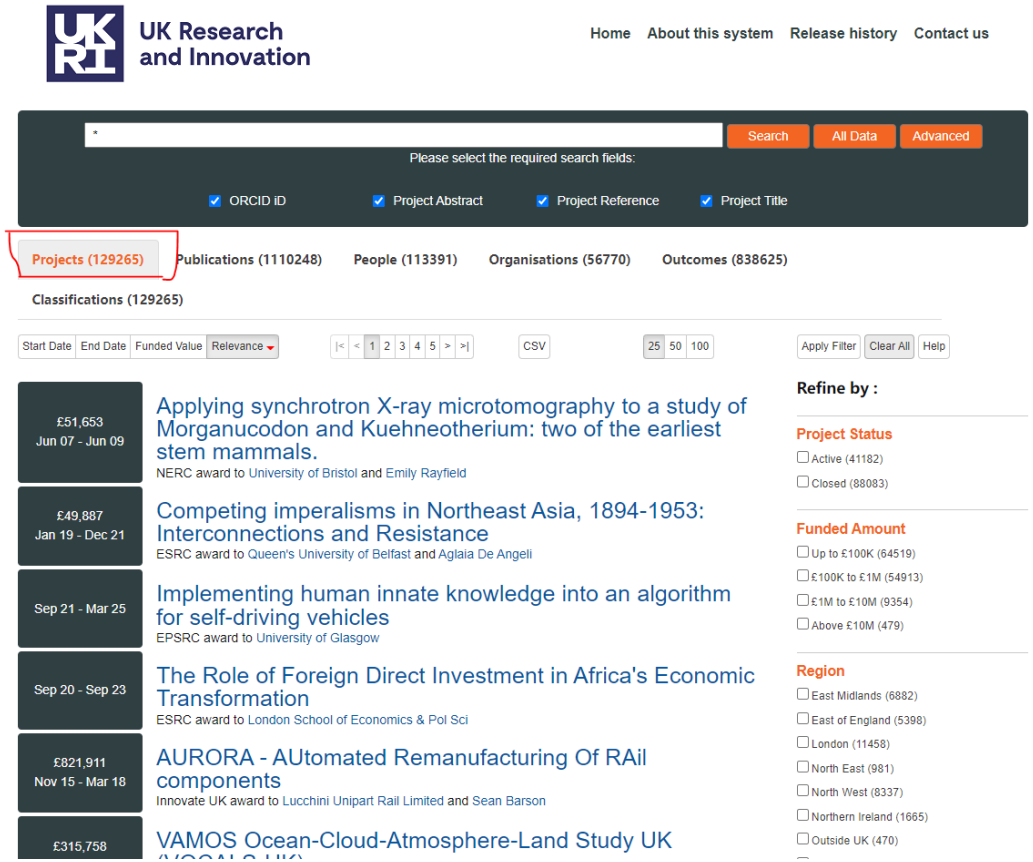
https://www.grants.gov/web/grants/search-grants.html



英国基金资助机构

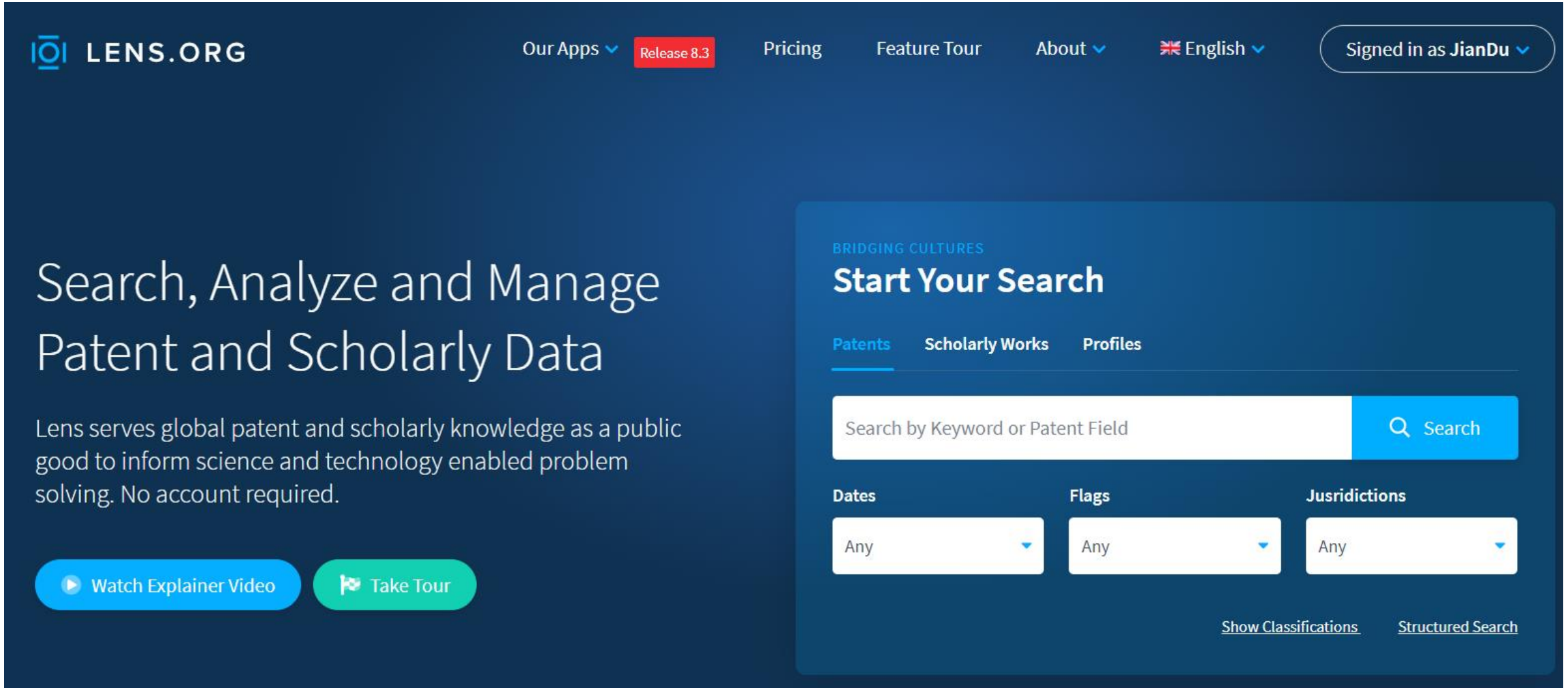
- 英国成立统一的科研资助与项目管理机构





https://gtr.ukri.org/

连接科学论文和发明专利：科学→技术



https://www.lens.org/

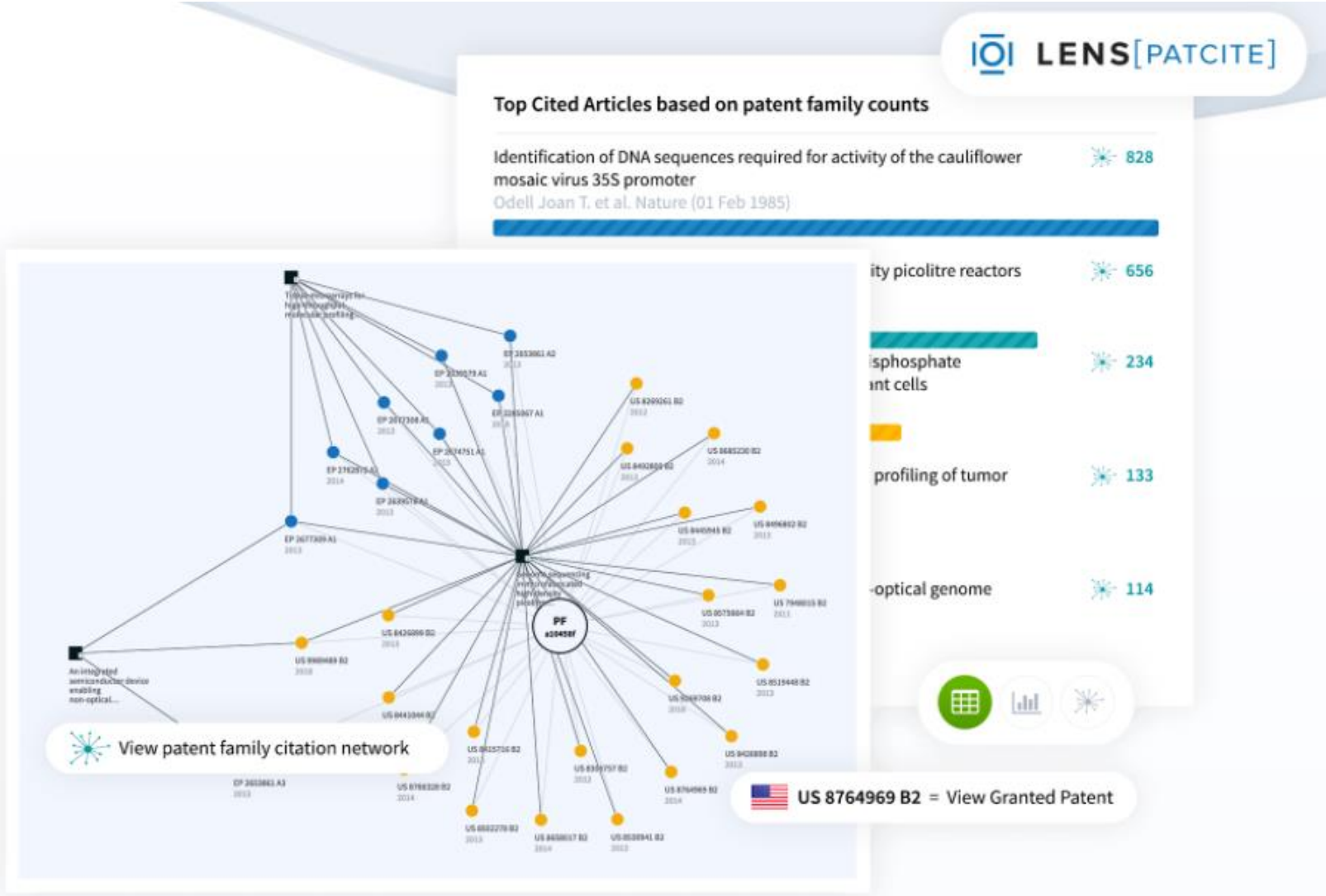
发现**精准合作伙伴**

发现哪些学术文章影响了哪些专利以及谁在使用您的学术作品。

这些地图动态且易于查询，可让您识别重要的联系，建立合作网络。该工具的粒度使您能够实时了解科学和学术如何塑造基于专利的发明，以及哪些研究文章、哪些科学家或研究人员，以及哪些机构可能对部分经济活动产生影响。

可供个别研究人员/发明人、大学部门、机构甚至资助组织使用的动态地图。

浏览下面的示例投资组合



https://www.lens.org/lens/patcite

Format: Abstract

Send to

Nat Biotechnol. 2014 Jun;32(6):569-76. doi: 10.1038/nbt.2908. Epub 2014 Apr 25.

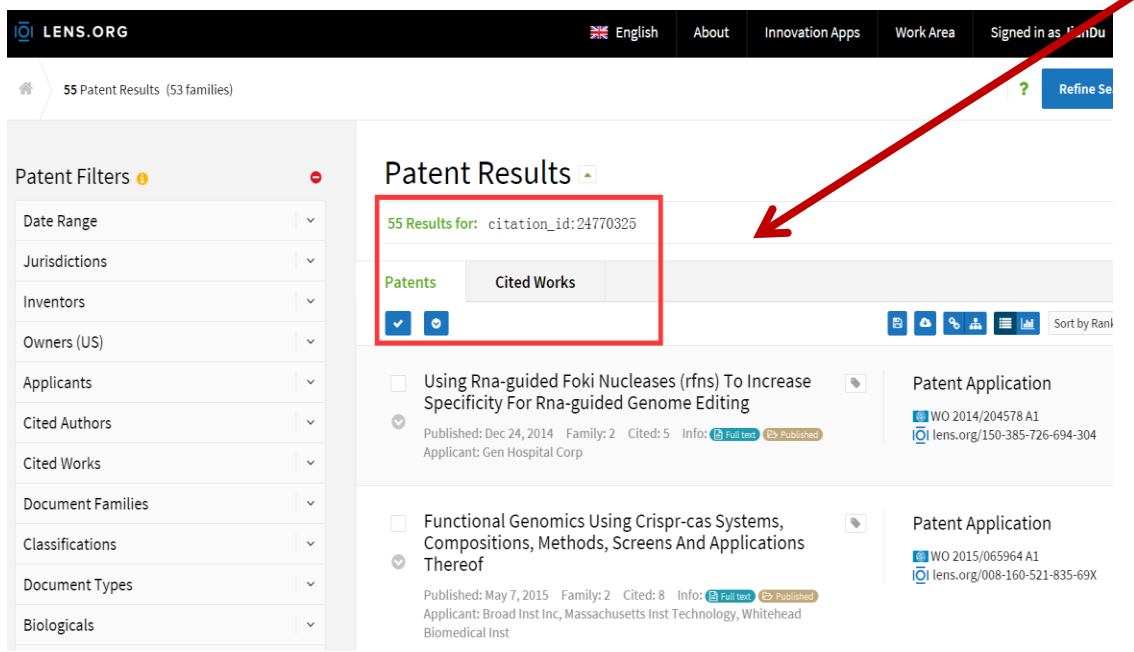
Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing.

Tsai SQ[1], Wyvekens N[2], Khayter C[2], Foden JA[2], Thapar V[3], Reyon D[1], Goodwin MJ[2], Aryee MJ[4], Joung JK[1].

Author information

Abstract

Monomeric CRISPR-Cas9 nucleases are widely used for targeted genome editing but can induce unwanted off-target mutations with high frequencies. Here we describe dimeric RNA-guided FokI nucleases (RFNs) that can recognize extended sequences and edit endogenous genes with high efficiencies in human cells. RFN cleavage activity depends strictly on the binding of two guide RNAs (gRNAs) to DNA with a defined spacing and



论文-专利引用大规模数据链接

https://www.lens.org/

nature biotechnology

Altmetric: 119 More detail »

Patents | OPEN | Published: 10 January 2018

Mapping the global influence of published research on industry and innovation

Osmat A Jefferson, Adam Jaffe, Doug Ashton, Ben Warren, Deniz Koellhofer, Uwe Dulleck, Aaron Ballagh, John Moe, Michael DiCuccio, Karl Ward, Geoff Bilder, Kevin Dolby & Richard A Jefferson

Nature Biotechnology **36**, 31–39 (2018) | Download Citation

一个用于评价学术研究对创新的影响的指标 精选

已有 7898 次阅读 2017-8-11 10:17 | 个人分类:科学计量学 | 系统分类:科研笔记 推荐到群组

一个用于评价学术研究对创新的影响的指标

——基于lens.org平台上专利引用论文的链接数据

杜建

今天早上看到8月10日关于自然指数-创新（Nature Index Innovation）的报道，我一直关注的Lens.org平台首次推出了评价产品。Lens.org是澳大利亚一家独立的非营利性机构-Cambia的服务平台，是Cambia和昆士兰州理工大学联合开发的。他们与美国国立医学图书馆（NLM）和Crossref合作，把全球所有的专利族和它们引用的科技论文的PMID号、DOI号映射好，可以方便地检索一篇论文被专利引用的情况。

我当时发现Lens.org是一次偶然的机会，因为常用PubMed检索，发现每篇论文的Linkout部分，多了一个Len.org的标志。当时恰好正在承担汤森路透和中信所科学计量学联合实验室资助的一项课题"基于科学与技术关联分析测度转化研究：方法与实证"，苦于做专利-论文引用的数据，Lens.org平台为研究提供了很大帮助。

因为这个平台是开放的，我曾担心他们做的专利引用论文的链接否精确，曾致信他们首席科学家Osmat A Jefferson教授，她告诉我，他们会有论文发表，介绍他们的方法论。

Lens.org

Analyze Linkages Between Academic Research And Inventions

We are sure that science and technology influence and enable industry. But how do we map this and measure it?

Influence Mapping provides an evidence base to guide decision-making and enables improved public policies and institutional practices. Using this toolkit you can track, filter, sort, and link scholarly articles that are cited in patents, examine the citing patents, and discover new partners and collaborators.

https://www.lens.org/lens/patcite

Your search query

Feng Zhang's influence on inventions

Created 27 Aug 2019 with
714 citation identifiers

178 QNPL found

Filter your search

● Citing Patents ■ Cited Articles

- Document type
- Date
- Jurisdiction
- Classification
- Applicant
- Owner
- Inventor
- Publication key
- NPL citation
- Family size
- Citations count

Search Results "Feng Zhang's influence on inventions"

Hide Summary | Start a new search

Query articles cited in patent literature (QNPL)

178 ■ Cited articles

535 Unmatched identifiers

Citing patents/families distribution

4,217 ● Patent documents
- 1,668 Grants
- 2,325 Applications
- 224 Others

2,578 ○ Corresponding patent families

Top Cited Articles based on patent family counts

Multiplex Genome Engineering Using CRISPR/Cas Systems ✳ 782
Cong Le et al. Science (New York, N.Y.) (03 Jan 2013)

Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas… ✳ 543
Zetsche Bernd et al. Cell (25 Sep 2015)

RNA-guided editing of bacterial genomes using CRISPR-Cas systems ✳ 375
Jiang Wenyan et al. Nature biotechnology (29 Jan 2013)

DNA targeting specificity of RNA-guided Cas9 nucleases ✳ 347
Hsu Patrick D. et al. Nature biotechnology (21 Jul 2013)

Double nicking by RNA-guided CRISPR Cas9 for enhanced genom… ✳ 339
Ran F. Ann et al. Cell (29 Aug 2013)

Cited articles | Citing patents

Showing 1 to 20 of 2,578 patent families | Number of QNPL

1. Protected guide RNAS (PGRNAS)

Earliest priority date: 12 Dec 2014

Feedback

科学中的未知知识

- 通过文献中自然语言描述文本寻找科学问题?
- 不确定性知识、未知知识?

临床内科杂志 2018 年 11 月第 35 卷第 11 期 J Clin Intern Med,

胃食管反流病与心房颤动关系的研究进展

肖蓉 李波

[关键词] 胃食管反流病; 心房颤动; 食管裂孔疝; 质

心房颤动(简称房颤)是临床上常见的心律失常疾病,其发病率在普通成年人中占 3%,随着老年化的进展患病率仍在升高[1]。据估计,到 2030 年欧洲房颤患者数量将由 1 400 万上升到 1 700 万,房颤使卒中风险增加 5 倍,死亡率增加 2 倍[2]。胃食管反流病(GERD)是消化系统的常见疾病,患病率约为 2.5% ~25.0%,亚洲较西方国家低,随着成人和儿童饮食方式的改变、超重及肥胖人群增多,GERD 逐渐呈现年轻化趋势[3-4]。由于食管和左心房在解剖位置上临近,随着对 GERD 和房颤发病机制研究的深入,近年来发现两者之间可能存在一定关系,GERD 可能促进房颤的发生并且建议将房颤作为 GERD 的食管外症状[5]。在 GERD 患者中使用质子泵抑制剂(PPI)后似乎可以使其房颤症状减轻,目前关于房颤、GERD 和 PPI 之间的关系仍存在争议,本文对 GERD 和房颤关系、GDED 触发房颤的作用机制及 PPI 与房颤的关系进行综述。

四、总结

临床医生应该意识到胃食管疾病与房颤之间的相互关系。鉴别并有效治疗 GERD 特别是食管炎,可能有助于减少房颤发病和症状,目前研究对房颤和 GERD 因果关系仍不确定,PPI 治疗是否减轻房颤仍有争议,还需要各项深入研究一进一步探讨。

“不确定性”在科学知识的表达和交流中起重要的作用

- 表达未知、假设、推测、争议和矛盾
- 不确定性的突然增加可能意味着新范式的出现

- 通过识别出处于争议、矛盾、未知等不确定性状态、悬而未决的知识，对于引导科学研究方向、优化资源配置提供了一个决策支持工具

Table 1. Frequencies of the uncertain cue words.

Cue words	Frequency in all SemMedDB sentences	
Unknown lexicon		
uncertain*	227,014	(10.57‰)
unknown	525,536	(24.48‰)
Hedging lexicon		
maybe	10,286	(0.48‰)
may	5,946,955	(276.96‰)
might	949,536	(44.22‰)
possible	1,751,994	(81.59‰)
potential	2,879,336	(134.10‰)
seems	333,677	(15.54‰)
perhaps	84,058	(3.91‰)
likely	1,052,986	(49.04‰)
sometimes	119,942	(5.59‰)
Conflicting lexicon		
conflict*	175,516	(8.17‰)
contradict*	46,639	(2.17‰)
controvers*	208,264	(9.70‰)
debat*	122,332	(5.70‰)
no consensus	17,907	(0.83‰)
questionable*	21,159	(0.99‰)
refut*	9,710	(0.45‰)

“*” stands for all of the possible derivations of the word.

Xin Guo, Yuming Chen, Jian Du, Erdan Dong. "Extracting and Measuring Uncertain Biomedical Knowledge from Scientific Statements." *Journal of Data and Information Science* (2022).

小结

- 搜索引擎精准检索
- 信息自动推送，提高信息获取效率
- 关键词检索（matching）：local perspective，依赖于标注质量
- 引文检索（Citing）：代际思维、自然而然的秩序、有利于学科交叉
- 文献前后端信息（资助、专利）检索，全域信息获取
- 科学前沿信息：不确定性信息的识别和获取

谢谢大家！

北京大学 健康医疗大数据国家研究院
NATIONAL INSTITUTE OF HEALTH DATA SCIENCE AT PEKING UNIVERSITY