THOMSON REUTERS

ADVANCING RESEARCH AND INNOVATION WITH THE WEB OF SCIENCE

Ning Ning, PhD

Solution Consultant, IP & Science
Thomson Reuters

ning.ning@thomsonreuters.com

A surfer swims at Barra da Tijuca beach in Rio de Janeiro. February 22, 2013. REUTERS/Ricardo Moraes

Innovation is about connecting the dots

1. collect diverse dots
2. cluster the related dots
3. connect the dots

Juan Antonio Solares, Innovating by playing "Connect the Dots", http://www.ixl-center.com, 2014

1. COLLECTING DIVERSE DOTS:
Challenge of information overload

50,000+
Academic Journals

7,500,000+
Conference papers

110,000
Conferences

2,000,000+
chemical structural formula

50,000,000+
Patents

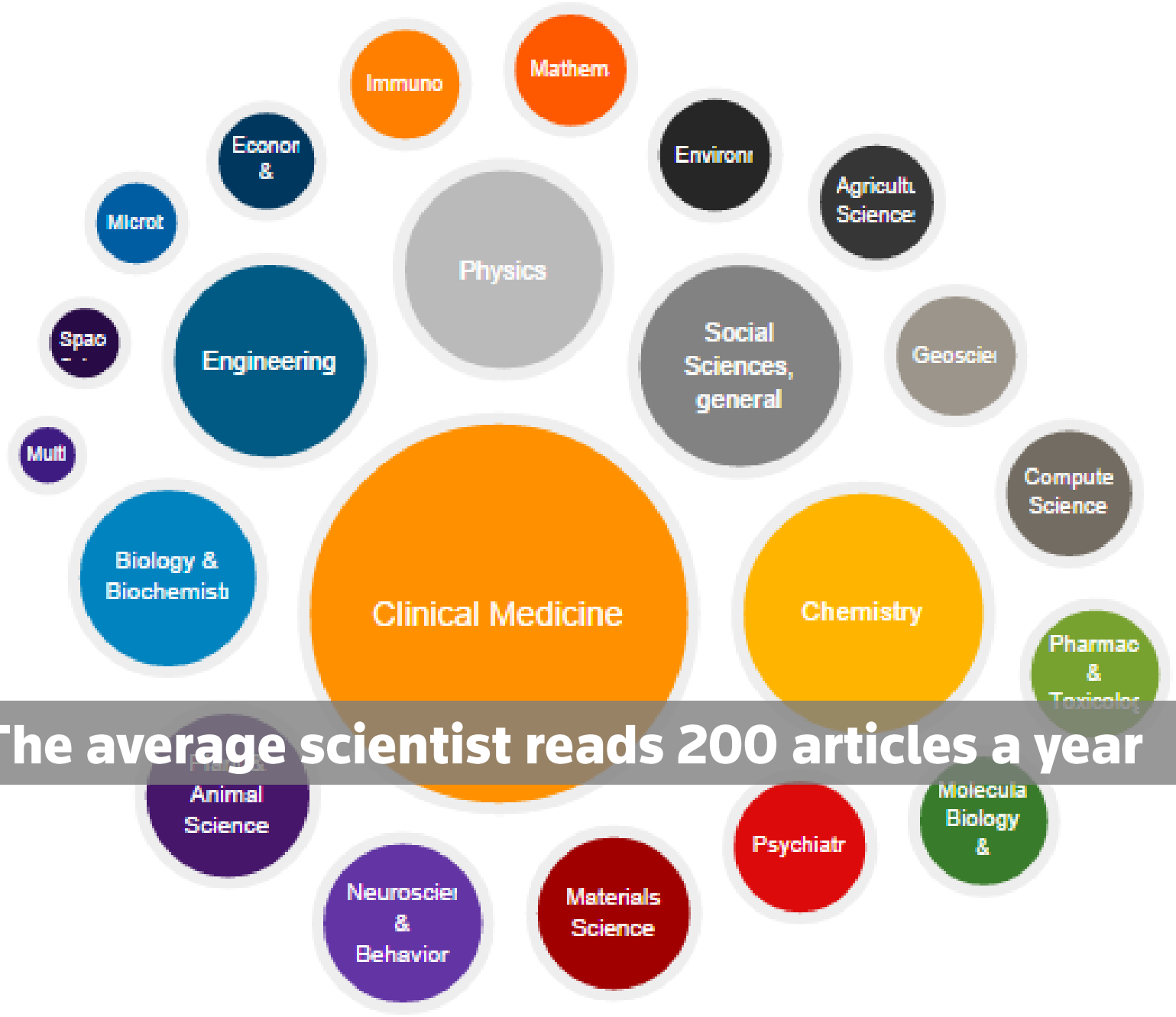
60,000+
Academic books

42,000+
International trademark applications (per year)

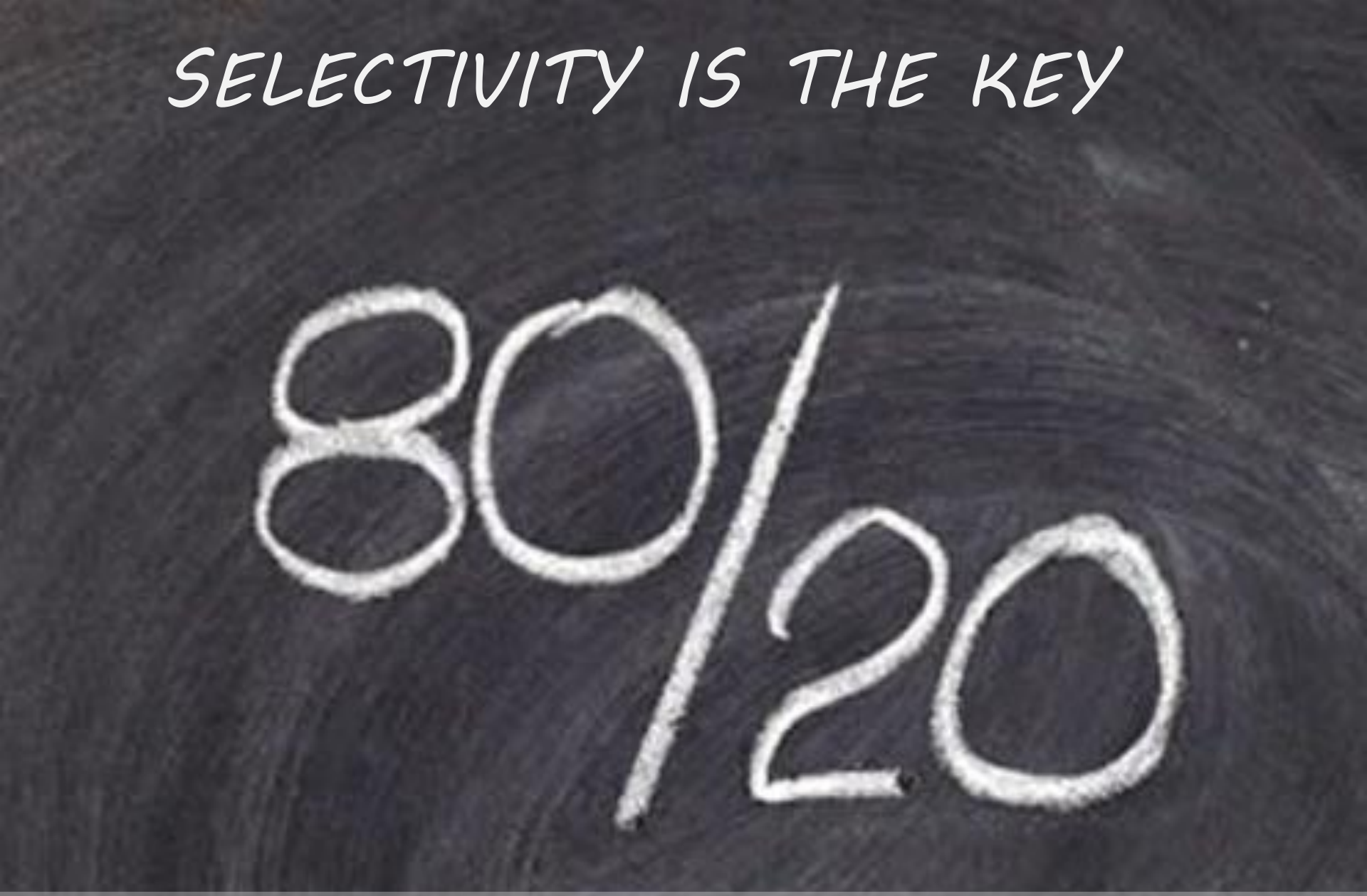
5,000+
Paper formats

400,000+
industrial standards

80 million+
Gene Sequences



The average scientist reads 200 articles a year



A relatively small number of journals publish the majority of significant scholarly results

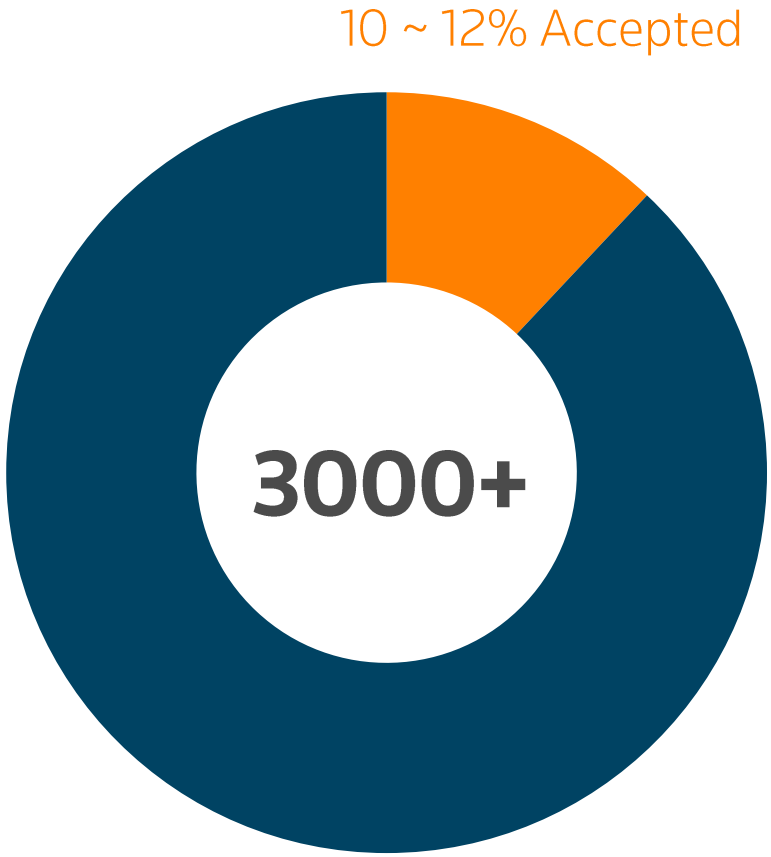
DELIVERING THE “200 ARTICLES”

THE SEA OF SCHOLARLY INFORMATION

THOMSON REUTERS JOURNAL SELECTION

- Journal Publishing Standards
- Editorial Content
- International Diversity
- Citation Analysis

THE CORE OF SCIENCE



WEB OF SCIENCE CORE COLLECTION

- It is a vast multidisciplinary collection of standardized metadata - source items and cited references - from a selected and rigorously curated collection of

12,500 Top Tier
JOURNALS
1900 - 2014

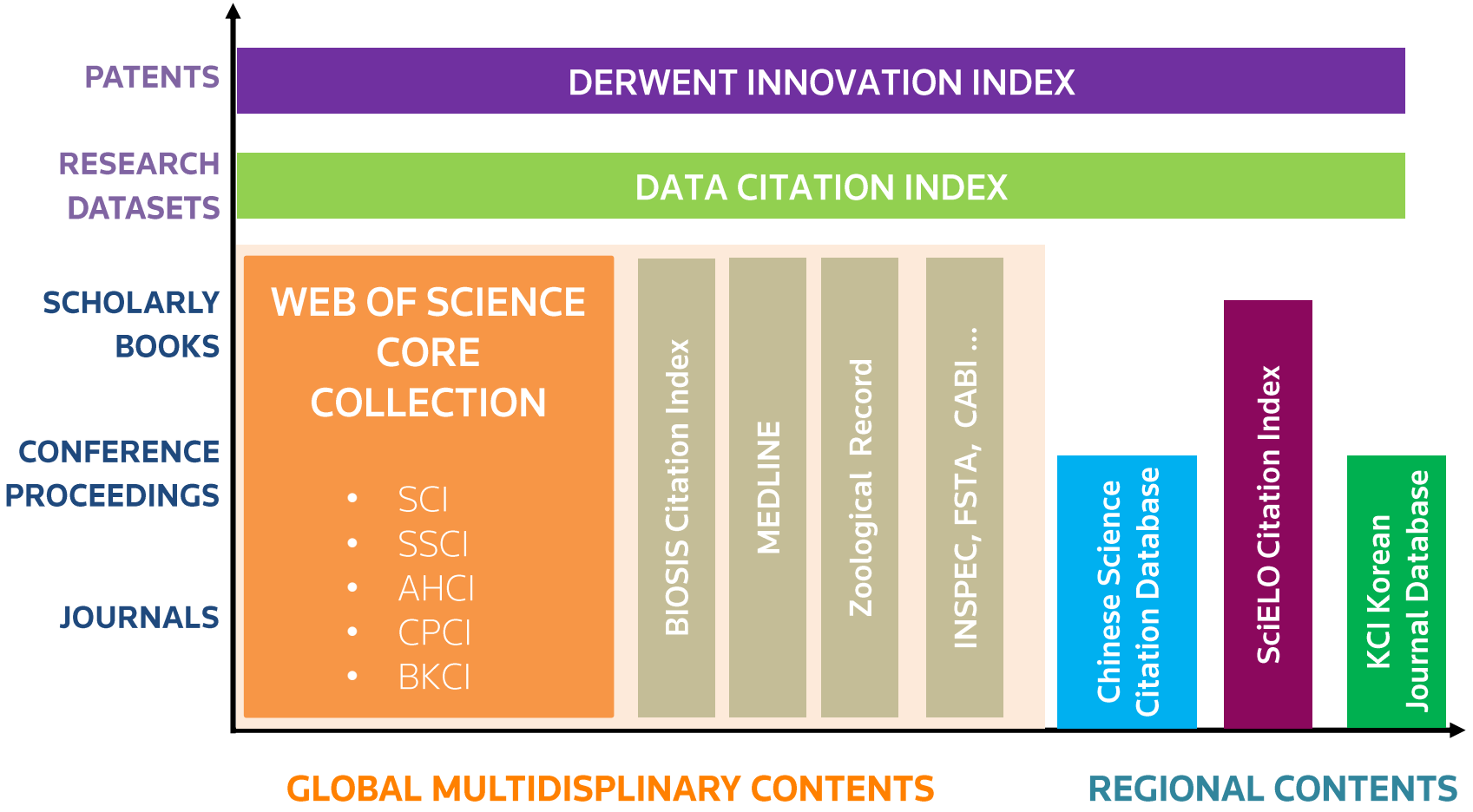
161,000
CONFERENCES
1990 - 2014

52,000
SCHOLARLY BOOKS
2005 - 2014

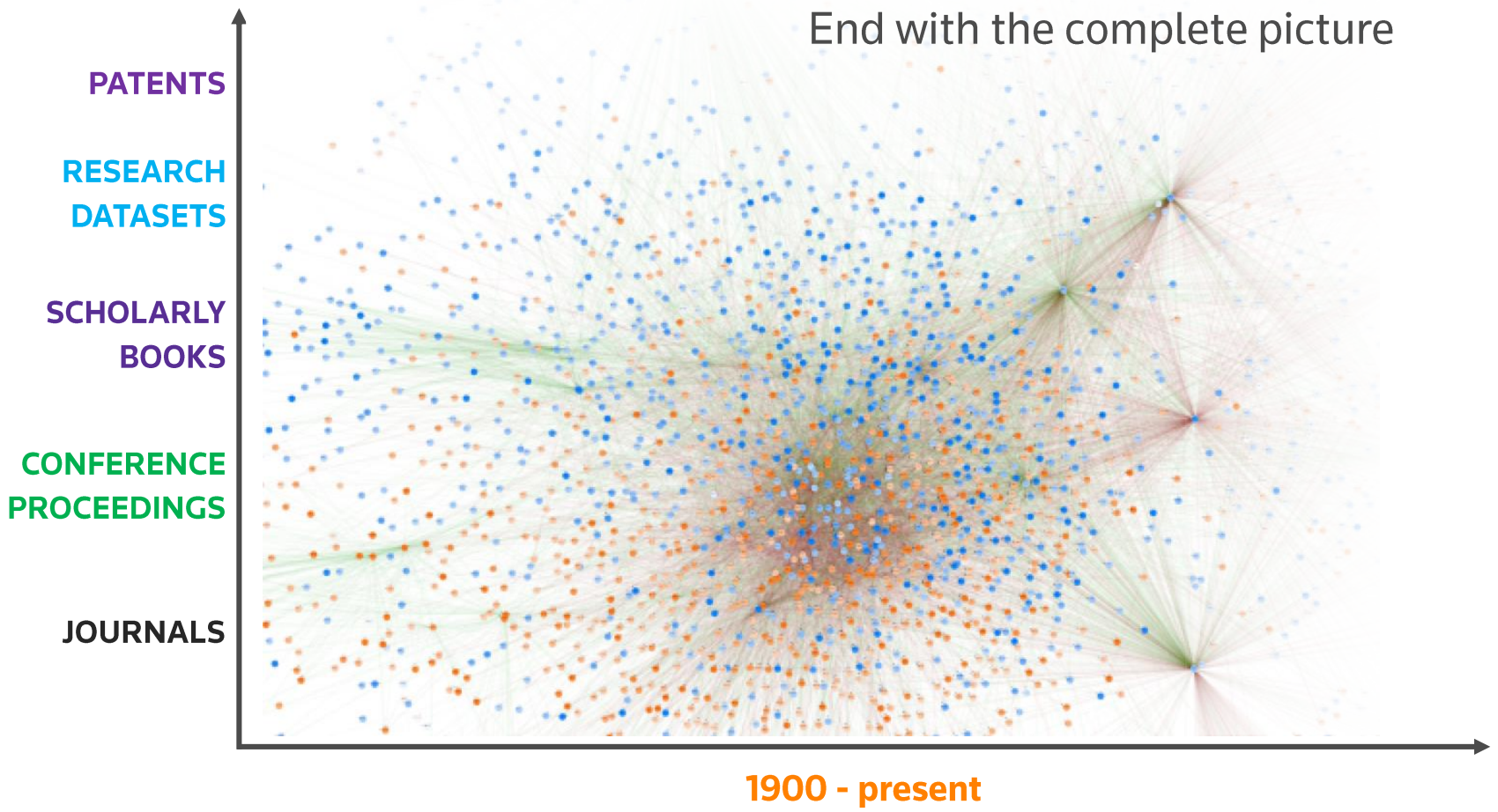
57 million source items • 1 Billion Cited References

Not to have everything, but to represent everything

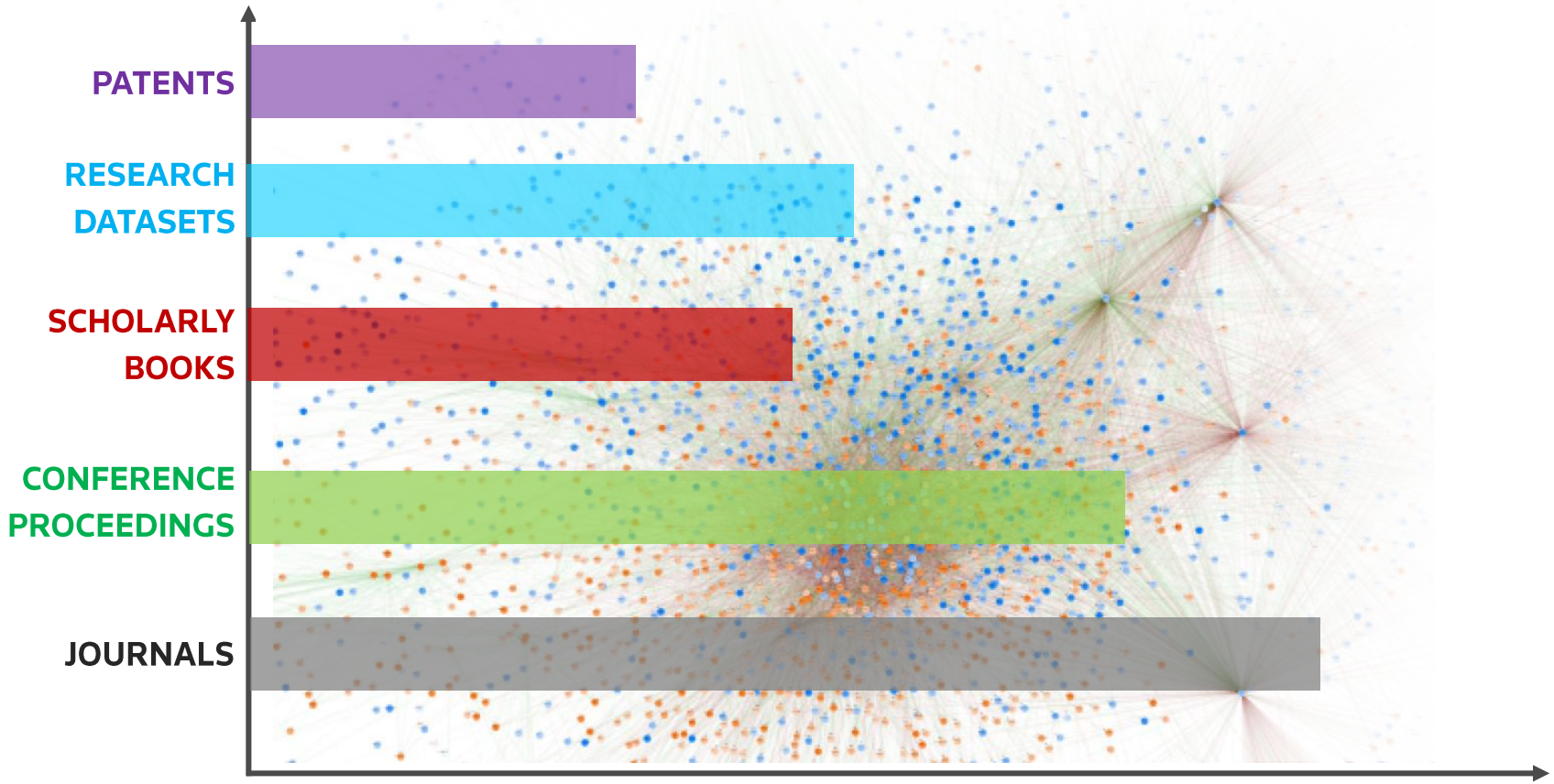
COLLECTING DIVERSE DOTS FOR INNOVATION



COLLECTING DIVERSE DOTS TO SEE THE COMPLETE PICTURE



COMPLETE VIEW OF RESEARCHERS' PUBLICATION PERFORMANCE




ALFR·
NOBEL
NAT·
MDCCC
XXXIII
OB·
MDCCC
XCVI

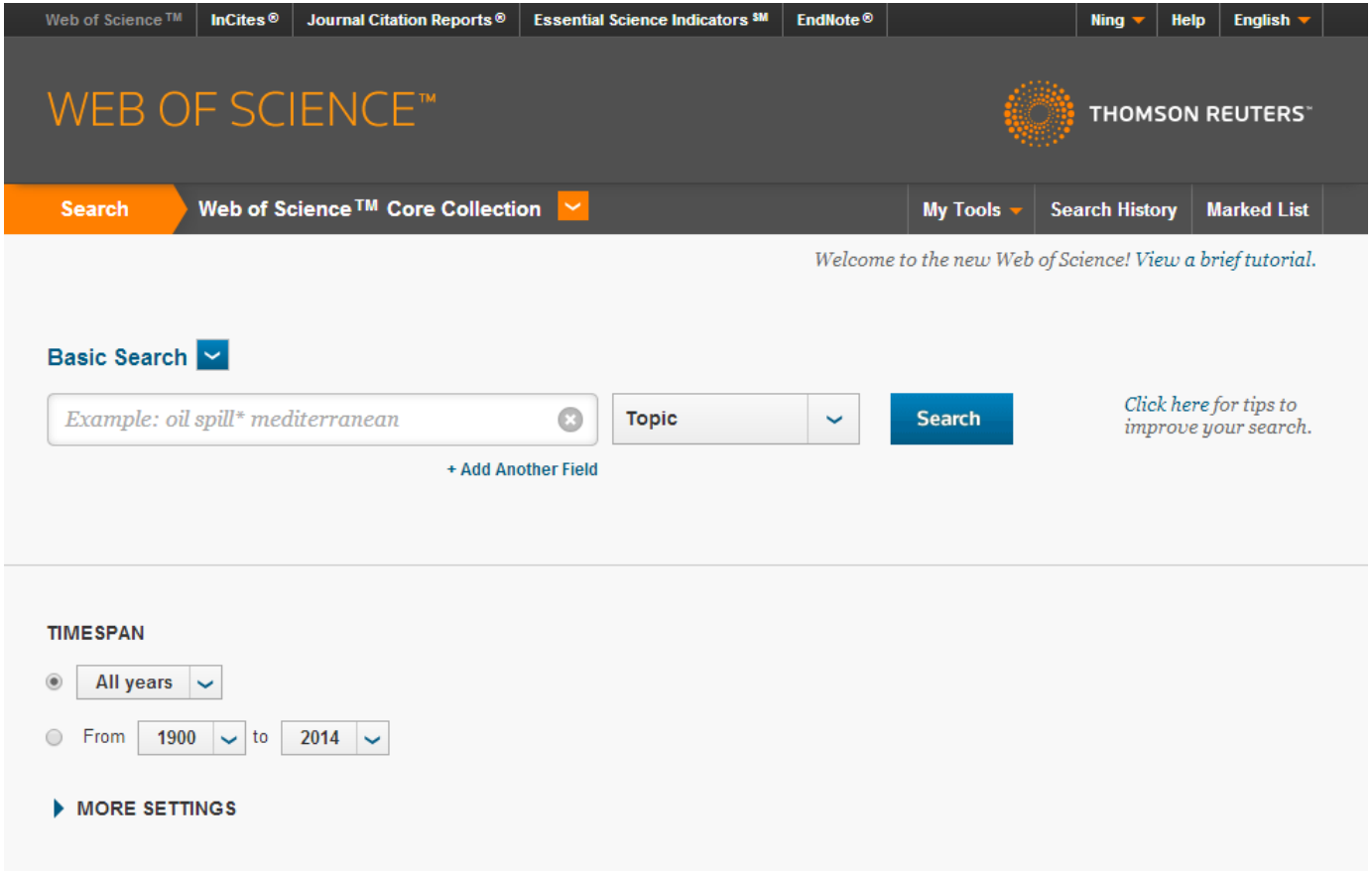
Since 2002, Thomson Reuters has accurately forecasted

37

Nobel Prize recipients

The crystal ball is…

mining the citations in Web of Science



1. COLLECTING DIVERSE DOTS:

- Selective yet comprehensive
- Broader, deeper

2. CLUSTER THE RELATED DOTS :

Putting the data into context

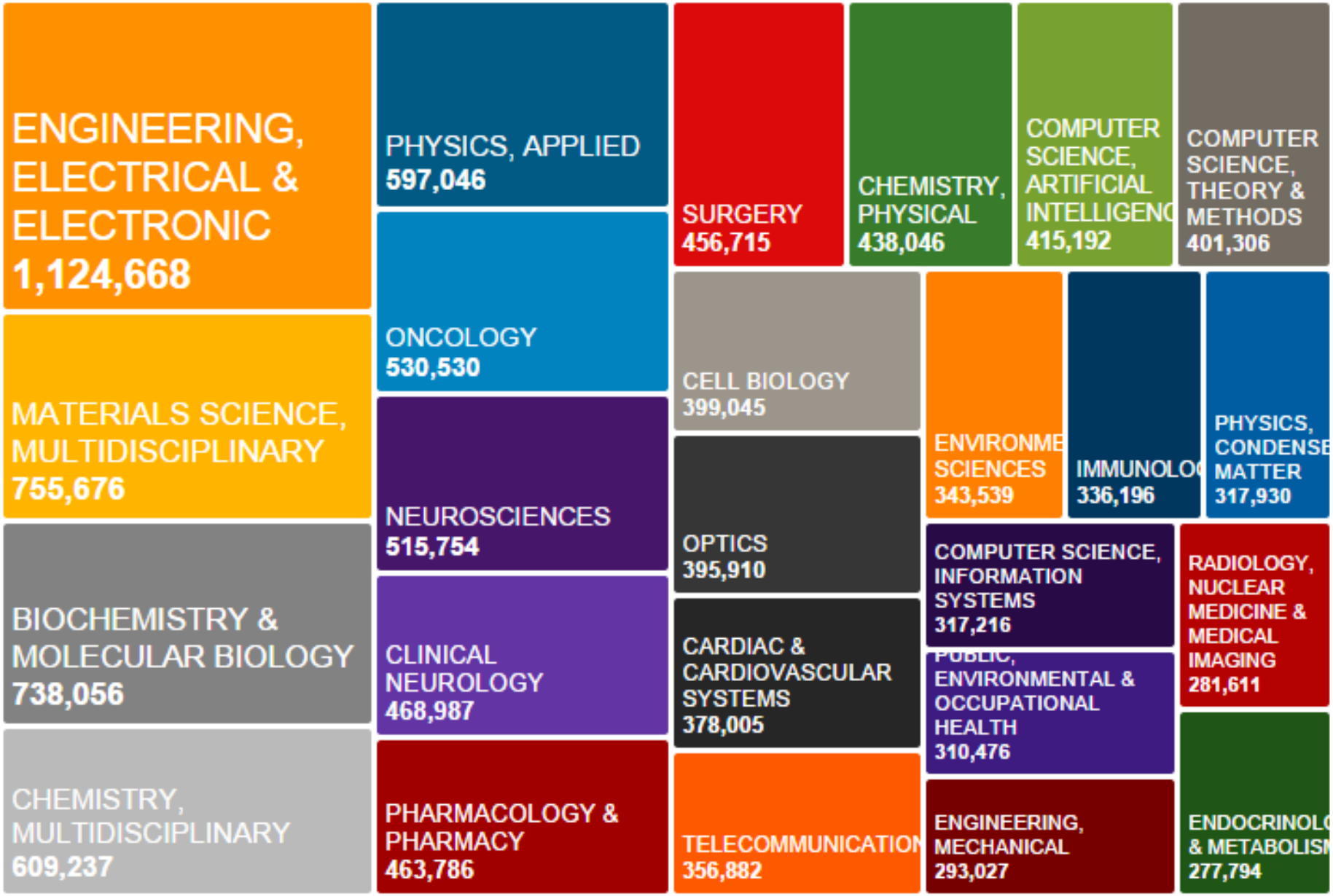
- By Category
- By Text Mining and Clustering
- By Citation

CLUSTER THE RELATED DOTS
By Categories

- Which of the following Web of Science categories has the most number of papers in the last 10 years globally?

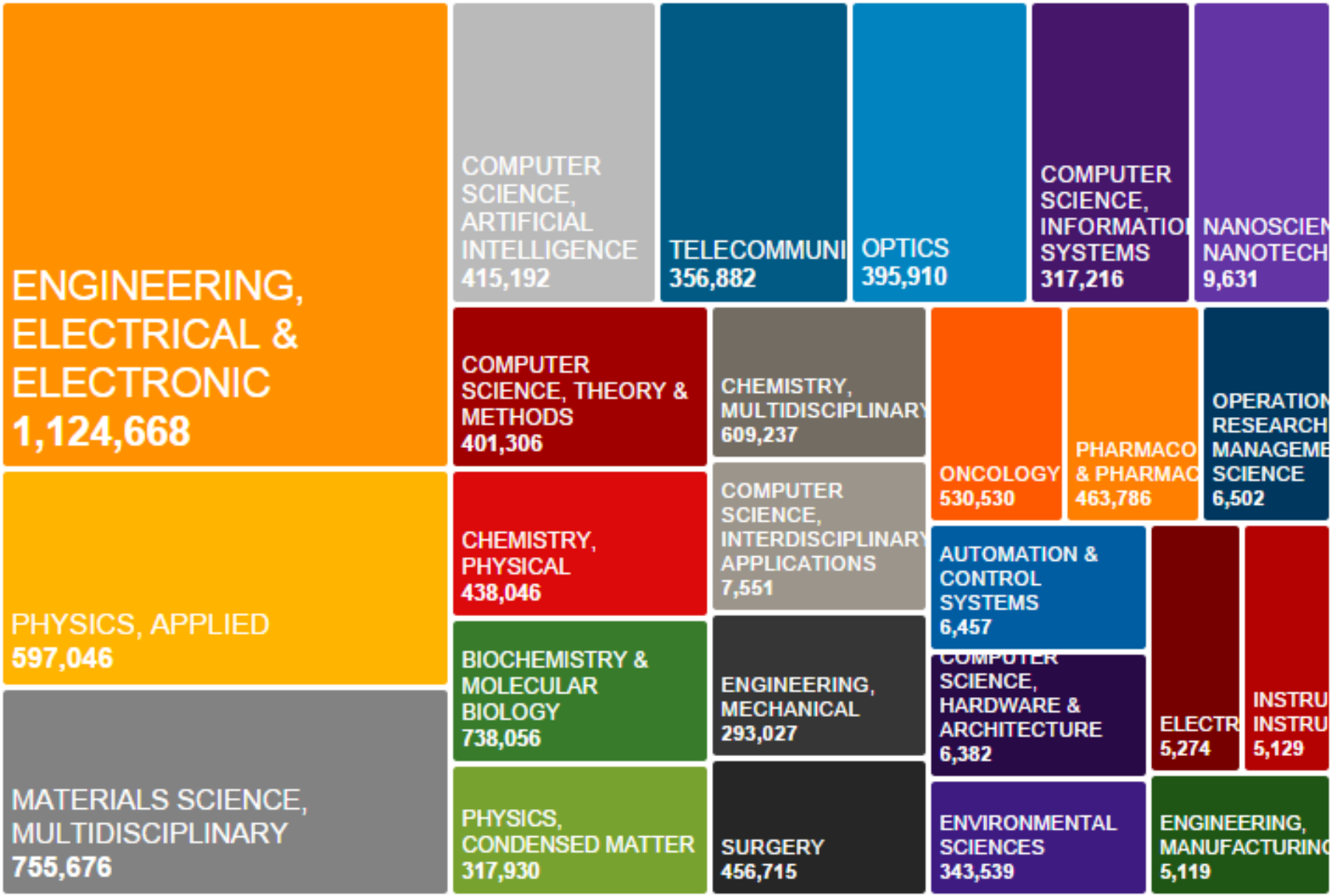
A. NEUROSCIENCE
B. MATERIALS SCIENCE, MULTIDISCIPLINARY
C. ENGINEERING, ELECTRICAL & ELECTRONIC
D. BIOCHEMISTRY & MOLECULAR BIOLOGY

Web of Science Top 25 Categories (2004-2013)



Source: InCites

Web of Science Top 25 Categories in Taiwan (2004-2013)



Source: InCites

Source: Journal Citation Reports

Benchmarking Journals in Specific Category

	Category	Edition	#Journals	Total Cites	Median Impact Factor	Aggregate Impact Factor ▾
1	MULTIDISCIPLINARY SCIENCES	SCIE	55	2,079,971	0.786	5.882
2	CELL BIOLOGY	SCIE	185	1,784,263	3.333	5.815
3	CHEMISTRY, MULTIDISCIPLINARY	SCIE	148	2,195,260	1.401	5.222
4	CELL & TISSUE ENGINEERING	SCIE	18	76,359	3.535	4.940
5	NANOSCIENCE & NANOTECHNOLOGY	SCIE	73	799,992	1.768	4.902
6	NEUROIMAGING	SCIE	13	118,110	2.374	4.732
141	MANAGEMENT	SSCI	172	356,261	1.160	1.699
142	EDUCATION, SPECIAL	SSCI	37	26,278	1.013	1.694
142	PALEONTOLOGY	SCIE	48	77,218	1.168	1.694
144	ENGINEERING, CIVIL	SCIE	124	277,434	0.956	1.670
145	BUSINESS	SSCI	110	257,996	1.388	1.658

Journal rankings and comparisons are meaningful only within each category - not between categories or domains.

Source: Journal Citation Reports

Benchmarking Journals in Specific Category

ENGINEERING, ELECTRICAL & ELECTRONIC

	Full Journal Title	JCR Abbreviated Title	Total Cites	Journal Impact Factor	Impact Factor without Journal Self Cites
1	IEEE WIRELESS COMMUNICATIONS	IEEE WIREL COMMUN	2,910	6.524	6.262
2	IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS	IEEE T IND ELECTRON	24,432	6.500	4.561
3	IEEE TRANSACTIONS ON FUZZY SYSTEMS	IEEE T FUZZY SYST	7,208	6.306	4.548
4	IEEE TRANSACTIONS ON POWER ELECTRONICS	IEEE T POWER ELECTR	17,852	5.726	3.379
5	IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE	IEEE T PATTERN ANAL	27,275	5.694	5.430
6	PROCEEDINGS OF THE IEEE	P IEEE	20,916	5.466	5.403
7	IEEE Industrial Electronics	IEEE IND			
8	PROGRESS IN QUANTUM	PROG QUANT	803	4.688	4.688
9	IEEE SIGNAL PROCESSING	IEEE SIGNAL PROC	5,580	4.481	4.406

- Researchers: decide what to read and where to publish
- Librarians: assess new or current journals in their collection
- Publishers: assess their journals compared to similar journals

Benchmarking Research Performance in Context

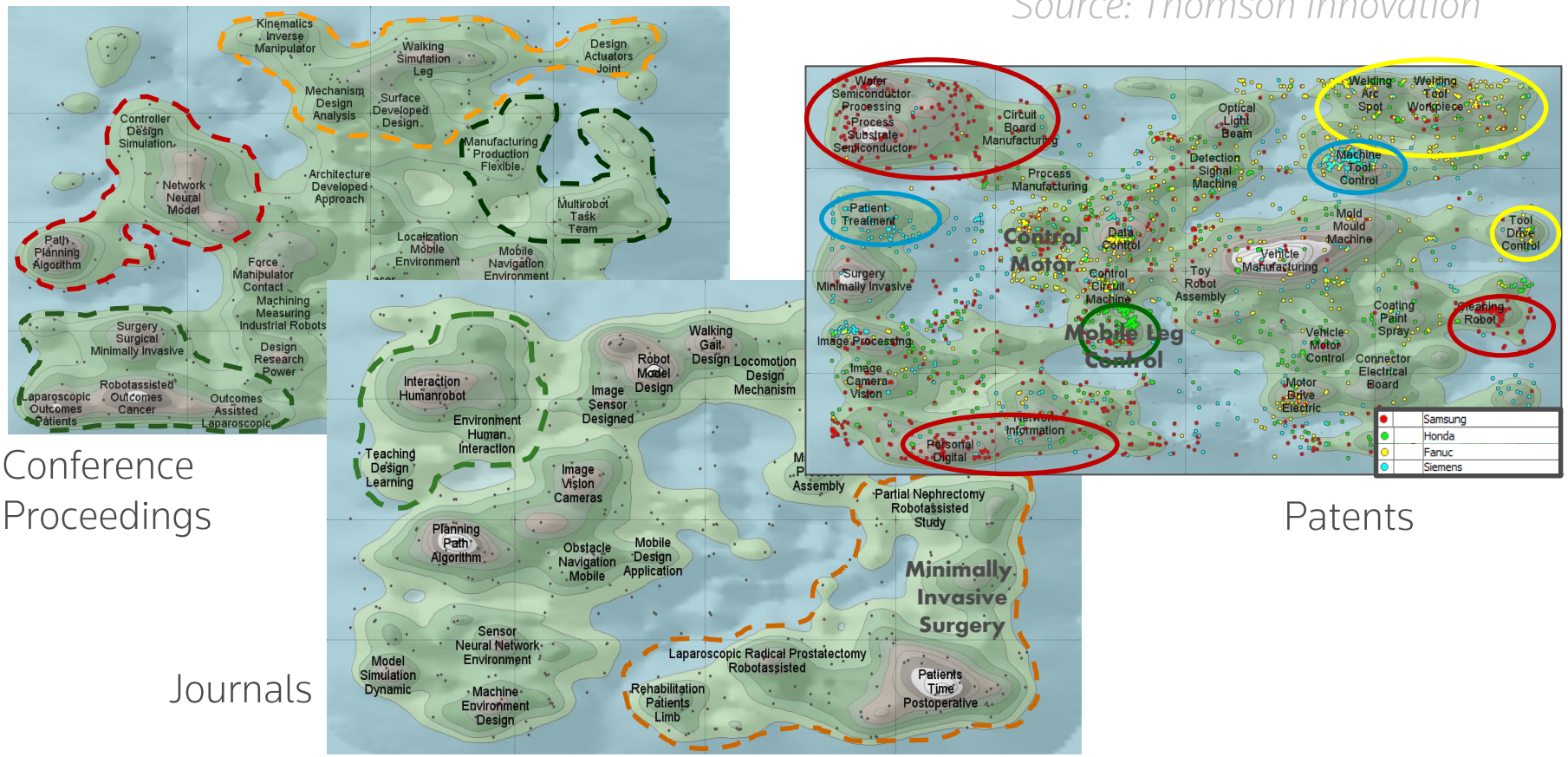
Citation rates vary among fields. What is good or average in mathematics is very different from what is good or average in biochemistry.

RESEARCH FIELDS	2004	2005	2006	2007	2008	2009	2010	2011
ALL FIELDS	22.50	20.77	18.68	16.68	14.30	12.06	9.51	6.71
AGRICULTURAL SCIENCES	16.71	15.45	13.93	11.86	9.49	7.87	6.19	4.25
BIOLOGY & BIOCHEMISTRY	32.89	29.90	26.69	23.33	20.22	17.19	13.19	9.19
CHEMISTRY	21.67	20.86	18.96	16.97	15.77	13.43	11.18	8.23
CLINICAL MEDICINE	26.33	24.62	21.81	18.49	15.74	13.19	10.17	7.11
COMPUTER SCIENCE	7.19	6.75	6.16	8.51	7.41	6.49	4.89	3.35
ECONOMICS & BUSINESS	17.57	16.02	13.82	11.16	8.64	6.98	5.09	3.22
ENGINEERING	10.49	9.42	9.23	8.65	7.45	6.82	5.42	3.90
ENVIRONMENT/ECOLOGY	26.44	24.78	21.59	19.01	16.26	13.06	10.32	7.34
GEOSCIENCES	21.42	19.80	18.38	15.47	13.74	11.83	8.97	6.56
IMMUNOLOGY	38.29	36.43	31.95	28.71	24.39	20.53	15.69	11.04
MATERIALS SCIENCE	15.03	14.77	13.70	13.78	11.96	10.67	9.07	6.70
MATHEMATICS	8.08	7.26	6.57	5.86	5.12	4.25	3.35	2.26
MICROBIOLOGY	31.66	29.54	25.92	22.65	19.10	16.26	12.90	8.42

Source: Essential Science Indicators

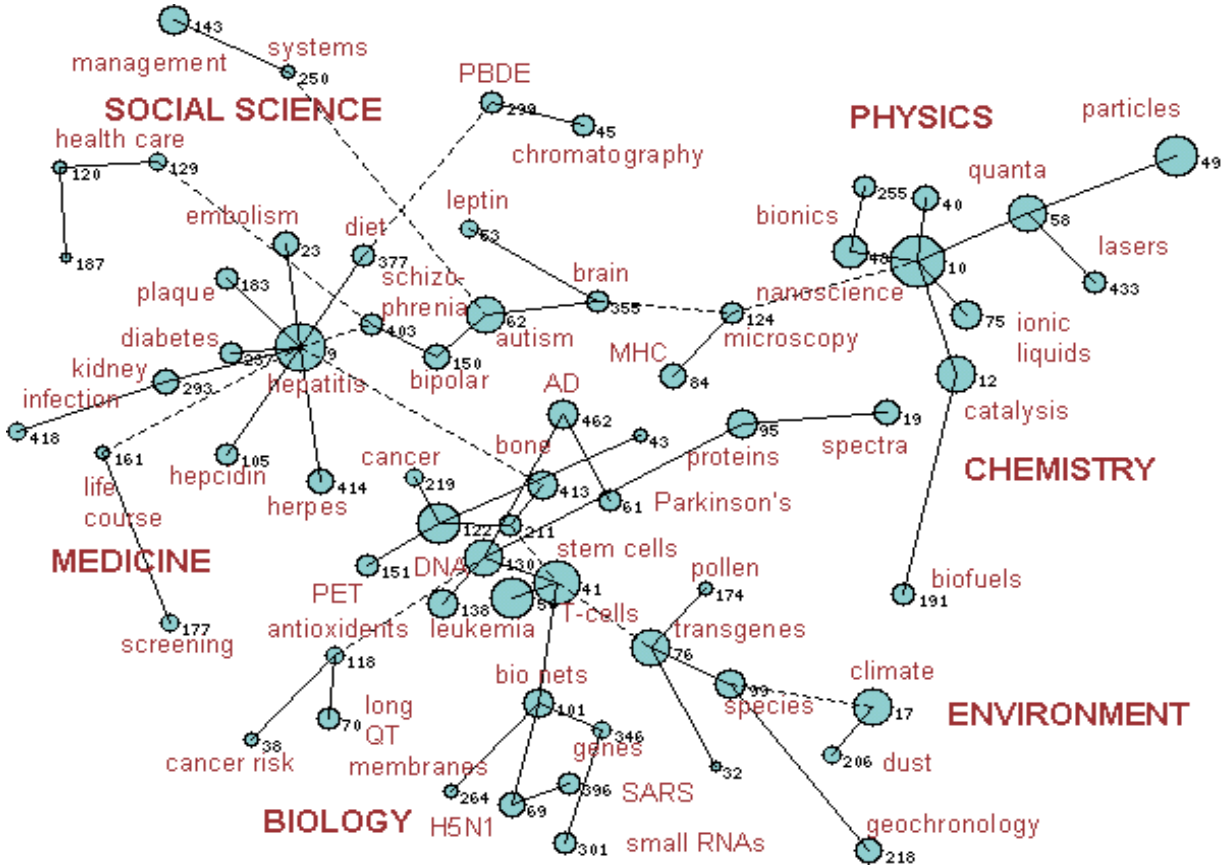
CLUSTER THE RELATED DOTS
By Text Mining and Clustering

Source: Thomson Innovation



LITERATURE AND PATENT LANDSCAPING – Robotics

CLUSTER THE RELATED DOTS
By Citations

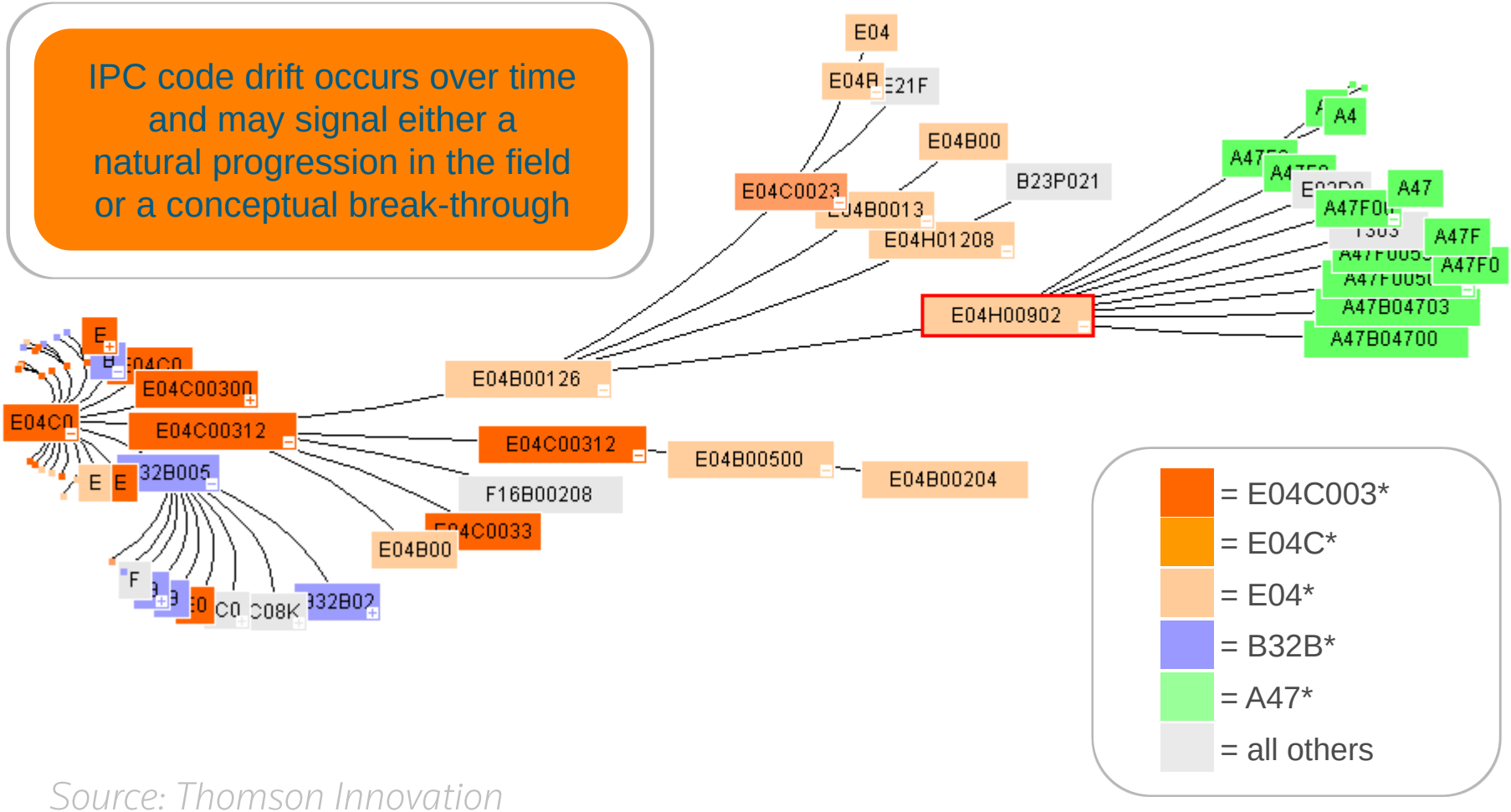


Map of the Structure of Global Science (May 2010)

Source: sciencewatch.com

Patent Citations: Technology Break-through

IPC code drift occurs over time and may signal either a natural progression in the field or a conceptual break-through



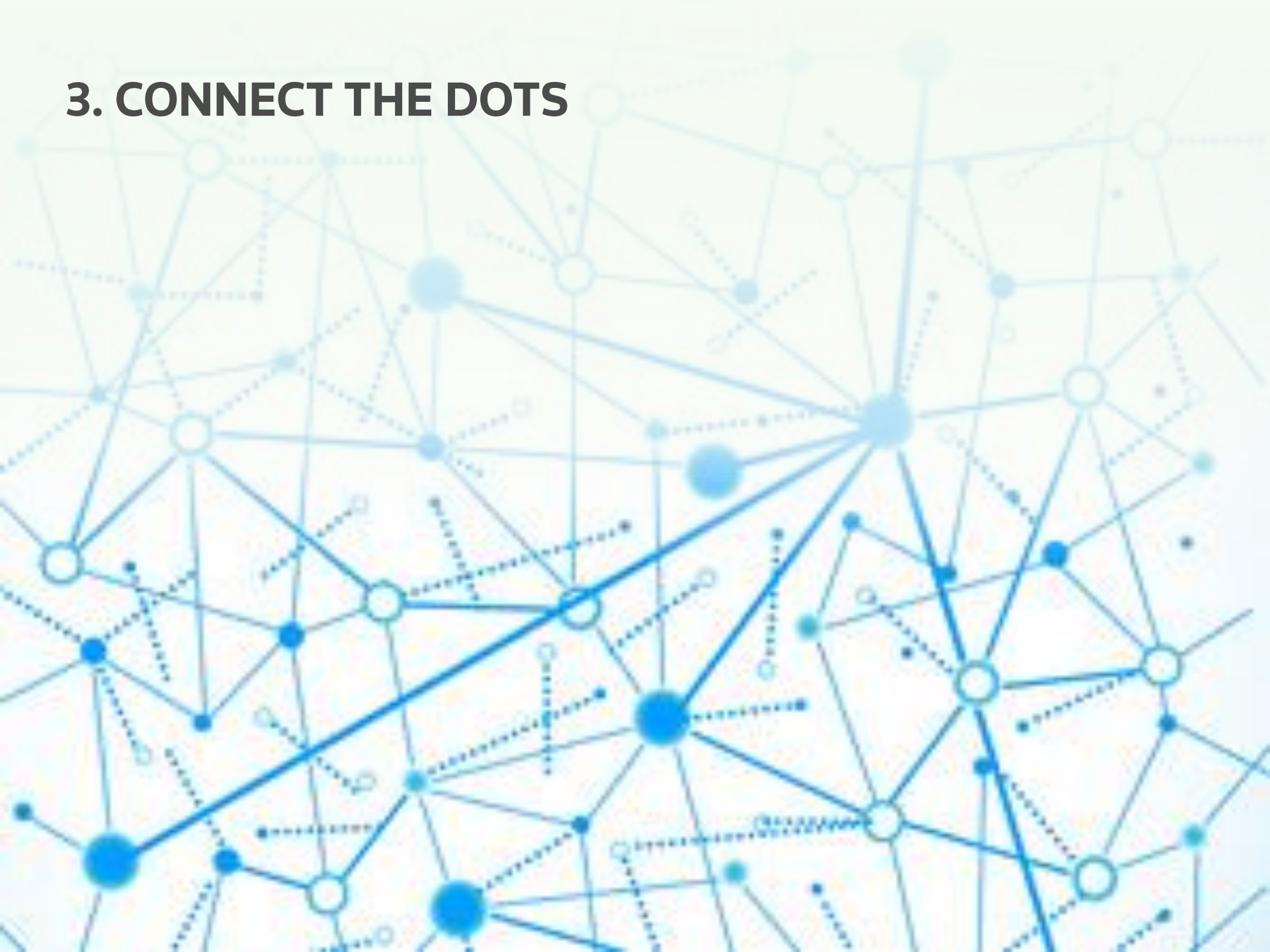
Source: Thomson Innovation

2. CLUSTER THE RELATED DOTS :

Putting the data into context

- By Category
- By Text Mining and clustering
- By Citation

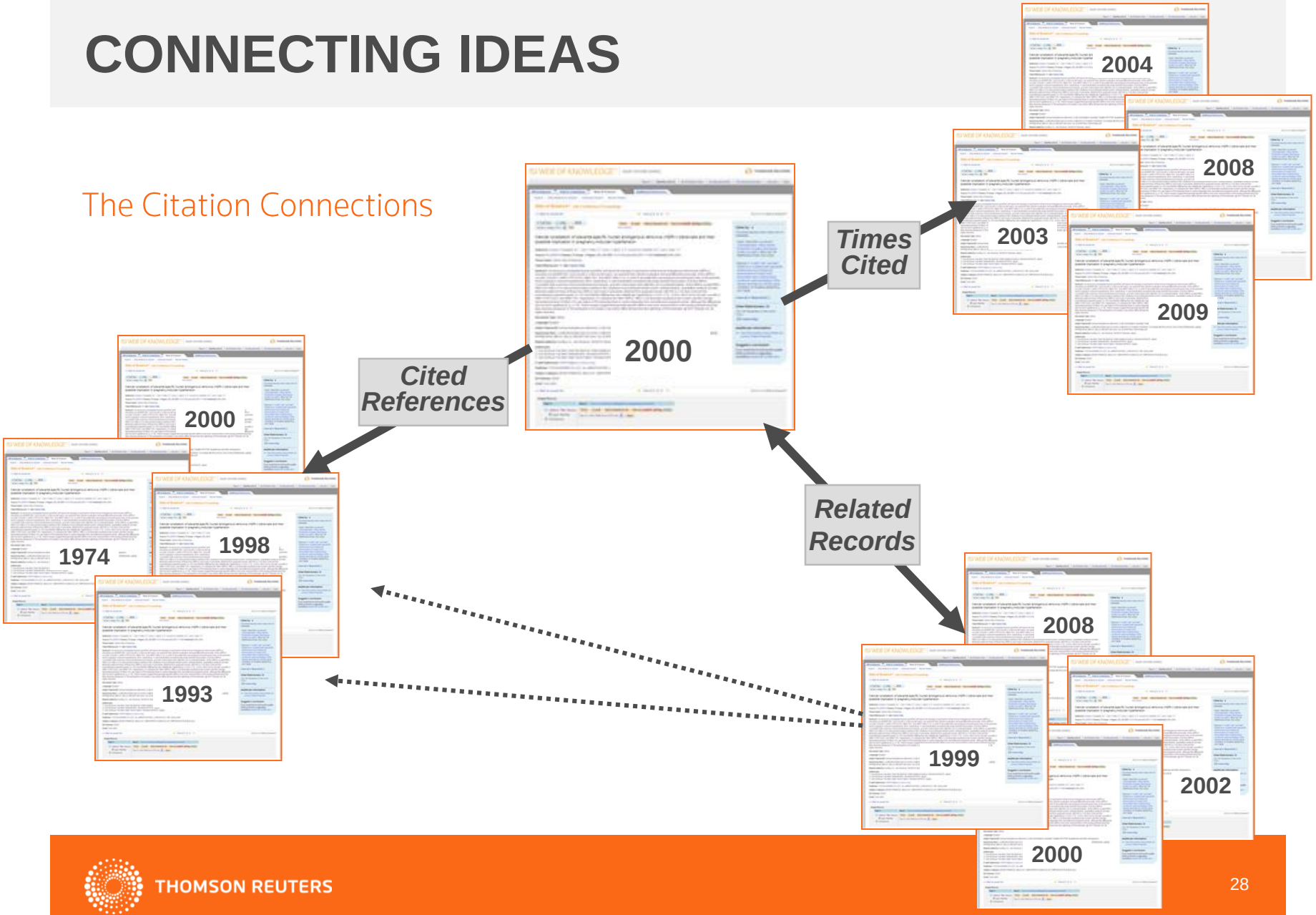
3. CONNECT THE DOTS



HOW?

CONNECTING IDEAS

The Citation Connections

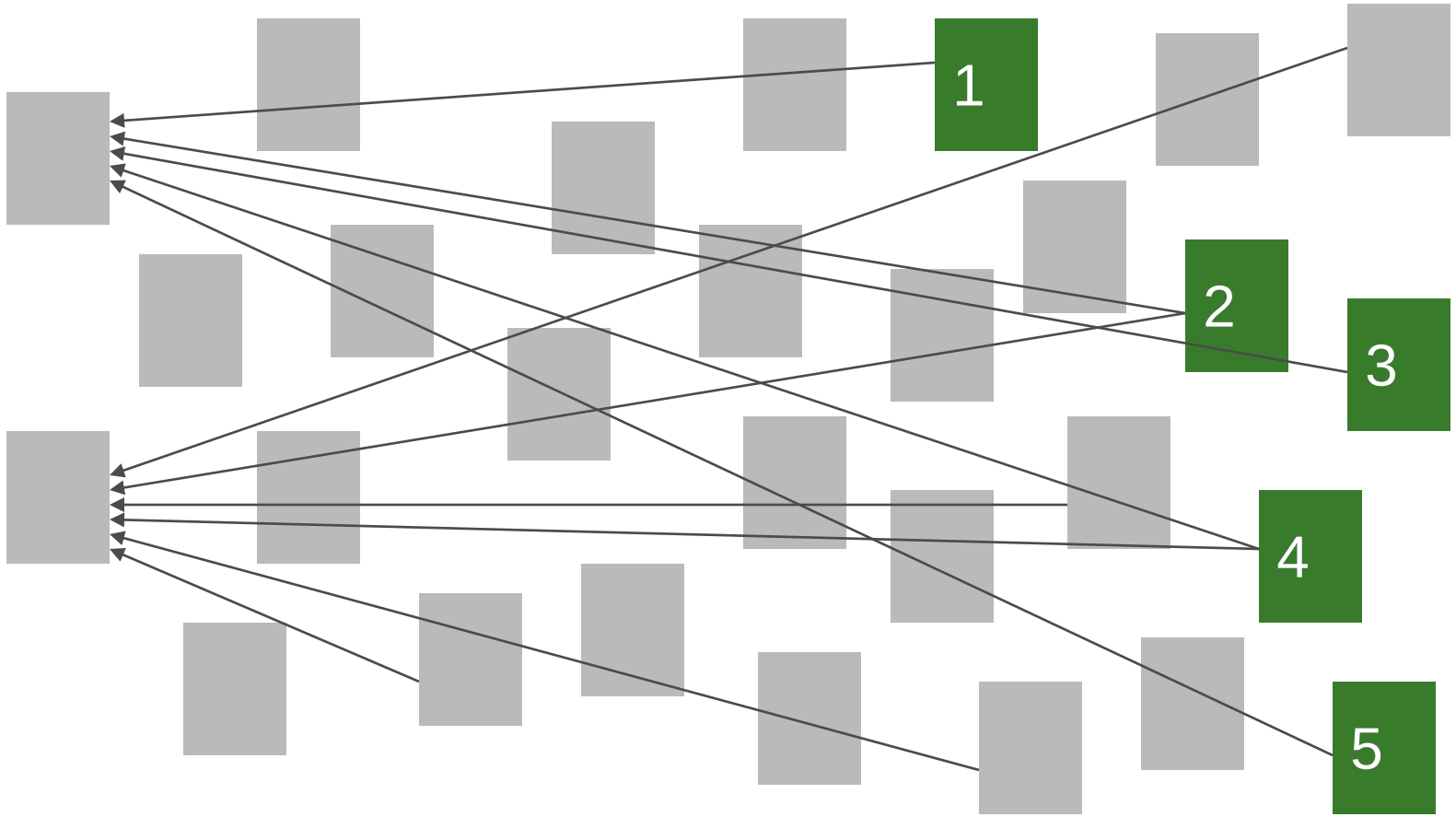


KEY BENEFIT OF CITATION IS NOT JUST ONLY TO ACQUIRE OR RETRIEVE BUT TO DISCOVER!

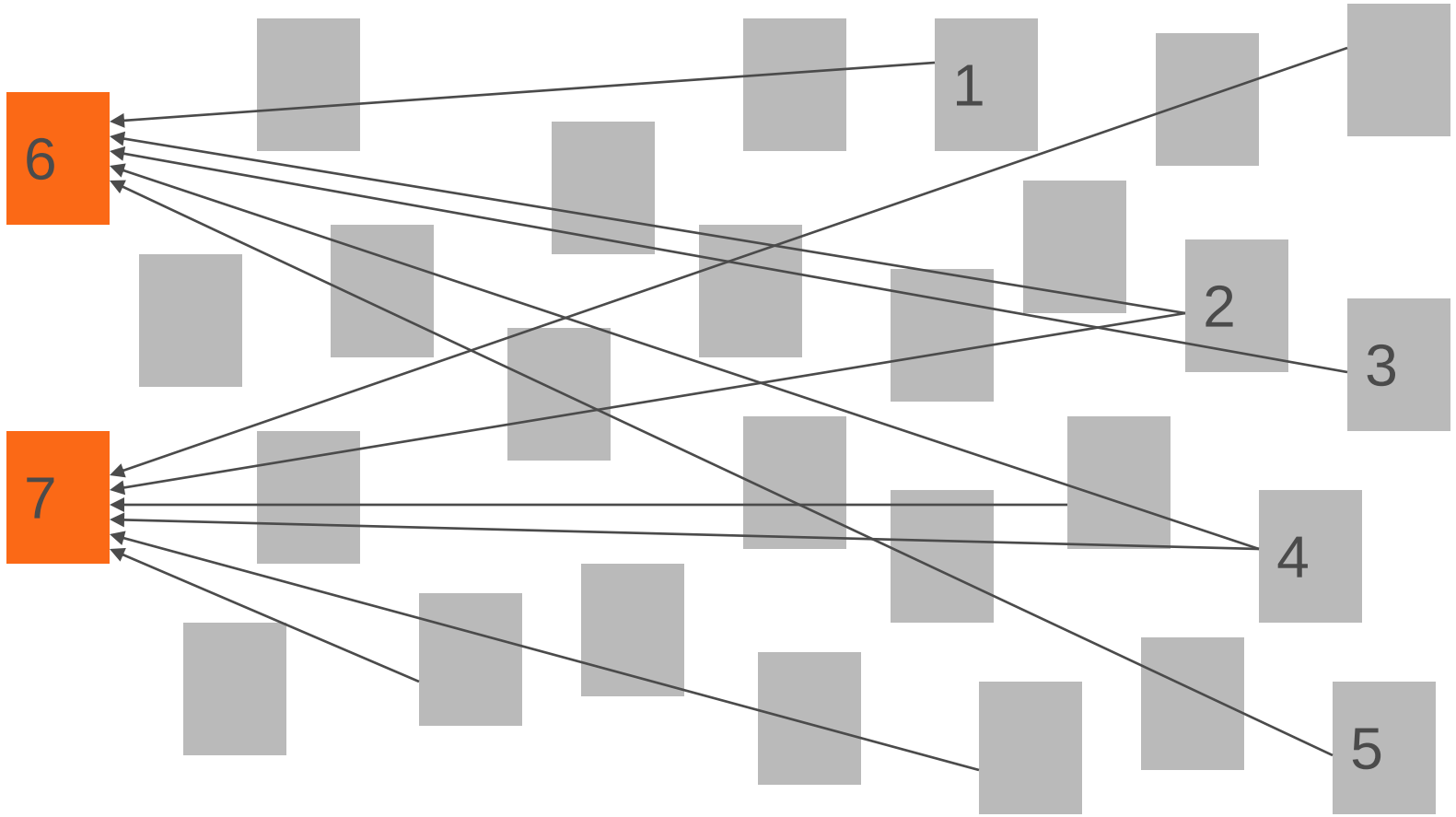
- To Acquire is straightforward, all researchers are doing it … I know what I need to know
- To Retrieve is common, any researcher can do it … I know what I don't know
- But … To Discover is challenging, you need to discover what others don't know … things that you didn't even know you don't know

Citations provides the "discovery links" that are essential to link up known ideas to discover unknown concepts

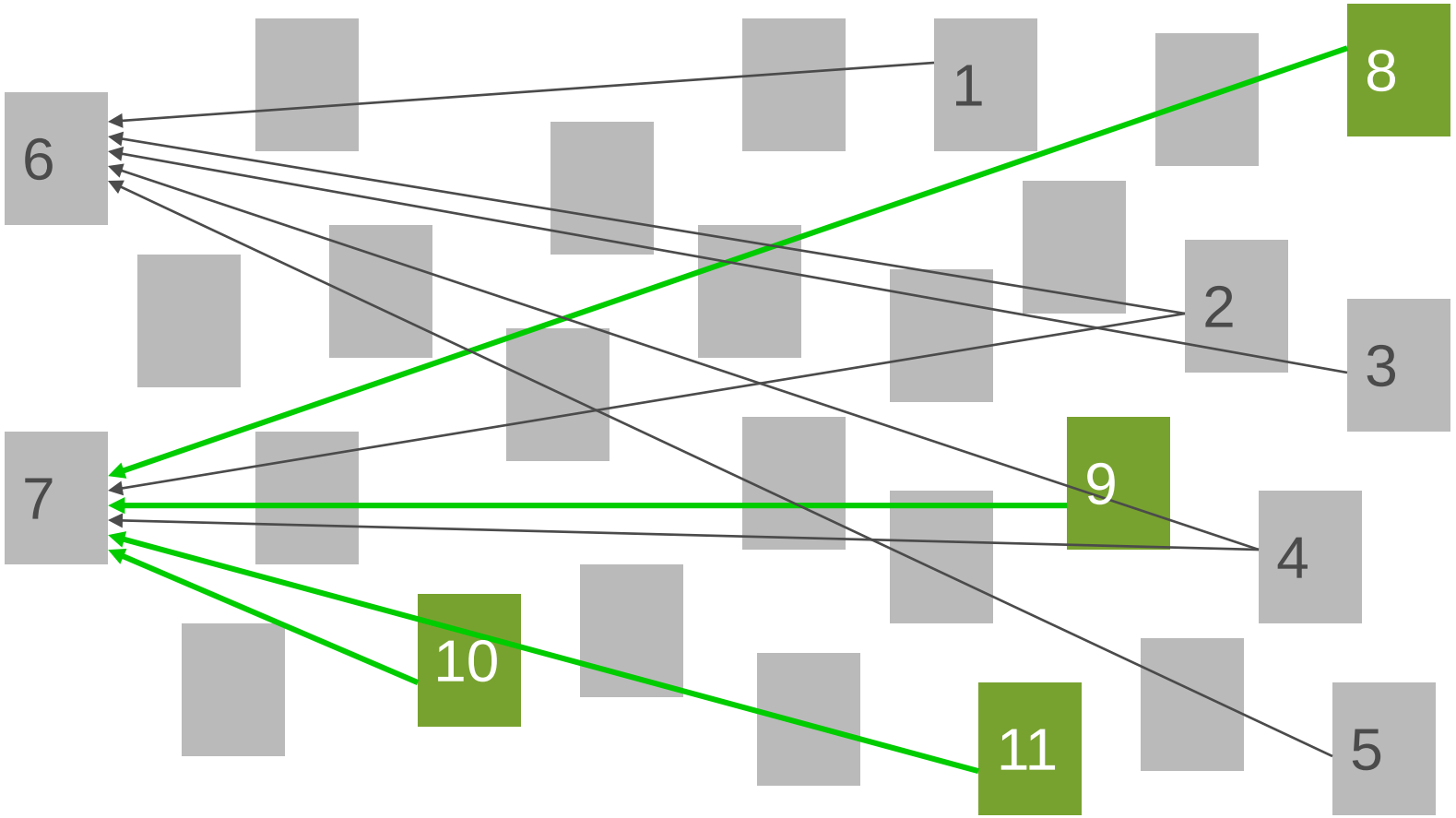
I know what I need to know – you are only acquiring information – any researcher can do it (1,2,3,4,5)



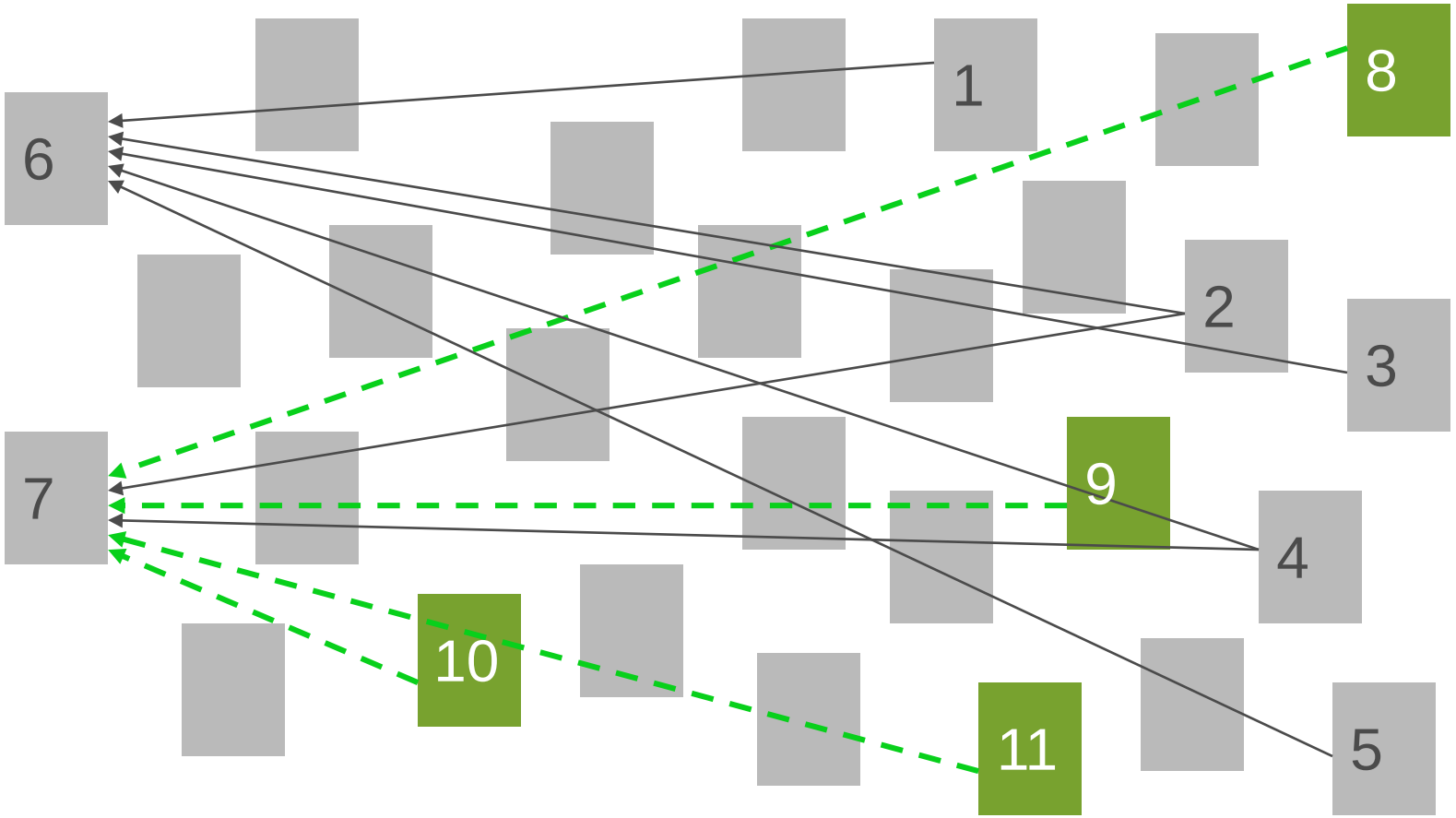
I know what I don't know – you are only retrieving information – every researcher is doing it (6,7)



If you could find what are those things that you didn't even know you don't know, you DISCOVER! (8,9,10,11)

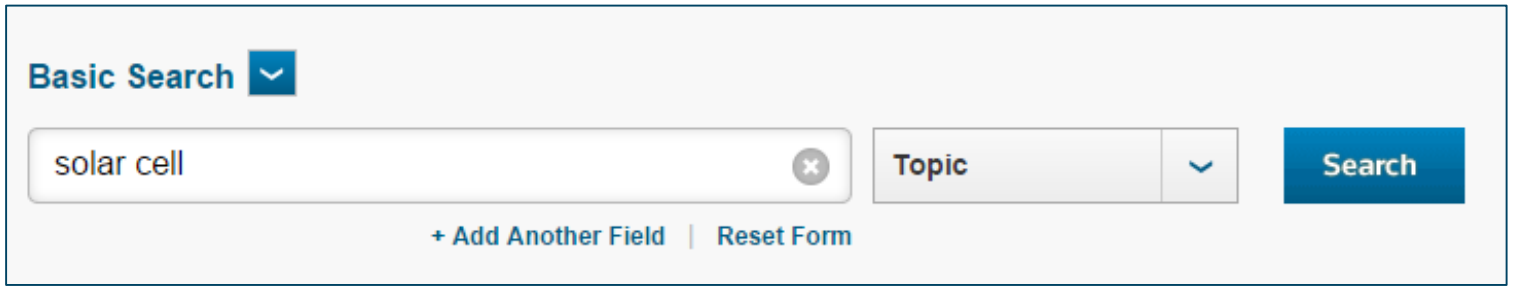


Without the "discovery links" and you lost the related materials (8,9,10,11)



CONNECTING COLLABORATORS

Source: Web of Science, Thomson Reuters



Conferences to attend ?

Field: Conference Titles
39TH IEEE PHOTOVOLTAIC SPECIALISTS CONFERENCE PVSC
246TH NATIONAL MEETING OF THE AMERICAN CHEMICAL SOCIETY ACS
PHYSICS SIMULATION AND PHOTONIC ENGINEERING OF PHOTOVOLTAIC DEVICES III
CONFERENCE ON PHOTONICS FOR SOLAR ENERGY SYSTEMS V
15TH INTERNATIONAL CONFERENCE ON GETTERING AND DEFECT ENGINEERING IN SEMICONDUCTOR TECHNOLOGY GADEST
CONFERENCE ON NEXT GENERATION NANO PHOTONIC AND CELL TECHNOLOGIES FOR SOLAR ENERGY CONVERSION IV
CONFERENCE ON MICRO NANO MATERIALS DEVICES AND SYSTEMS
3RD INTERNATIONAL CONFERENCE ON CRYSTALLINE SILICON PHOTOVOLTAICS SILICONPV
2ND INTERNATIONAL CONFERENCE ON ENERGY AND ENVIRONMENTAL PROTECTION ICEEP 2013
2ND INTERNATIONAL CONFERENCE ON INNOVATION COMMUNICATION AND ENGINEERING ICICE

Potential Collaborators ?

Field: Authors	Record Count	% of 76003
GRATZEL M	491	0.646 %
GREEN MA	374	0.492 %
KIM J	308	0.405 %
KIM JH	305	0.401 %
YAMAGUCHI M	289	0.380 %
ANONYMOUS	279	0.367 %
YANG Y	277	0.364 %
ZHANG Y	269	0.354 %
LI YF	257	0.338 %
KIM H	249	0.328 %

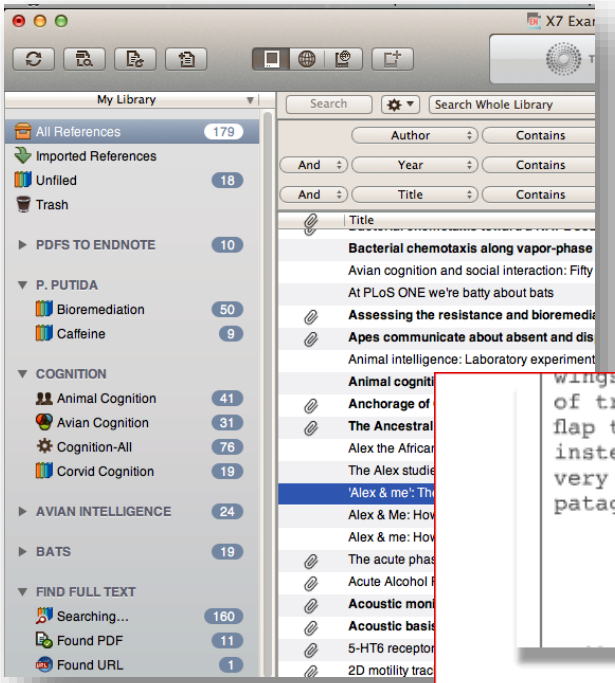
Institutes to collaborate ?

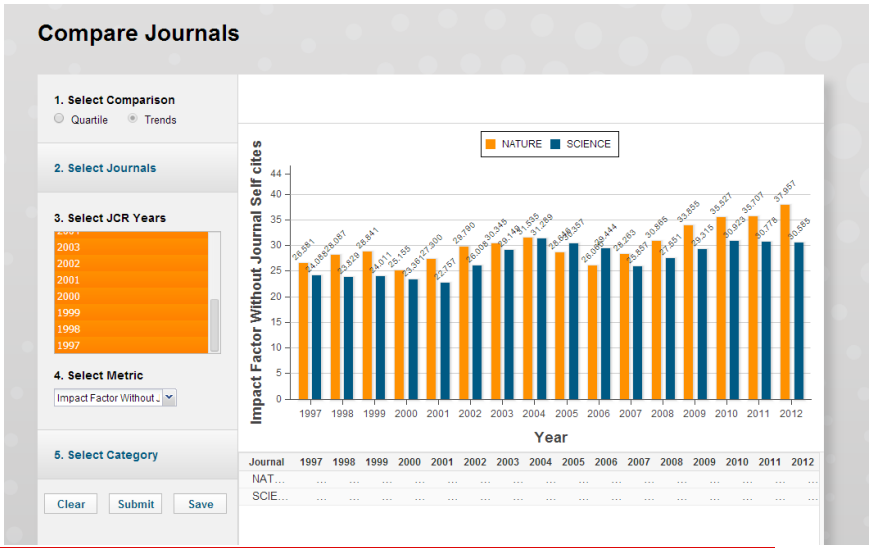
Field: Organizations-Enhanced	Record Count	% of 76003
CHINESE ACADEMY OF SCIENCES	2376	3.126 %
NATIONAL RENEWABLE ENERGY LABORATORY USA	1562	2.055 %
UNIVERSITY OF CALIFORNIA SYSTEM	1543	2.030 %
HELMHOLTZ ASSOCIATION	1046	1.376 %
NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE TECHNOLOGY AIST	898	1.182 %
SWISS FEDERAL INSTITUTE OF TECHNOLOGY LAUSANNE	799	1.051 %
UNIVERSITY OF NEW SOUTH WALES	782	1.029 %
FRAUNHOFER GESELLSCHAFT	778	1.024 %
NATIONAL TAIWAN UNIVERSITY	766	1.008 %
OSAKA UNIVERSITY	684	0.900 %

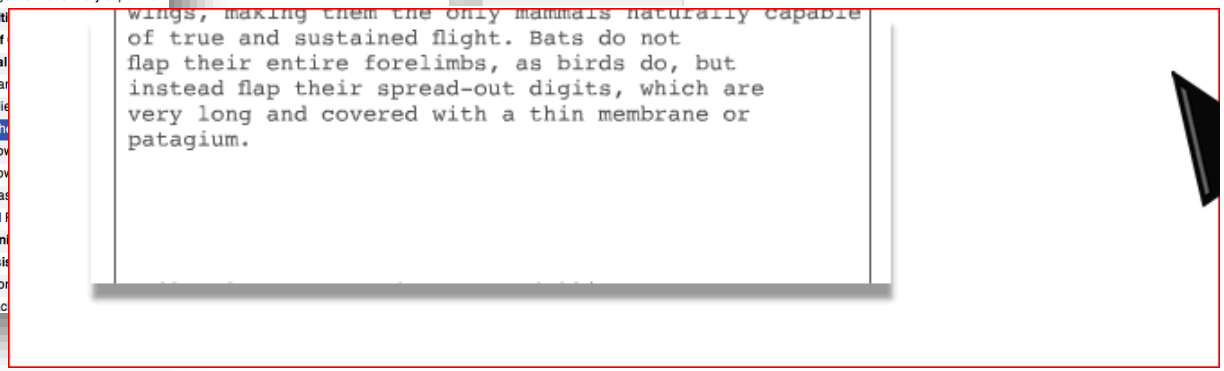
CONNECTING RESEARCH TO PUBLICATION

Source: EndNote, Journal Citation reports

From managing research information to paper writing to submission







CONNECTED RESEARCHERS' WORKFLOW