

Lens PatSeq:

An open facility for biotech and patent professionals

September 17, 2020

Osmat A. Jefferson

Director of Lens Product Development, Cambia.
Professor of Science and Technology, QUT.



- Global crises demand better innovation practices.
- Problem solving requires effective partnerships and collective action.
- Open evidence, including patent knowledge, is critical for new solutions.
- Trust in public institutions to advance the public interest needs to be restored.



SARS-CoV-2: Navigating information overload with Lens.org

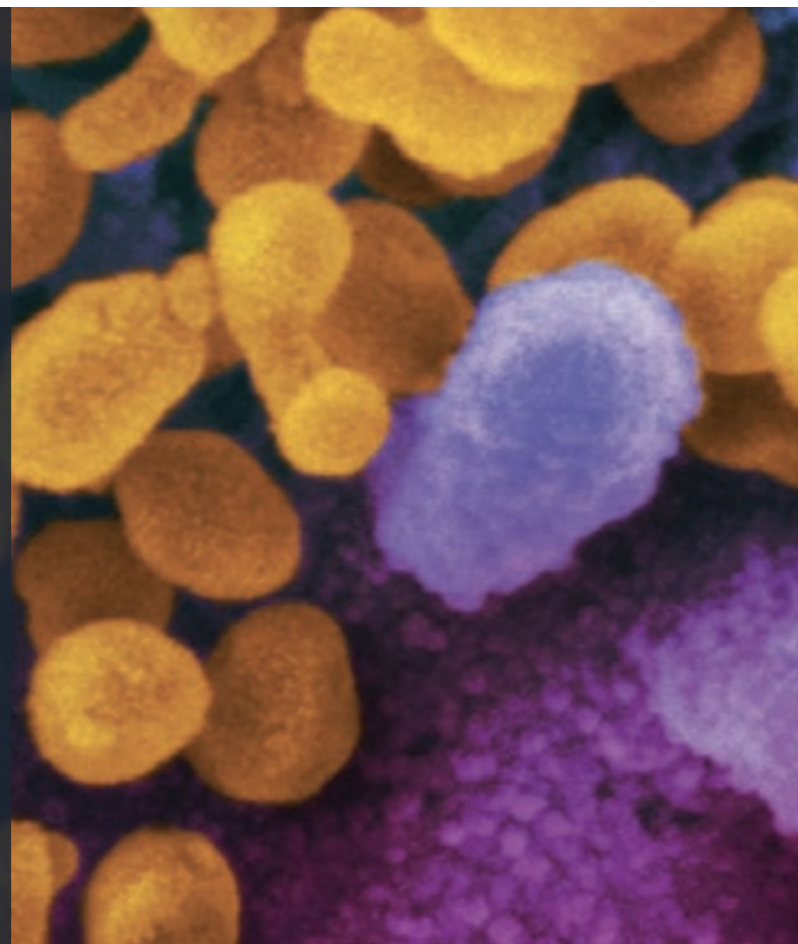
Evidence must be used to guide interventions for the COVID-19 pandemic, but that evidence must be comprehensive, credible and shared for it to be effective. This report focuses on how global scientific research and patenting activity relevant to SARS-CoV-2, including its genetics and pathogenesis, can be discovered and made transparent, open, shareable and navigable to help inform how it could be translated into intervention options.

We demonstrate how the platform Lens.org can be used to conduct reliable systematic scholarly and patent searches, refine and build domain-based collections to share publicly, and set alerts to receive notifications for updates. We also provide some analyses/comments on some of these collections. Our aim is to enable researchers to find clarity into the discovery, analysis, and translation of COVID-19 scientific and patent knowledge into products, practices or services to help mitigate the current crisis

Credit for background picture: a screenshot for SARS-CoV-2 from a photo in [Immune Matter](#), im Spring 2020



Authored By: Osmat Azzam Jefferson, Simon Lang, Kenny Williams, Aaron Ballagh, Deniz Koellhofer, Ben Warren, Bernard Schellberg, Roshan Sharma, Richard Jefferson



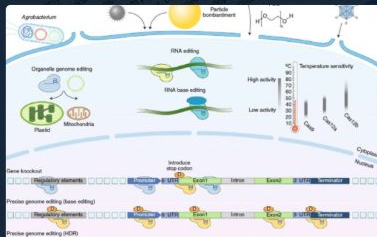
BETA 0.5

Introducing Lens Reports

Lens Reports is a new facility for creating evidence-based, open, sharable and reusable reports linked to real Lens data and analytics. The next evolution of Lens, search, collect, analyze, annotate and now present your findings in a flexible, data-driven reporting interface.

[View Example Report](#)

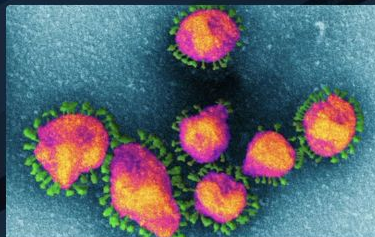
Request Early Access



GENETICS

Mapping innovation trajectories of CRISPR-Cas9 Technology

Fast and easy to implement, cheap and components are readily accessible, its versatility means that it delivers an end product. In plant cells, the technology can be applied to....



VIROLOGY

Human Coronaviruses: Patent and Research Landscape

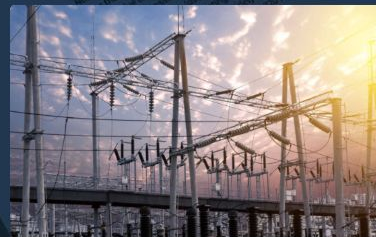
The rapid outbreak of Wuhan coronavirus is pressing health authorities across the globe to better understand how this virus and other coronaviruses spread among humans...



AGRICULTURE

Regenerative Agriculture: Open evidence for its role in solving climate change

Regenerative agriculture refers to a set of practices that can build organic matter back into the soil, effectively storing more water and drawing more carbon out of the atmosphere....



ENERGY

Ultra-Low-Cost Electricity Storage

Increasing penetrations of intermittent renewables, such as wind and solar power, on the world's power grids will require the deployment of growing amounts of energy storage...

Patent sequence search similarity results for SARS-CoV-2 Spike protein

To interpret, understand, and value the combined effect of a genetic sequence on a biological invention, the Lens provides an application, PatSeq Finder, that allows sequence-based similarity searches within the context of almost all known patent sequence disclosures to date. And by linking that information to that from the scientific literature, we can predict contextual uses in patents and potential products that are in the pipeline. Here, we show sequence similarity search results on the whole spike protein and identify three patent owners, Amgen Inc, Academia Sinica, and Crucell Holland BV who have referenced sequences similar to SARS-CoV-2 spike protein in their granted patent claims. More details on these granted patents and other patent applications are below.

Start another search

Results

Query: MFVFL

Location in document

Showing 1 to 4 of

Sequence

Patent Application US 8106170 B2 [VIEW DOCUMENT](#)

SEQUENCE LISTING INFORMATION

[SEQ ID NO 115](#)

SEQUENCE TYPE	AA
SEQUENCE LENGTH	1,255
ORGANISM	Artificial

1 alignment were found:

Alignment 1

E-VALUE	0	LENGTH	1,277aa
Query	<div style="width: 76.0%; background-color: #e91e63; height: 10px;"></div> S		
Hit	<div style="width: 100%; background-color: #e91e63; height: 10px;"></div> C		

US, Published J

Applicants: TER

Organism: Artifi

Summary

Compositions against SARS-coronavirus and uses thereof

APPLICANTS [TER MEULEN JAN HENRIK, VAN DEN BRINK EDWARD NORBERT, ...](#) More

INVENTORS [TER MEULEN JAN HENRIK, VAN DEN BRINK EDWARD NORBERT, ...](#) More

PUBLICATION DATE **Mon, 30 January 2012**

FILING DATE **Wed, 09 November 2005**

▶ ABSTRACT

▼ CLAIMS

1. A composition comprising at least two immunoglobulins that are capable of specifically binding to Spike (S) protein of Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and neutralizing SARS-CoV, wherein the first immunoglobulin comprises a heavy chain variable region comprising the amino acid sequence of SEQ ID NO: 4 and a light chain variable region comprising the amino acid sequence of SEQ ID NO: 8, and wherein the second immunoglobulin comprises a heavy chain variable region comprising the amino acid sequence of SEQ ID NO: 6 and a light chain variable region comprising the amino acid sequence of SEQ ID NO: 10, and wherein the relative amounts of the at least two immunoglobulins are such that they exhibit a synergistic neutralizing activity.

Alignments

average and medians

☐

☐

Lens.org and Cambia

- Cambia, founded in 1992, started Patent Lens in 2001, the precursor of lens.org, the world's first free and open full text patent search capability.
- In 2006, launched the world's first public patent sequence search capability for United States patent applications.
- Since 2013, Lens.org has served the world most comprehensive public Patent Sequence platform to navigate biological patents from over 17 jurisdictions.
- In 2011, with seed money from USPTO, Cambia began extracting and resolving non-patent literature strings in collaboration with NIH-NCBI and Crossref.
- By 2014, Cambia linked non-patent literature to patents and began serving the data, created [PatCite](#) and [In4M](#) metric to map influence of research on industry and foster meaningful partnerships.
- Lens honors privacy and confidentiality and its data is fully open, shareable and reusable.
- The platform has been up 24/7 for over 20 years.



CURRENT OFFERINGS

Discovery, Analytics, And Management Tools

API & Data Facility:

Scholarly API, PatSeq bulk data, Patent API*

Reports:

Assemble your saved queries and collections with other knowledge in a dynamic and interactive report*

Collections:

based on Scholarly Works and Patents Lens searches and analyses.

In4M:

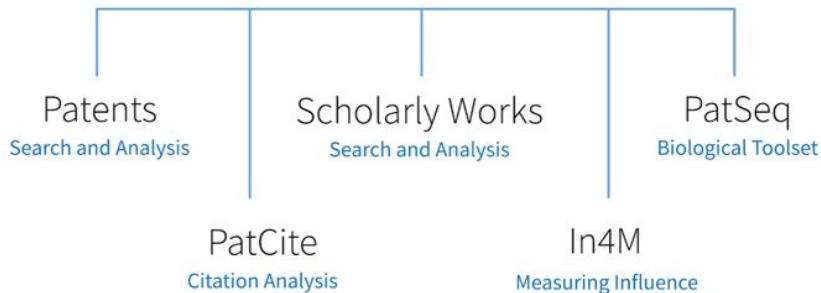
International Industry & Innovation Influence Mapping

* coming soon



LENS.ORG

Solving The Problem Of Problem Solving™



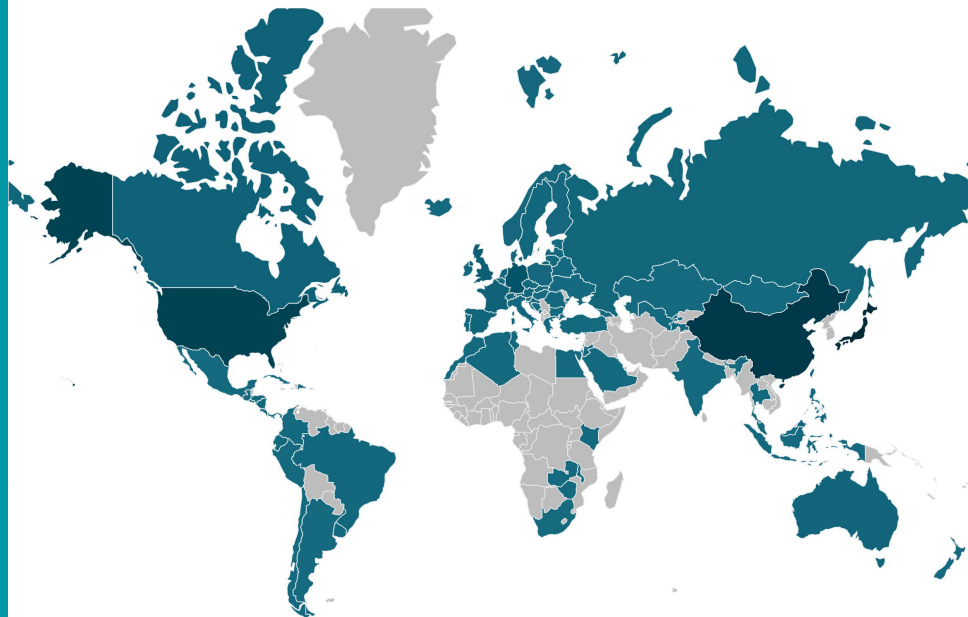


September 2020

Lens Patent Data

125.4 Million Patent Records:

- 105 jurisdictions
- 68.8M patent families
- 722k biological patents
- 360.6M patent sequences

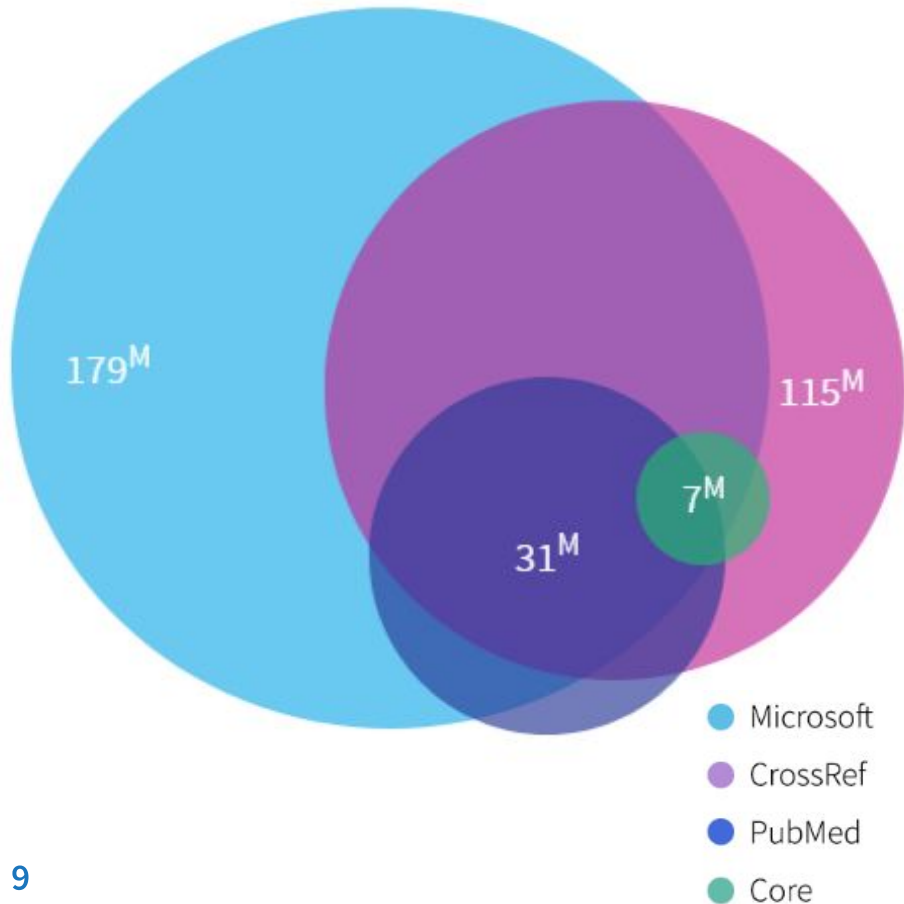


September 2020

Lens Scholarly Data

220.4 million research publication records:

- 112.7M journal articles
- 17.5M Books and book chapters
- 7.3M conference proceedings
- 4.2M works cited in patents
- 74.1M works cited by other scholarly works
- 1.6B scholarly citations

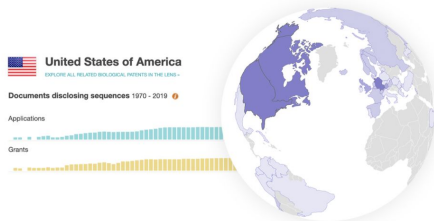
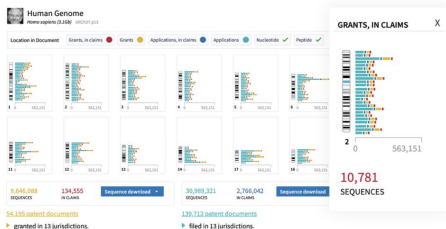


PatSeq Data

Compare biological patent holdings and view sequence disclosures across jurisdictions over time.

Patent Sequence Data is released to the public from 3 major sources: 1) national patent offices, 2) collaborating 3rd party public sequence listings repositories, and 3) regional or global intellectual property organizations.

[Start using](#) [Watch Video](#)

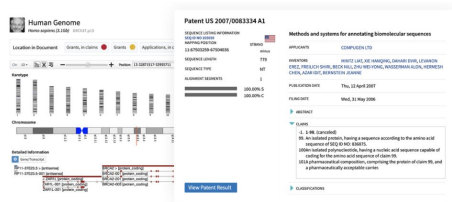


PatSeq Explorer

Navigate and analyse patent-disclosed sequences mapped onto genomes and chromosomes.

The PatSeq Explorer offers a dynamic, embeddable, and multi-level view of mapped patent sequences on a genome of a specific organism. We currently offer patent sequences mapped onto the human, mouse, maize, rice, and soybean genomes, with other genomes to be mapped soon.

[Start using](#) [Watch Video](#)

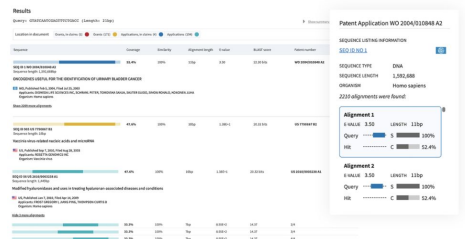


PatSeq Analyzer

Compare patenting activity at the chromosomal locus or gene level.

The PatSeq Analyzer is a modified genome viewer tool that is now a stand alone facility but links to PatSeq Explorer to give you a full view on a patent sequence details.

[Start using](#) [Watch Video](#)



PatSeq Finder

Use input DNA or protein sequences to find matches in the Lens PatSeq database.

The PatSeq Finder is a sequence similarity search tool based on BLAST, allowing you to search the Lens patent sequence (PatSeq) databases for matches to a sequence of your interest. This tool is unique since it enables you to conduct sequence-based searches within more than 250 million patent sequences that we serve in either a nucleotide-based or protein-based databases.

[Start using](#)

* PatSeq Text Not Pictured



PATSEQ

Explore Biological Sequences in Patents

DNA and protein in patents are crucial to understand and harness new science for health, agriculture and the environment.

Lens hosts the world's largest publicly available database and toolkit for biological patents, with internal transparency.

Patent sequence search similarity tool

Using DNA, RNA, or Protein sequences, search-privately and securely- for similar sequences that are disclosed in the Lens patent databases. The tool is based in part on the public NIH-NCBI BLAST program. In the results view, compare sequence alignments or patent claims and/or export/embed findings.

Your Search Params

TYPE

peptide

LENGTH

223aa

Database

NAME

patseq-aa

ENTRIES

71,001,000

SIZE

10,351,155,755 aa

BLAST search

TOOL

blastp

SUBSTITUTION MATRIX

BLOSUM62

MAX E VALUE

10

FILTER

disabled

MAX NO. HITS

500

SHORT QUERY OPTIMISATION

disabled

MASK

disabled

LOWERCASE

disabled

show

PatSeq Finder

Want to save this search? Enter a name here

Submit search

Enter sequence

RYQPTESIVR FPMITNLCPF GEVFNATIFA SIYHNNRISL SNCVADSYSL YNSASFSTFK CYGVSPTKLN DLCTFNAYAD
SFVINGEDIVR QIAPQTKGI ADYNYKLPOD FTGCVJANNS NLDSKGVGN WNYLYRLFRK SNLKPFRDI STEIQVAGST
PCNGVEGNC YFLQSYGQF PTNGVQGPY RRVVLSLH HAPATVCGPK KSTNLVKNKC VNF

or upload a FASTA sequence file:

Choose File

 No file chosen

Press tab or [click here to pre-populate](#) the form based on this query

Open query subrange options

Sequence database

Amino Acid db

71,001,000 sequences
Last updated: Aug 24, 2020

Nucleotide database

289,558,390 sequences
Last updated: Aug 24, 2020

Sequence type

Nucleotide

Protein

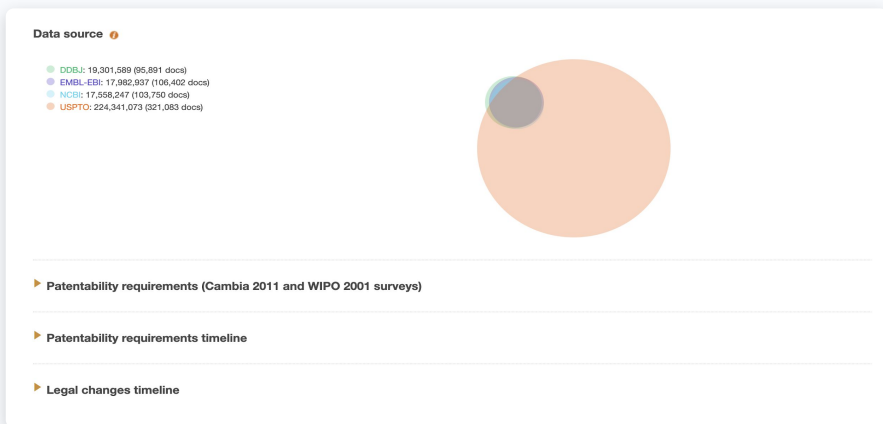
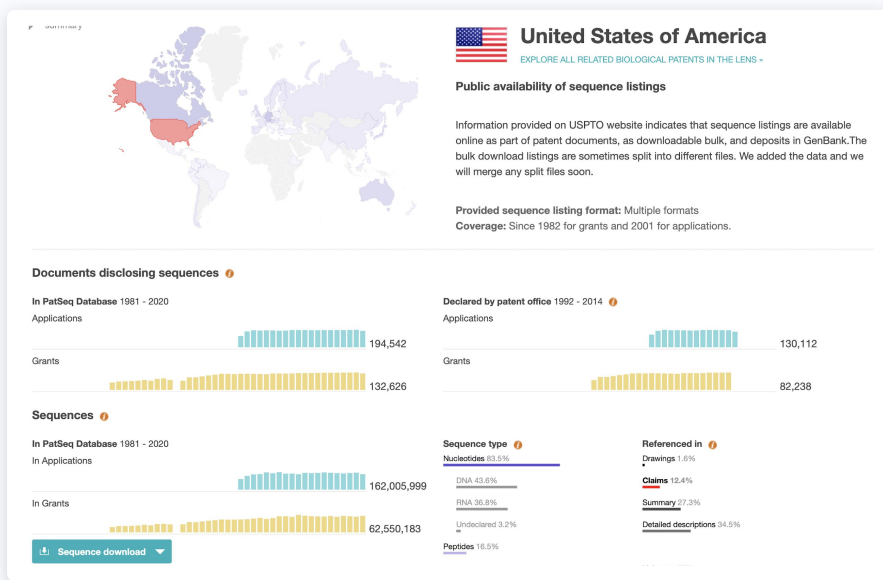
Search strategy

Open advanced options

Want to save this search? Enter a name here

Submit search

[illegible]



PATSEQ

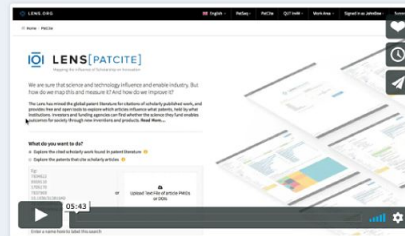
Transparency in PatSeq data

1. Check which jurisdictions make their data publicly available in machine searchable format.
2. Compare biological patent holdings with those disclosed by patent offices.
3. Verify the level of overlap among the various data sources including those from public databases.

PATCITE

Open Influence Mapping Facility

Using open persistent identifiers, funders, investors, researchers, or departments can discover and analyse who in the industry has cited their works and map that influence on invention and innovation.



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.

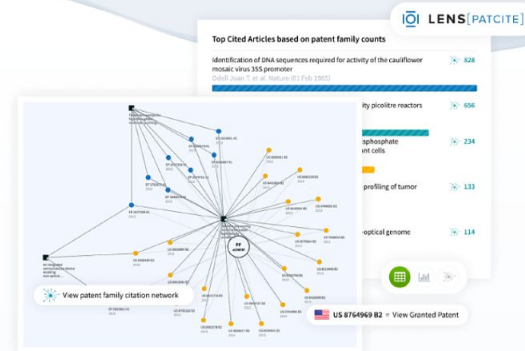
Discover Precise Partners

Discover which scholarly articles have influenced what patents and who is using your scholarly work.

Dynamic and interrogable, these maps allow you to identify important linkages, build networks of collaborations. The granularity of this tool allows you to gain real-time insights into how science and scholarship are shaping patent-based inventions and which research article, which scientists or researchers, and potentially, which institutions have influence over a subset of economic activity.

Dynamic maps that can be used by individual researchers/inventors, university departments, institutions, or even funding organizations.

[Browse example portfolios below](#)



Use Case Scenarios

Discover which scholarly articles have influenced what patents. PatCite is a granular tool suited for diverse user groups and here are some PatCite scenarios.



Funders & Investors

Who is pursuing important inventions based on the science you fund?



Research Institutions

What technology sectors are being influenced by your scholarly work?



Tech Transfer Offices

Which commercial entities are using your institution's research?



Commercial Entities

What research has influence your IP and patent portfolio?



Individual Scientist

What inventions being developed are informed by your work?



Individual Researcher

Who is using my research and how am I influencing inventions?

California Institute of Technology

10.2 Analysis of biological materials	6.1 Audio-visual technology	9.0 Basic communication processes	4.2 Basic materials chemistry	11.2 Biotechnology
7.2 Chemical engineering	4.1 Civil engineering	5.3 Computer technology	5.0 Digital communication	4.8 Electrical machinery, apparatus, energy
5.9 Engines, pumps, turbines	4.1 Environmental technology	5.9 Food chemistry	10.2 Furniture, games	6.7 Handling
7.3 Instruments-Control	3.1 IT methods for management	2.9 Machine tools	5.4 Macromolecular chemistry, polymers	4.3 Materials, metallurgy
7.5 Measurement	10.5 Mechanical elements	3.2 Medical technology	9.7 Micro-structural and nano-technology	7.2 Optics

Swiss Federal Institute of Technology Lausanne

4.2 Analysis of biological materials	10.3 Audio-visual technology	8.5 Basic communication processes	3.8 Basic materials chemistry	3.5 Biotechnology
6.5 Chemical engineering	4.5 Civil engineering	6.7 Computer technology	5.4 Digital communication	10.4 Electrical machinery, apparatus, energy
5.9 Engines, pumps, turbines	3.7 Environmental technology	0.8 Food chemistry	8.5 Furniture, games	3.3 Handling
5.7 Instruments-Control	2.9 IT methods for management	6.1 Machine tools	3.8 Macromolecular chemistry, polymers	7.4 Materials, metallurgy
8.5 Measurement	6.1 Mechanical elements	5.1 Medical technology	14.6 Micro-structural and nano-technology	8.3 Optics

Korea Advanced Institute of Science and Technology

2.1 Analysis of biological materials	6.7 Audio-visual technology	6.7 Basic communication processes	3.1 Basic materials chemistry	3.2 Biotechnology
4.1 Chemical engineering	3.1 Civil engineering	4.1 Computer technology	4.5 Digital communication	7.4 Electrical machinery, apparatus, energy
2.8 Engines, pumps, turbines	4.4 Environmental technology	1.6 Food chemistry	2.0 Furniture, games	2.0 Handling
3.4 Instruments-Control	2.6 IT methods for management	6.7 Machine tools	5.3 Macromolecular chemistry, polymers	6.6 Materials, metallurgy
4.3 Measurement	3.2 Mechanical elements	3.3 Medical technology	12.8 Micro-structural and nano-technology	6.9 Optics
2.2 Organic fine chemistry	4.1 Other consumer goods	4.5 Other special machines	1.9 Pharmaceuticals	8.6 Semiconductors
8.3 Surface technology, coating	6.5 Telecommunications	4.2 Textile and paper machines	7.5 Thermal processes and apparatus	2.6 Transport

In4M:

International Industry & Innovation Influence Mapping

An open, transparent and dynamic tool to explore, expose and measure the degree to which an institution's work product influences outcomes for society, in myriad forms.



COLLECTIVE

The Lens Collective

Open. Inclusive. Comprehensive. Enabling.

Initially supported by the Bill & Melinda Gates Foundation, AP Sloan Foundation, The Rockefeller Foundation and Amazon Web Services.

The Lens Collective allows government agencies, research institutions, foundations and enterprise to join, support, shape, use and benefit from a sustainable open infrastructure.

Launching in January 2021.
Beta members joining now.