Harnessing Generative AI for Research Analytics and Communicating the Societal Impact of Research

Ann Gabriel, Senior Vice President, Global Strategic Networks, Elsevier
Arthur Ellis, Senior Advisor, Elsevier
October 5, 2023

Elsevier’s Linked Data and Analytics

Largest scholarly publisher, with expertise in machine learning, semantic analysis, scientific taxonomies

Elsevier’s Scopus database:

- 90M peer-reviewed papers and conference proceedings from 7,000 publishers
- 18M author profiles—academia, industry, government
- 100,000 affiliation profiles of research-producing institutions worldwide
- 1.7B cited references
- 130M patent records from 100+ patent offices worldwide

Allows for linking data with high recall and precision

Foundation for Elsevier’s journey toward trusted Generative AI tools for researchers and teachers

Outline

Examples of machine language-linked mapping

- Publications and Patents
- Proposal Concept Outlines and Award Abstracts
- Funding-to-commercialization linkages

Generative AI

- Elsevier’s approach and principles
- Demonstration of *Scopus-AI*

Platforms for Demonstrating Societal Impact of Research: Links to UN Sustainable Development Goals (SDGs)

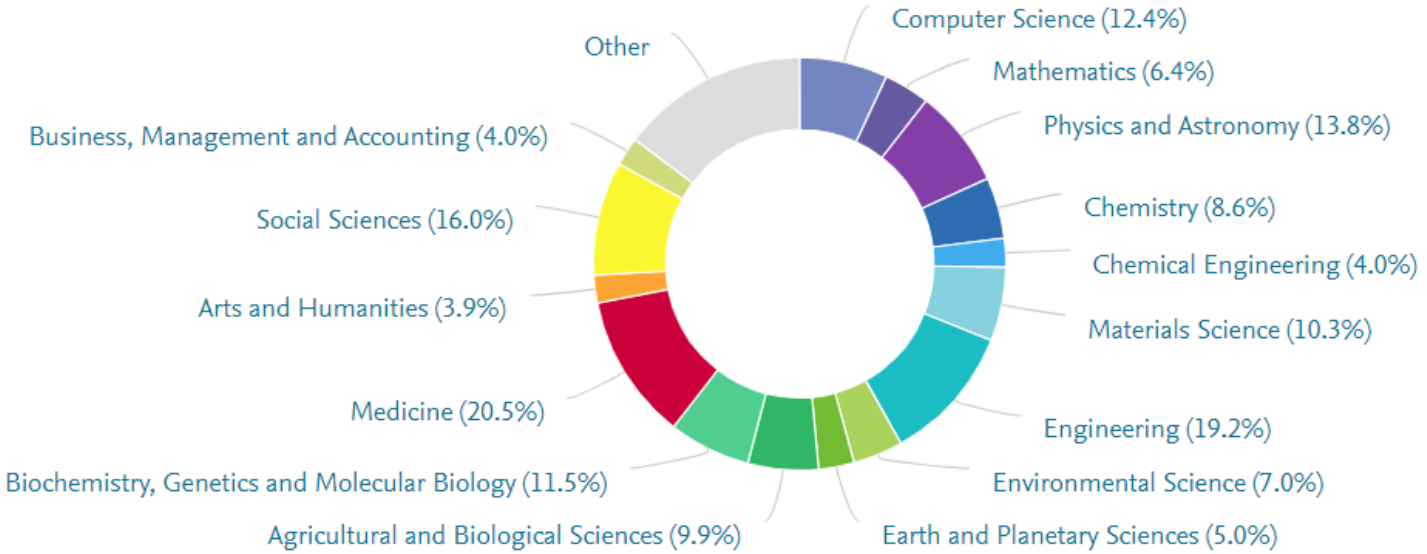
Links are made using extensive keyword queries supplemented with machine learning

https://sdgs.un.org/goals
https://elsevier.digitalcommonsdata.com/datasets/87txkw7khs/1

HBCU Research Overview, 2013-2022

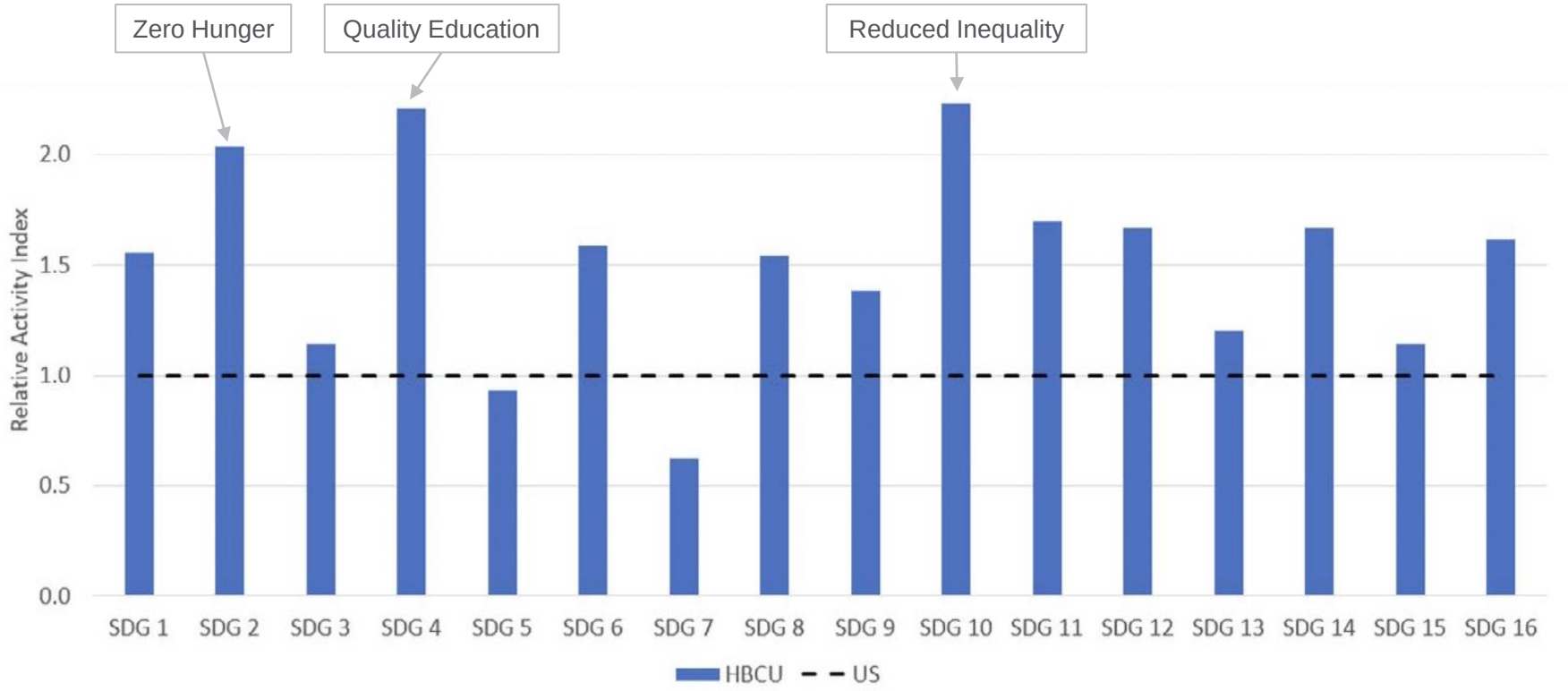
Overall research performance

48,767 ▲	26,008 ▲	1.35
Scholarly Output ⓘ	Authors	Field-Weighted Citation Impact ⓘ



Source: SciVal

Relative Activity Index of HBCU SDG Contributions



Data source: SciVal

CUNY Linkage of Patents to UN SDGs

Sectors: UN Sustainable Development Goals - Goals Sector Size: Portfolio Size

Source: PatentSight

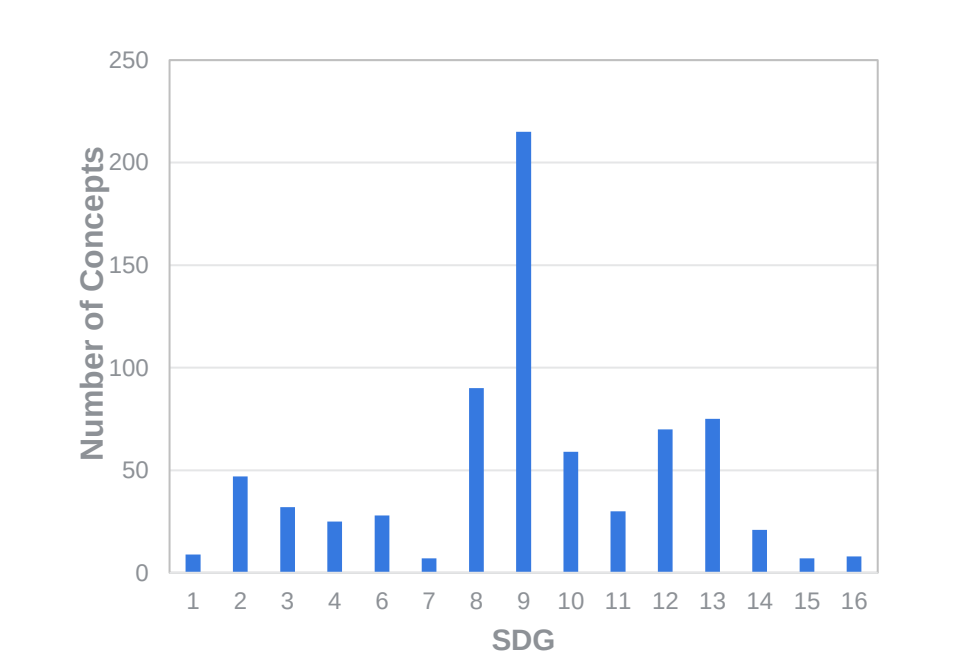
10 Key Technology Areas of the CHIPS and Science Act of 2022

- Artificial Intelligence
- Advanced Computing & Semiconductors
- Quantum Information Science & Technology
- Robotics & Advanced Manufacturing
- Disaster Prevention & Mitigation
- Advanced Communications
- Biotechnology
- Cyberinfrastructure & Cybersecurity
- Advanced Energy & Industrial Efficient Technologies
- Advanced Materials

NSF/TIP website; https://www.congress.gov/bill/117th-congress/house-bill/4346/text

NSF: Technology, Innovation and Partnerships (TIP) Directorate

TIP's New Signature Program: Regional Innovation Engines ("Engines") to advance translational and use-inspired research



"ENGINES" - SUBMITTED PROJECT CONCEPTS MAPPED TO SDGs

Titles, keywords, and brief

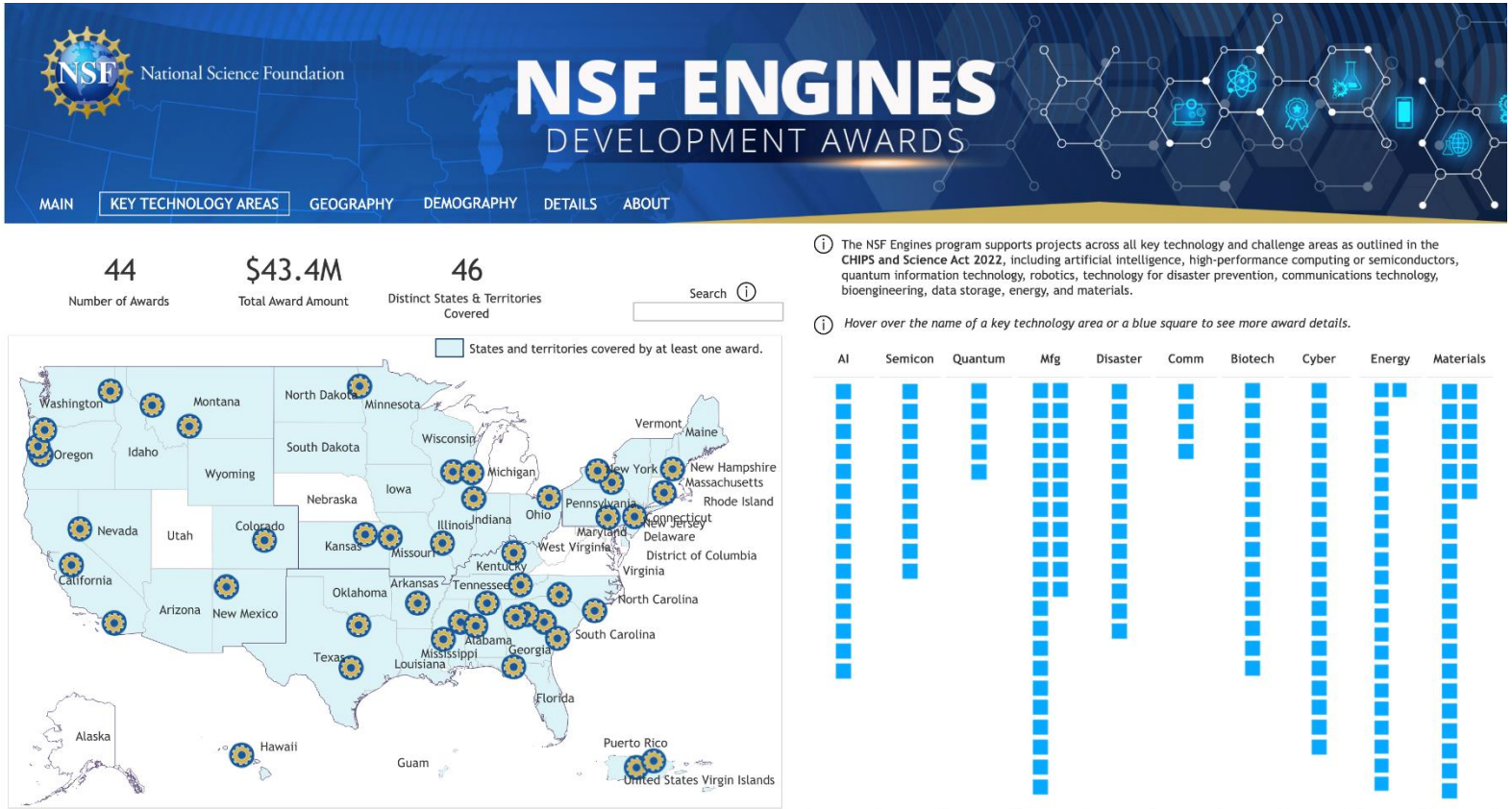
Concept Outlines of publicly submitted

project ideas were used to

link 442 of 679 projects (about 2/3)

to one or more SDGs.

Interactive Map of the Engines Development Awards Portfolio



Source: https://new.nsf.gov/funding/initiatives/regional-innovation-engines

Engines Award 2304059: Advancing Circular Bioeconomy Technologies in the North San Joaquin Valley

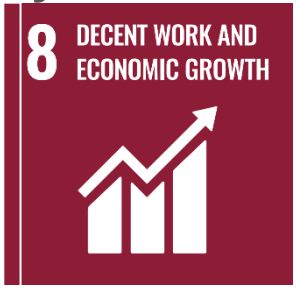
https://www.cbioinnovation.org/

SDG Mapping: Advancing Circular Bioeconomy Technologies in the North San Joaquin Valley

This Regional Innovation Engines Development Award is focused on building bioindustrial manufacturing capability in California's North San Joaquin Valley (NSJV), an integrated economic region consisting of Merced, Stanislaus, and San Joaquin Counties. NSJV is uniquely positioned to lead in bioindustrial manufacturing due to an unparalleled combination of large-scale food and agricultural activity, manufacturing capacity, and proximity to hubs of biotechnology innovation. NSJV is a historically underserved region, with poverty and unemployment rates well above the state average, low relative educational attainment, and significant environmental justice and public health challenges. By building regional specialization in bioindustrial manufacturing, the NSJV can diversify and strengthen its regional economy by moving up the value chain, generate higher-quality and accessible family-sustaining jobs for local residents, advance the national bioeconomy and the transition to a net-zero carbon economy, and promote a more sustainable food and agriculture system that reduces environmental and public health harm to surrounding communities. Biomanufacturing uses bioscience and biotechnology - especially the use of fermentation - to create useful products, including chemicals, materials, and fuels, from organic material. Yet there is a gap in translation of basic discoveries into bioproduction and bioengineering applications with public benefit, particularly translation into scalable solutions that meet the market demands and the scope of global sustainability challenges. There are gaps in areas of innovation related to feedstock diversification, advanced biomanufacturing process optimization, and equipment of the future for producing new bioproducts at scale. At the same time, there is a new opportunity for growing thousands of high-quality jobs that offer economic mobility - including for workers without a four-year degree - in a region where only one-third of current jobs provide pathways to prosperity. With intentional policies and coordination, bioindustrial manufacturing presents the opportunity to decrease racial disparities, address environmental injustice, and build wealth by integrating inclusion into this strategy through workforce development efforts. This Development Award will enable a broad range of regional partners, along with a network of strategic collaborators and advisors, to develop actionable and strategic plans for this Engine.

https://sdg-analyzer.azurewebsites.net/

SDG8:51% SDG9:24% SDG1:7%



51%

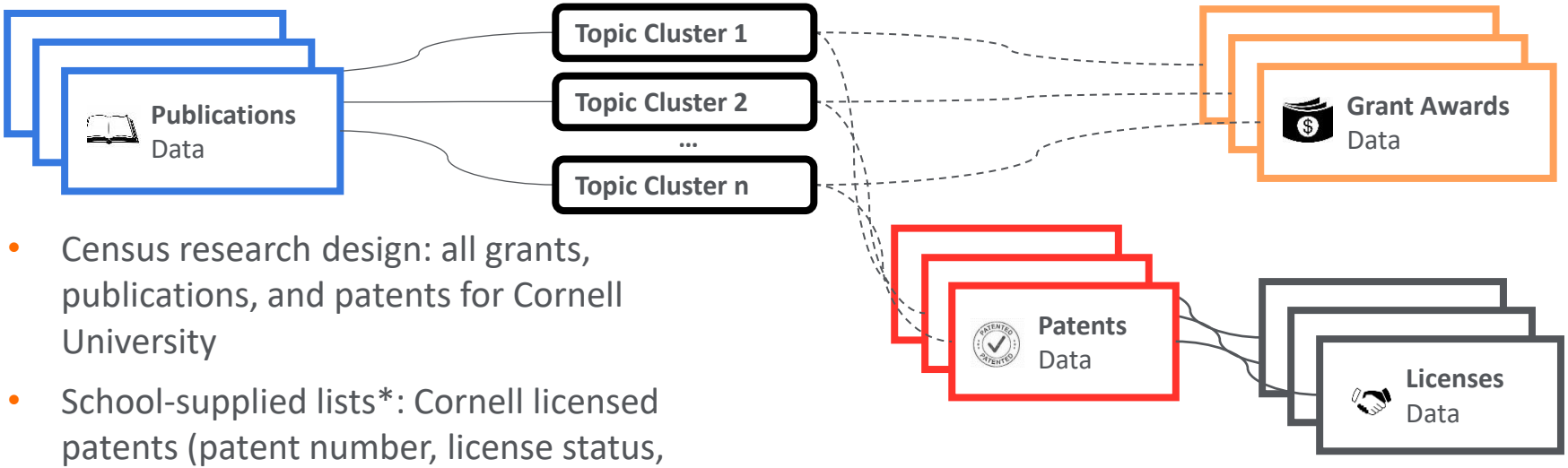
24%

7%

Commercialization Pipelines of Cornell R&D

For Cornell University, what are the paths for research areas to commercialization success?

METHODS



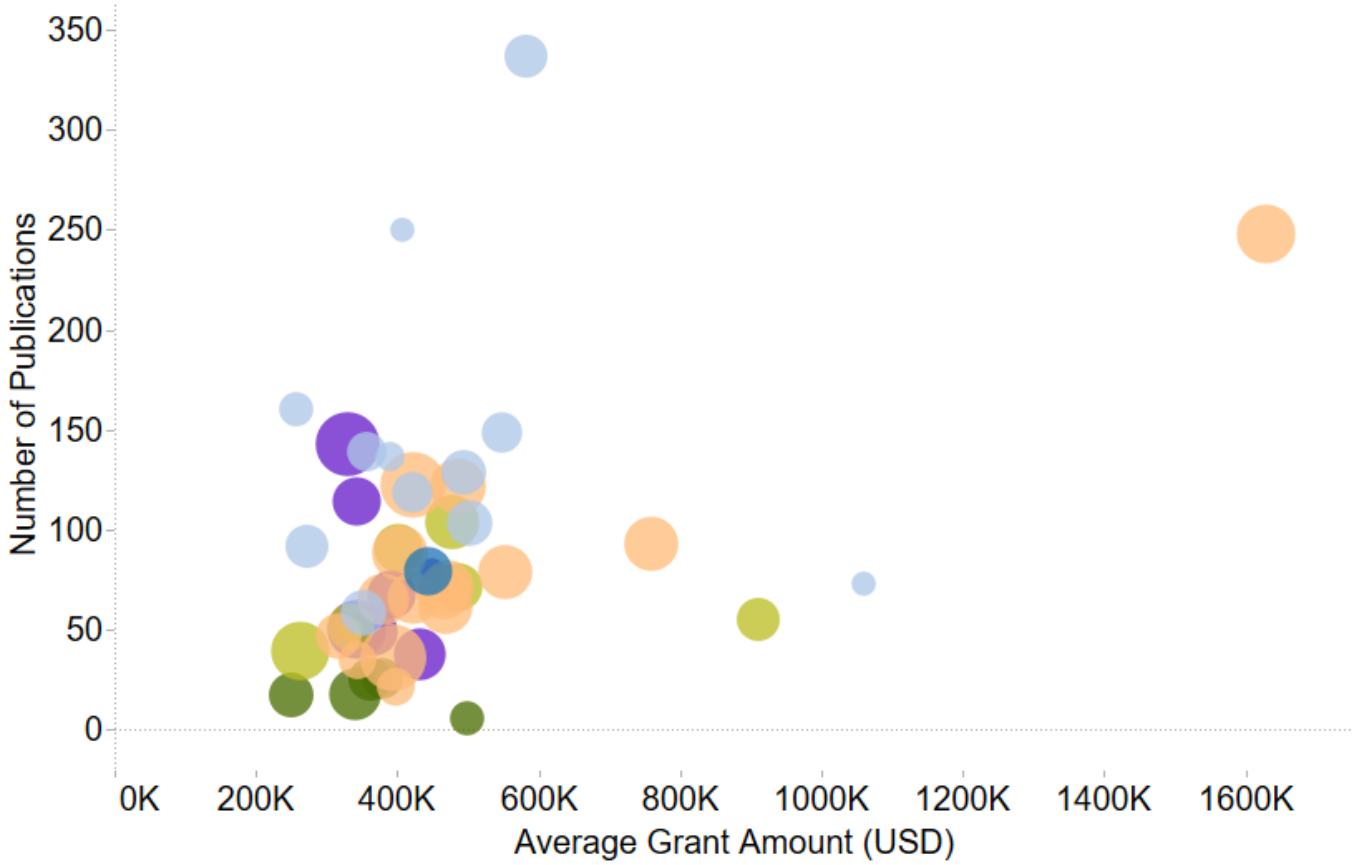
- Census research design: all grants, publications, and patents for Cornell University
- School-supplied lists*: Cornell licensed patents (patent number, license status, licensee small business status, licensee startup status)
- Research topic proxy: STM model for topic-modelling

- Data covers 11,812 grants, 62,674 publications, 1,693 patent families, and 1,185 licenses

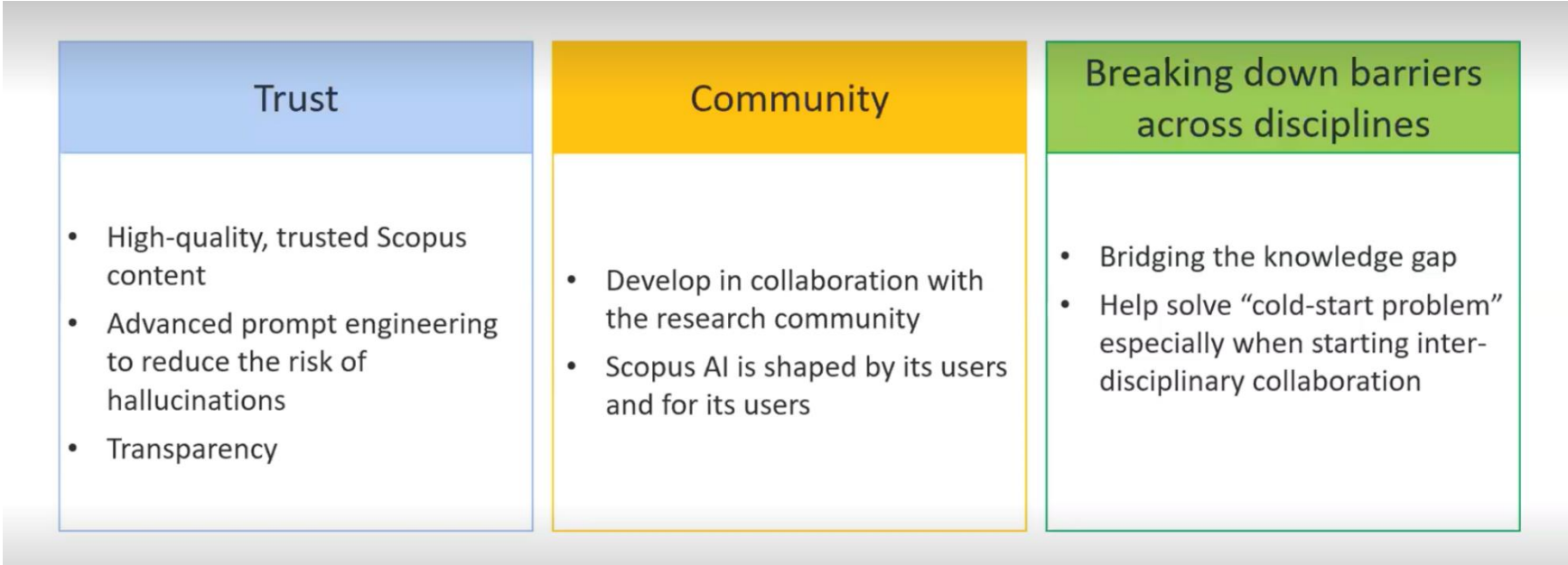
* Ethics approval was obtained from the Cornell Institutional Review Board before proceeding with data collection

Methodological Value for Ex-ante Analysis

RESULTS



Elsevier's Generative AI Development Principles



Scopus AI currently uses OpenAI's ChatGPT model hosted on Microsoft's Azure
Powered by generative AI and other Large Language Model technology with Scopus search capabilities
Frontend: Mix of Javascript and CSS
Backend: Python, Java, Elasticsearch, and Langchain
No exchange of data with public ChatGPT

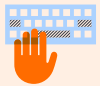
Elsevier's Responsible AI Principles

1. We consider the real-world impact of our solutions on people.
2. We take action to prevent the creation or reinforcement of unfair bias.
3. We can explain how our solutions work.
4. We create accountability through human oversight.
5. We respect privacy and champion robust data governance.

https://beta.elsevier.com/about/policies-and-standards/responsible-ai-principles?trial=true

Scopus - AI: How does it work?

Step 1:
Curation of high-quality Scopus content



Step 2:
Query formulation

Step 3:
Scopus AI: Vector Search & Result Generation

Step 4:
Language Learning Model (LLM) summary generation

Step 5:
Cited references for validation and transparency

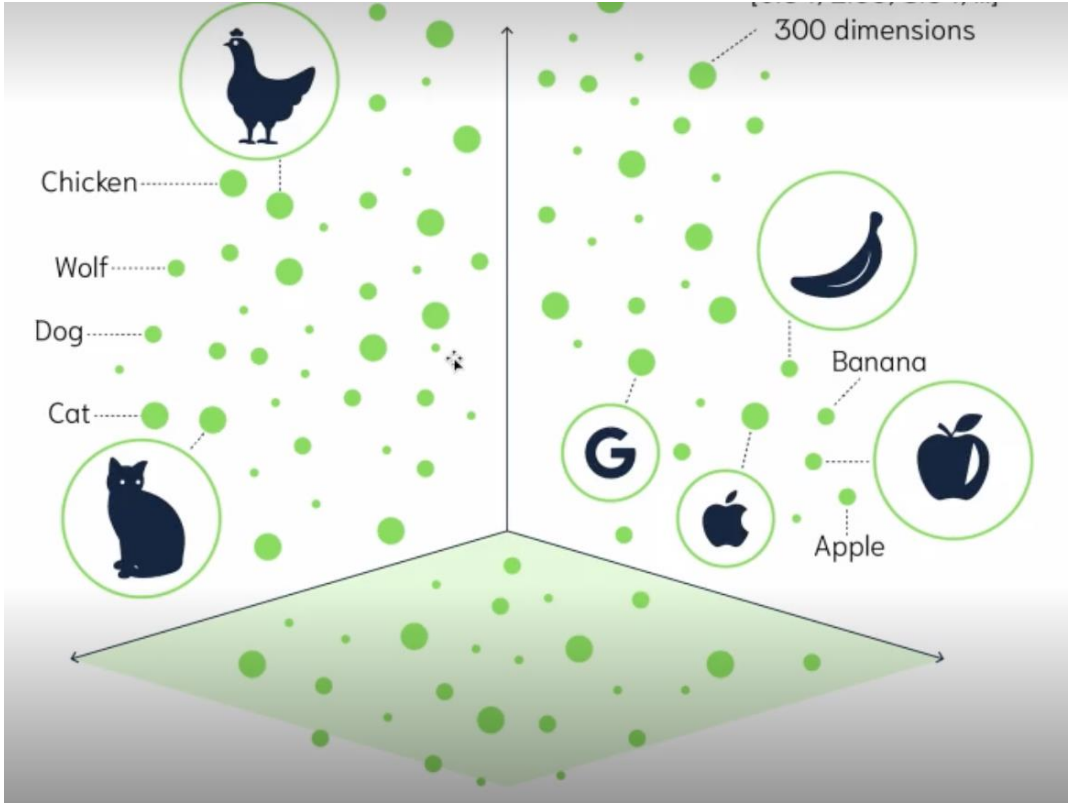
Step 6:
“Go deeper” and “Keyword graph” development

Vector Search

“How to be a better human” becomes, e.g., [100, -200, 5]

Vector Search: 22M Scopus abstracts, 2018-2022

Accommodates semantic, contextual and similarity searches



Scopus AI: What does it deliver?

Natural language queries

Ease your search: Researchers can ask questions about a subject in a natural, conversational manner.

Visual representation of entities

See the big picture: Our tool visually maps search results, offering a comprehensive overview that allows researchers to navigate complex relationships easily.

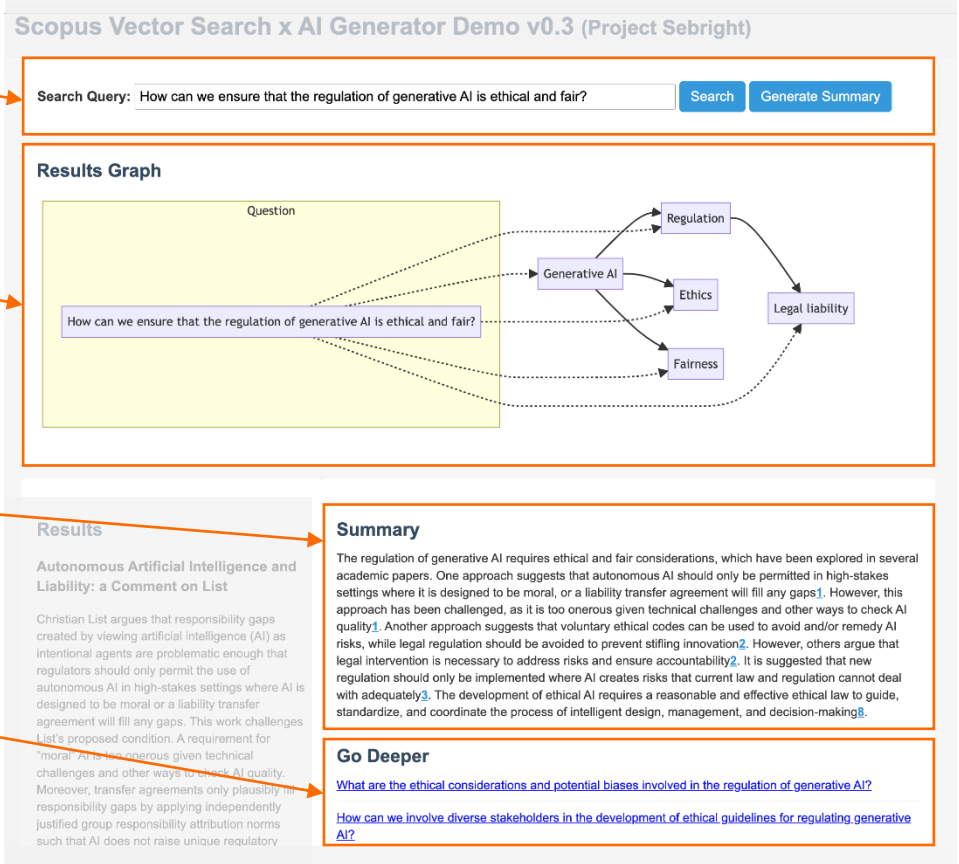
Summary with Scopus references

Instant overview: Skip the lengthy reading. Scopus AI gives you a concise and trustworthy summary with academic references for each search.

Deeper query exploration

Uncover more: Explore beyond the surface. Our AI offers relevant queries for further exploration, leading to hidden insights in various research fields.

*Interface and functionality subject to change



Display

Summary

Sol-gel dip-coating is a technique used to prepare titanium dioxide (TiO2) thin films. The sol-gel method offers benefits such as reliability, reproducibility, and controllability in producing high-quality nanostructured thin films [2]. TiO2 thin films can be prepared using various deposition techniques, including dip-coating, spin coating, spray pyrolysis, and electrophoretic sol-gel deposition [3]. These TiO2 thin films exhibit properties such as chemical stability, reusability, hydrophilicity, and activation by sunlight or simulated solar radiation, making them suitable for commercial devices that keep public surfaces clean [3]. The sol-gel dip-coating method allows for the formation of TiO2 thin films with controlled characteristics, such as surface morphology and thickness [9]. The films can be calcined at high temperatures to enhance their properties, such as adhesion and photodegradation [7]. Overall, sol-gel dip-coating is a versatile technique for the preparation of TiO2 thin films with various applications in areas such as photocatalysis and surface cleaning [3].

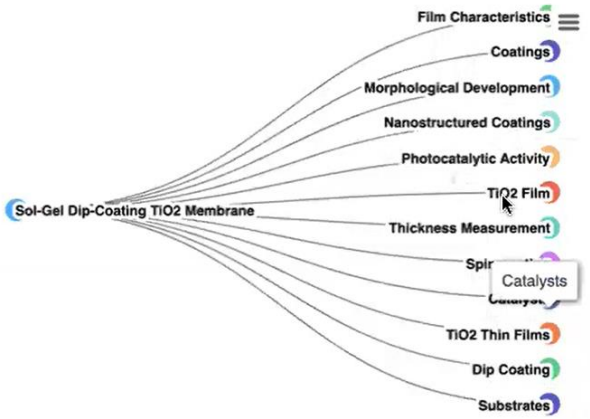
Go Deeper

What are the key steps involved in the sol-gel dip-coating process for creating TiO2 membranes?

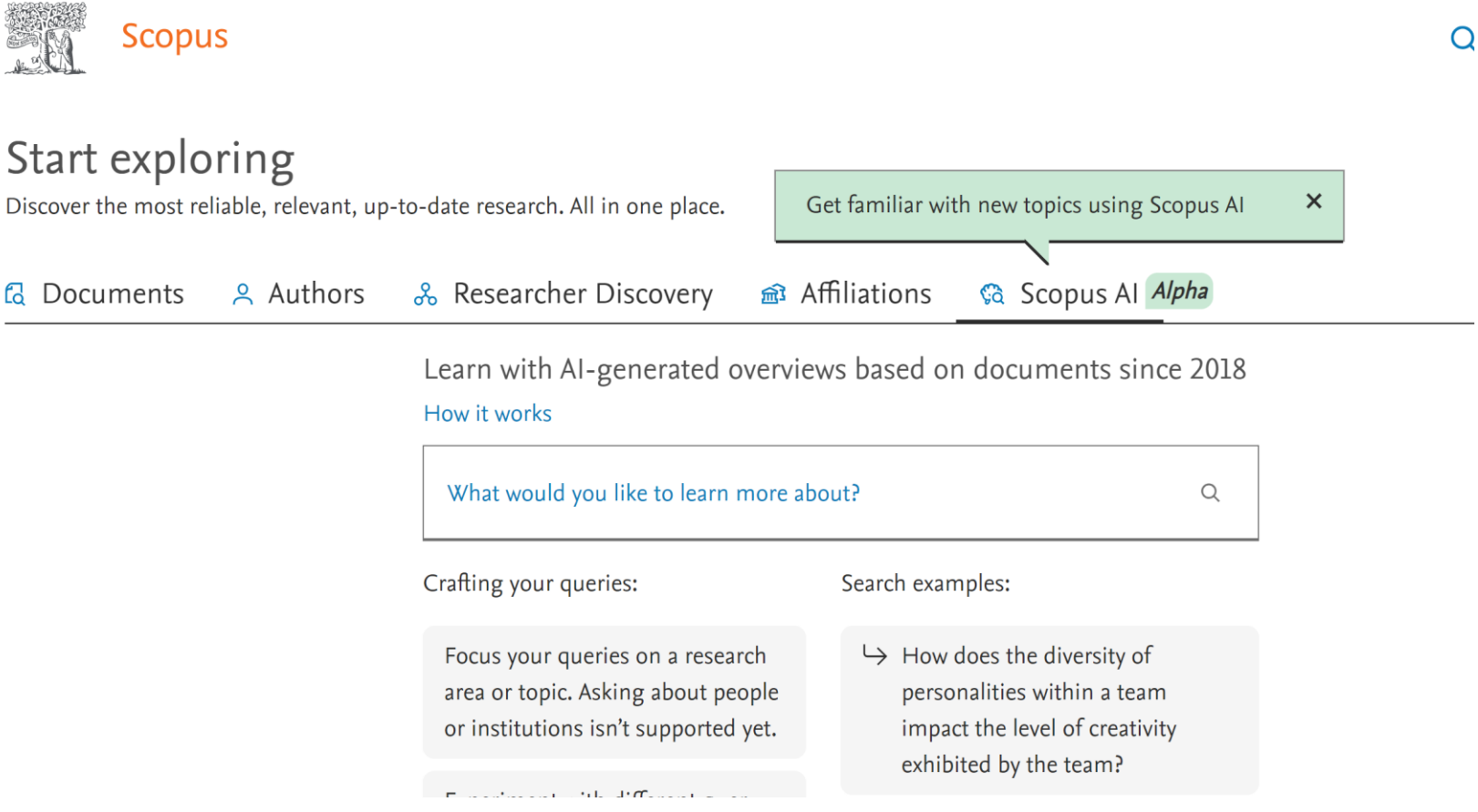
How does the thickness of a sol-gel dip-coated TiO2 membrane affect its performance in applications such as water purification or solar cells?

What are the main factors that influence the porosity and surface morphology of sol-gel dip-coated TiO2 membranes?

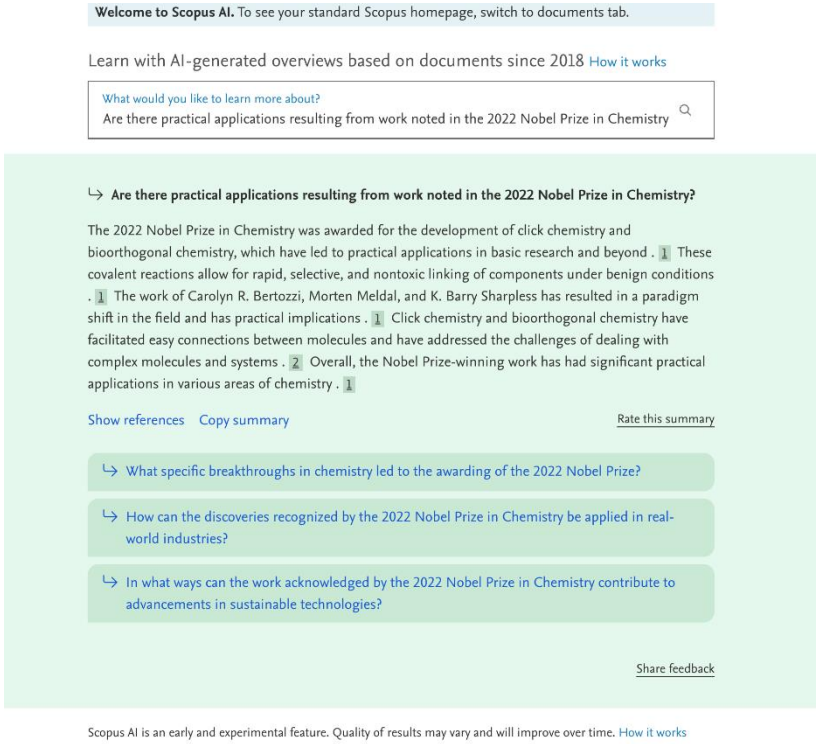
Keyword Graph



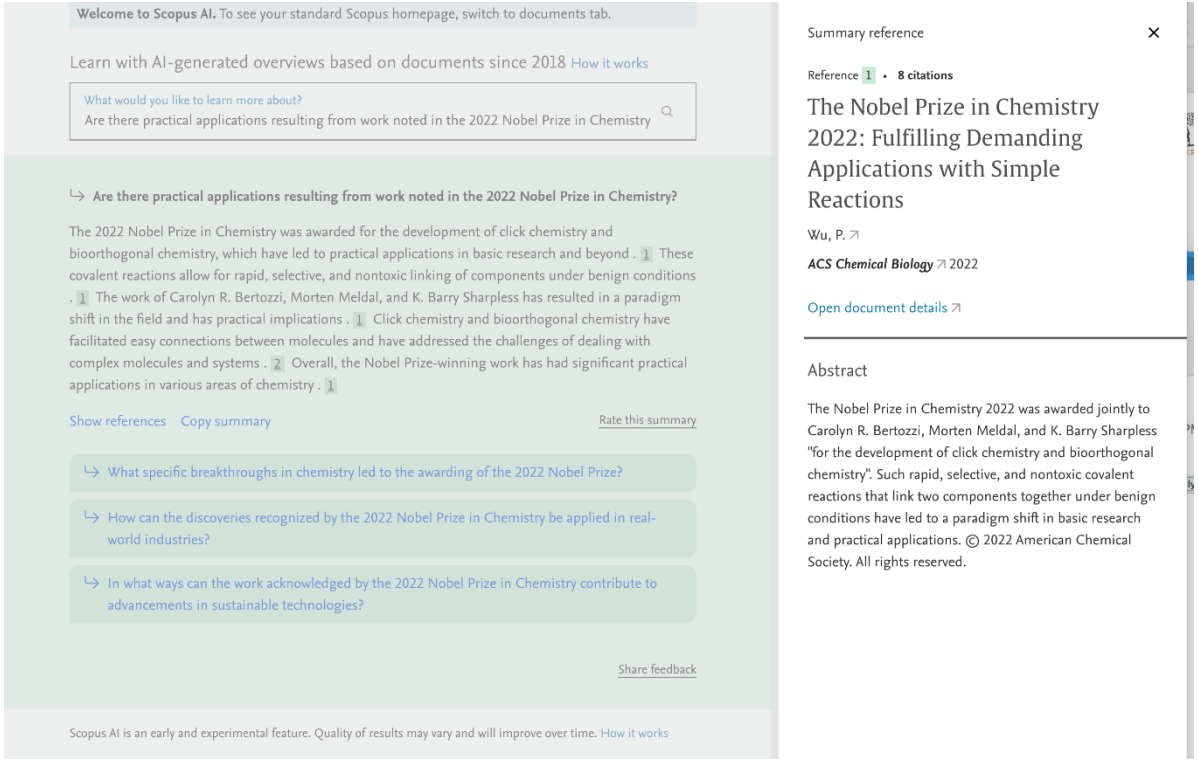
Live Demo


Scopus
Start exploring
Discover the most reliable, relevant, up-to-date research. All in one place.
Get familiar with new topics using Scopus AI
Documents
Authors
Researcher Discovery
Affiliations
Scopus AI Alpha
Learn with AI-generated overviews based on documents since 2018
How it works
What would you like to learn more about?
Crafting your queries:
Focus your queries on a research area or topic. Asking about people or institutions isn't supported yet.
Search examples:
How does the diversity of personalities within a team impact the level of creativity exhibited by the team?

Are there practical applications resulting from work noted in the 2022 Nobel Prize in Chemistry?



Are there practical applications resulting from work noted in the 2022 Nobel Prize in Chemistry?



Acknowledgments

- Daniel Calto, Elsevier
- Yashiko Kakita, Elsevier
- Jesse Mudrick, Elsevier
- Adrian Raudaschl, Elsevier
- Tina Zdawczyk, Elsevier
- Mark Hurwitz, Cornell University
- Alice Li, Cornell University

Additional Information

Ann Gabriel

A.Gabriel@Elsevier.com

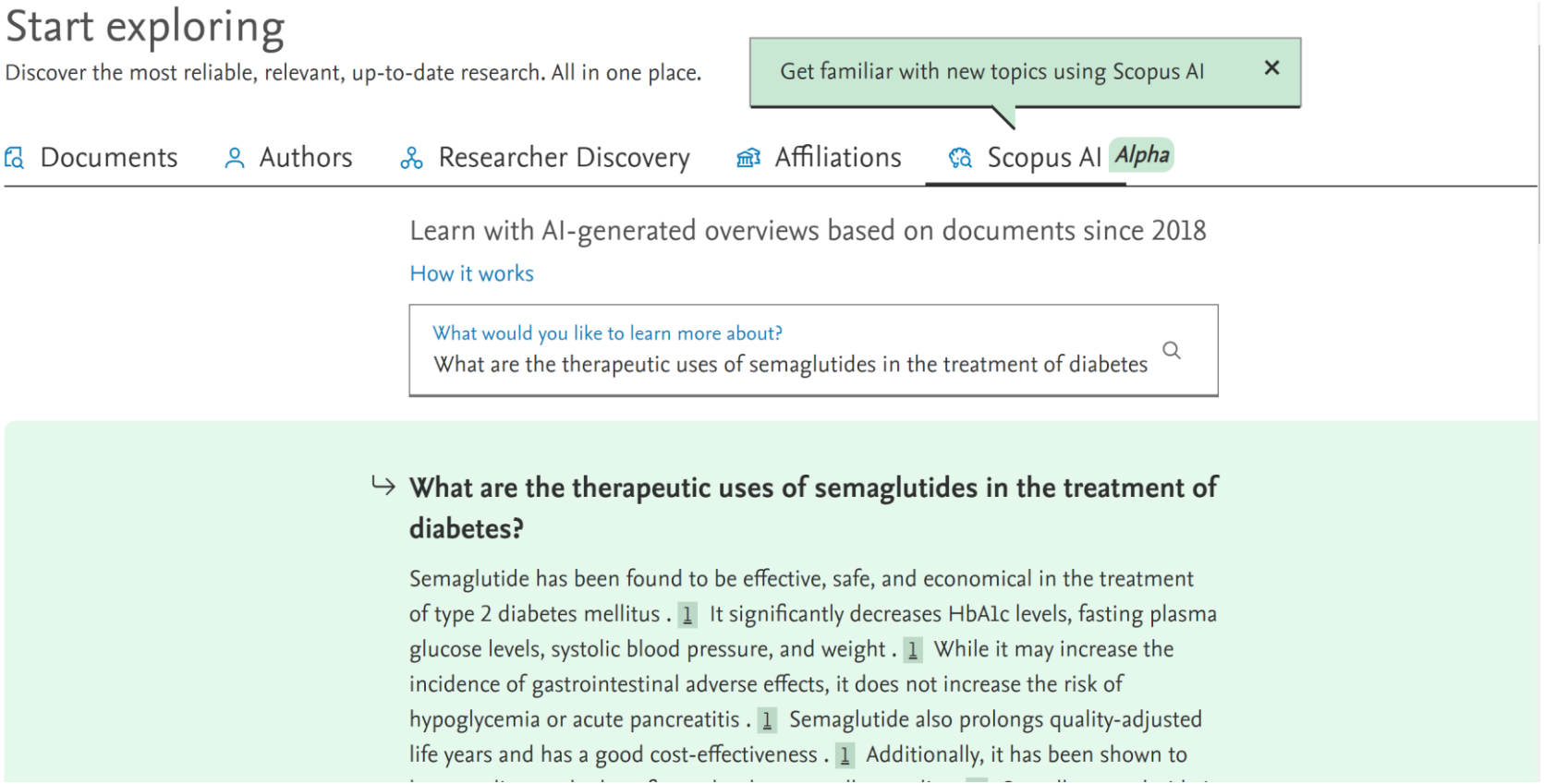
Arthur Ellis

A.Ellis@Elsevier.com

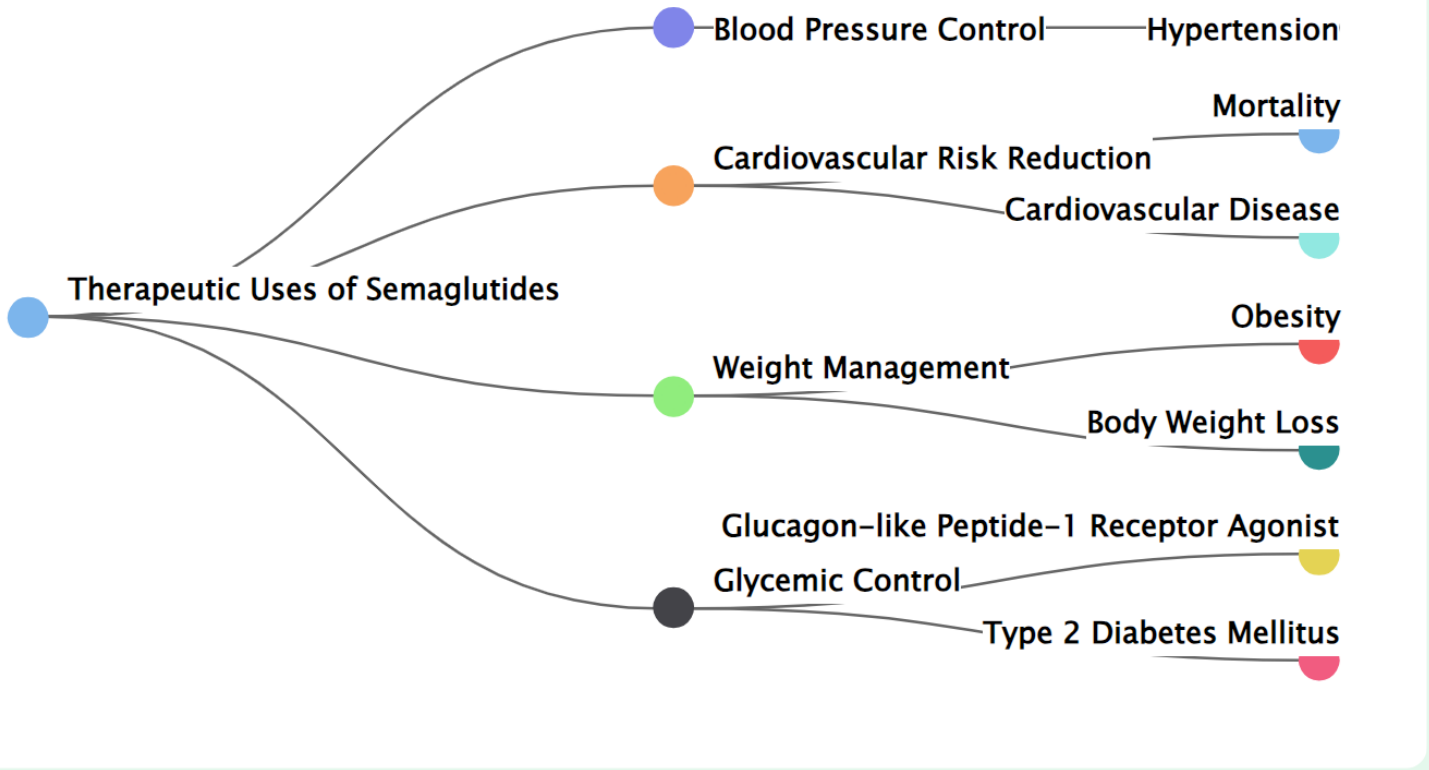
Appendix

- Static version of Live Demo

Query


Start exploring
Discover the most reliable, relevant, up-to-date research. All in one place.
Get familiar with new topics using Scopus AI
Documents
Authors
Researcher Discovery
Affiliations
Scopus AI Alpha
Learn with AI-generated overviews based on documents since 2018
How it works
What would you like to learn more about?
What are the therapeutic uses of semaglutides in the treatment of diabetes
What are the therapeutic uses of semaglutides in the treatment of diabetes?
Semaglutide has been found to be effective, safe, and economical in the treatment of type 2 diabetes mellitus . 1 It significantly decreases HbA1c levels, fasting plasma glucose levels, systolic blood pressure, and weight . 1 While it may increase the incidence of gastrointestinal adverse effects, it does not increase the risk of hypoglycemia or acute pancreatitis . 1 Semaglutide also prolongs quality-adjusted life years and has a good cost-effectiveness . 1 Additionally, it has been shown to

Visualize it



Summary References

Summary references

Reference 1 • **0 citations**

Rapid Health Technology Assessment of Semaglutide in Treatment of Type 2 Diabetes Mellitus

Li, C., Gao, R.

Chinese Journal of Modern Applied Pharmacy 2023

Show abstract

Reference 2 • **1 citation**

Oral semaglutide – Rybelsus®, the first GLP-1 receptor agonist for oral use in clinical practice

Karásek, D.

Vnitrni Lekarstvi 2022

Show abstract

Go deeper

↳ Can semaglutides be used as a standalone treatment or is it typically prescribed in combination with other diabetes medications?

Semaglutide can be used as a standalone treatment for type 2 diabetes mellitus . 1 It has been shown to significantly decrease HbA1c levels, fasting plasma glucose levels, systolic blood pressure, and body weight . 1 Semaglutide is also safe to use, with a low risk of hypoglycemia and acute pancreatitis . 1 Additionally, semaglutide has good cost-effectiveness and can prolong quality-adjusted life years . 1 Therefore, semaglutide is an effective and economical option for the treatment of type 2 diabetes mellitus . 1

Visualize it Show references