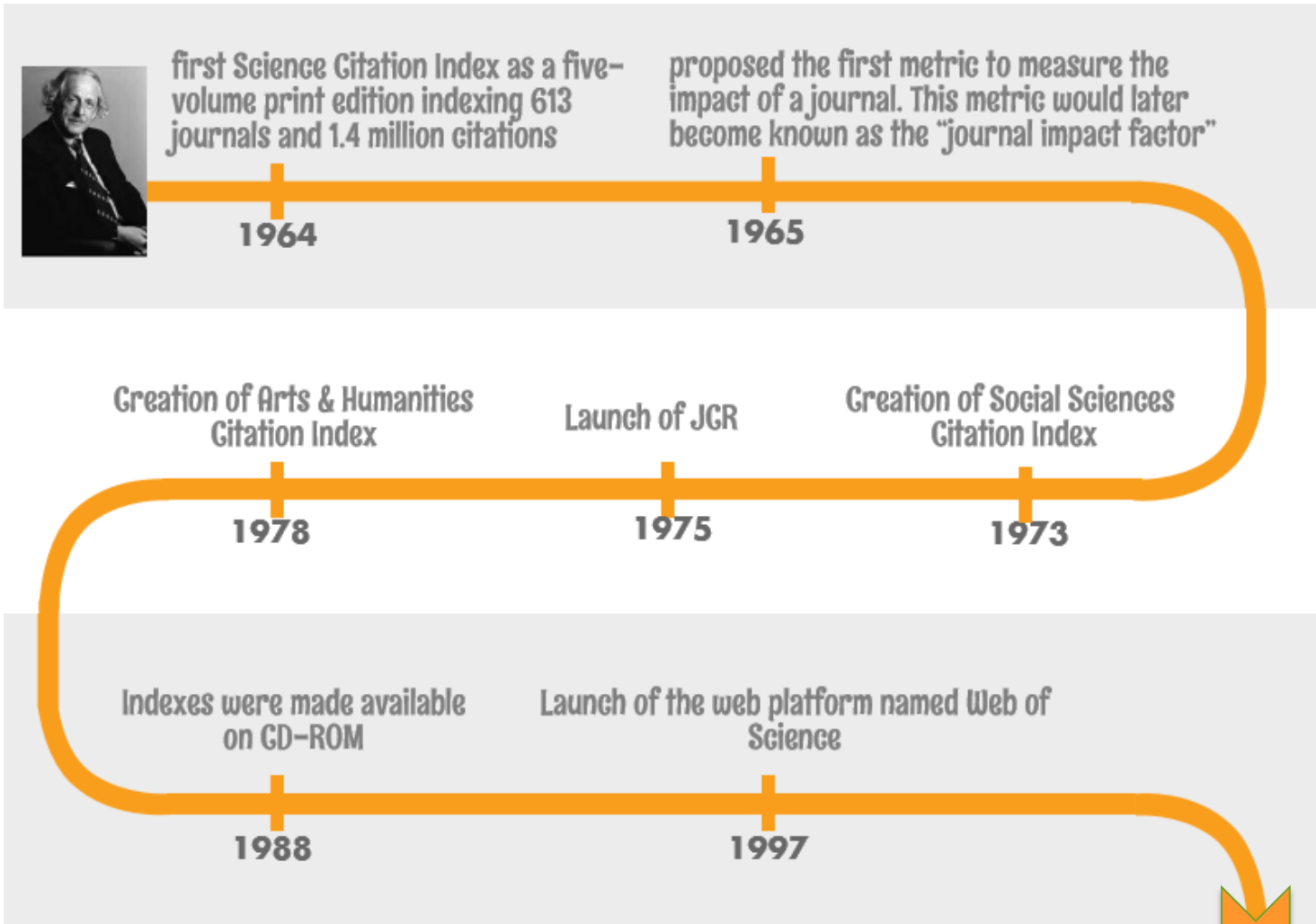# Web of Science™

The world's most trusted citation index covering the leading scholarly literature.

MASSIMILIANO CARLONI
September 2016
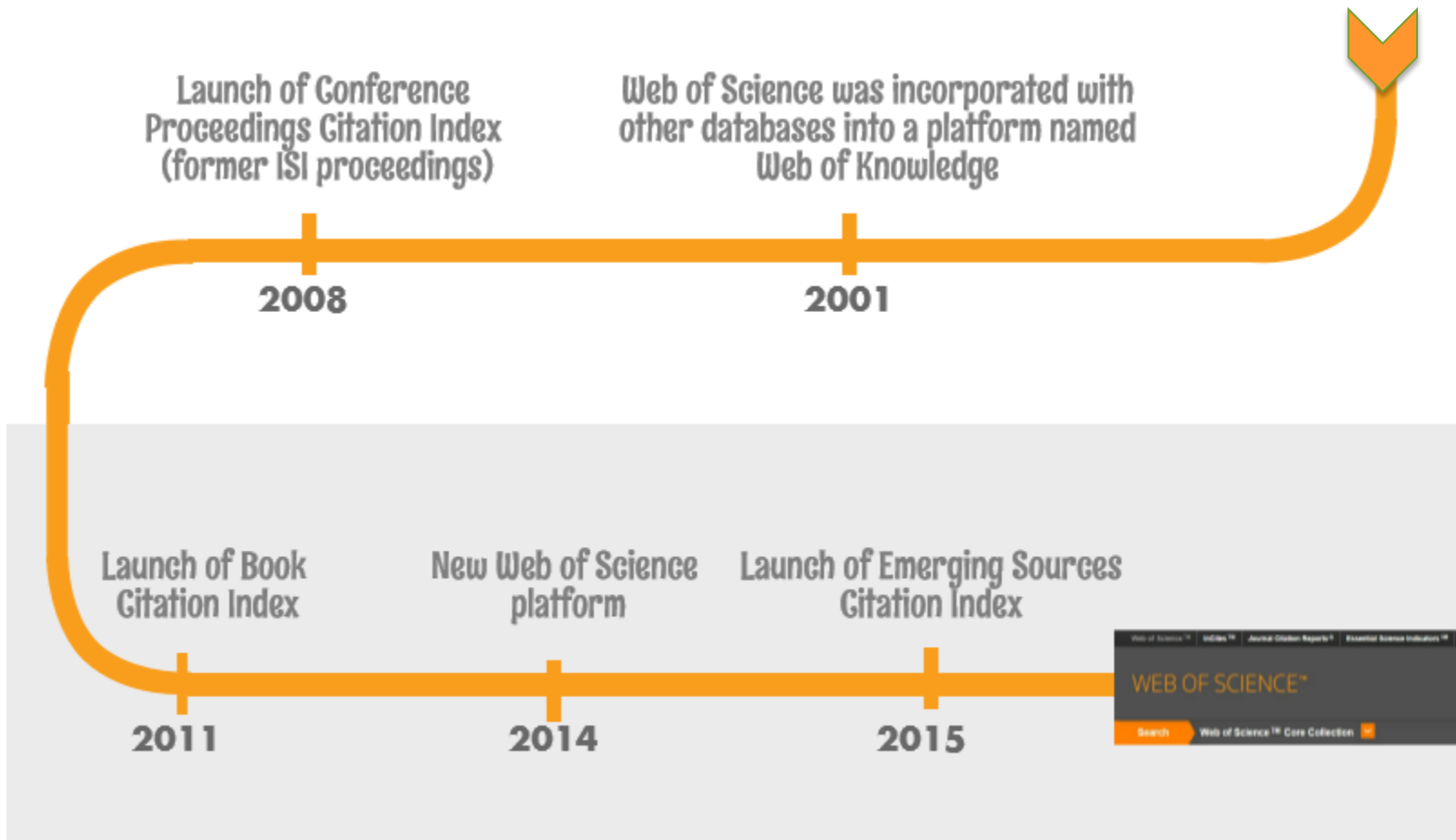
THOMSON REUTERS™

Web of Science: a bit of history

first Science Citation Index as a five-volume print edition indexing 613 journals and 1.4 million citations

proposed the first metric to measure the impact of a journal. This metric would later become known as the "journal impact factor"

1964

1965

Creation of Arts & Humanities Citation Index

Launch of JGR

Creation of Social Sciences Citation Index

1978

1975

1973

Indexes were made available on CD-ROM

Launch of the web platform named Web of Science

1988

1997

THOMSON REUTERS

Web of Science: a bit of history

Launch of Conference
Proceedings Citation Index
(former ISI proceedings)

Web of Science was incorporated with
other databases into a platform named
Web of Knowledge

2008

2001

Launch of Book
Citation Index

New Web of Science
platform

Launch of Emerging Sources
Citation Index

2011

2014

2015

WEB OF SCIENCE™

THOMSON REUTERS

Web of Science Core Collection

Science Citation Index Expanded (SCI-E) 1900

Social Sciences Citation Index (SSCI) 1900

Arts & Humanities Citation Index (AHCI) 1975

12,812 journals

~9 million records

Conference Proceedings Citation Index (CPCI) 1990
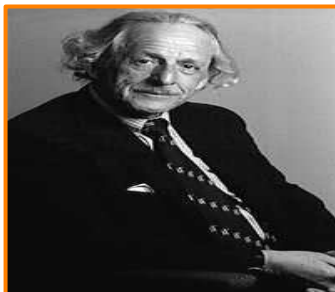
Current Chemical Reactions (CCR) 1985

63 million records

Index Chemicus (IC) 1993

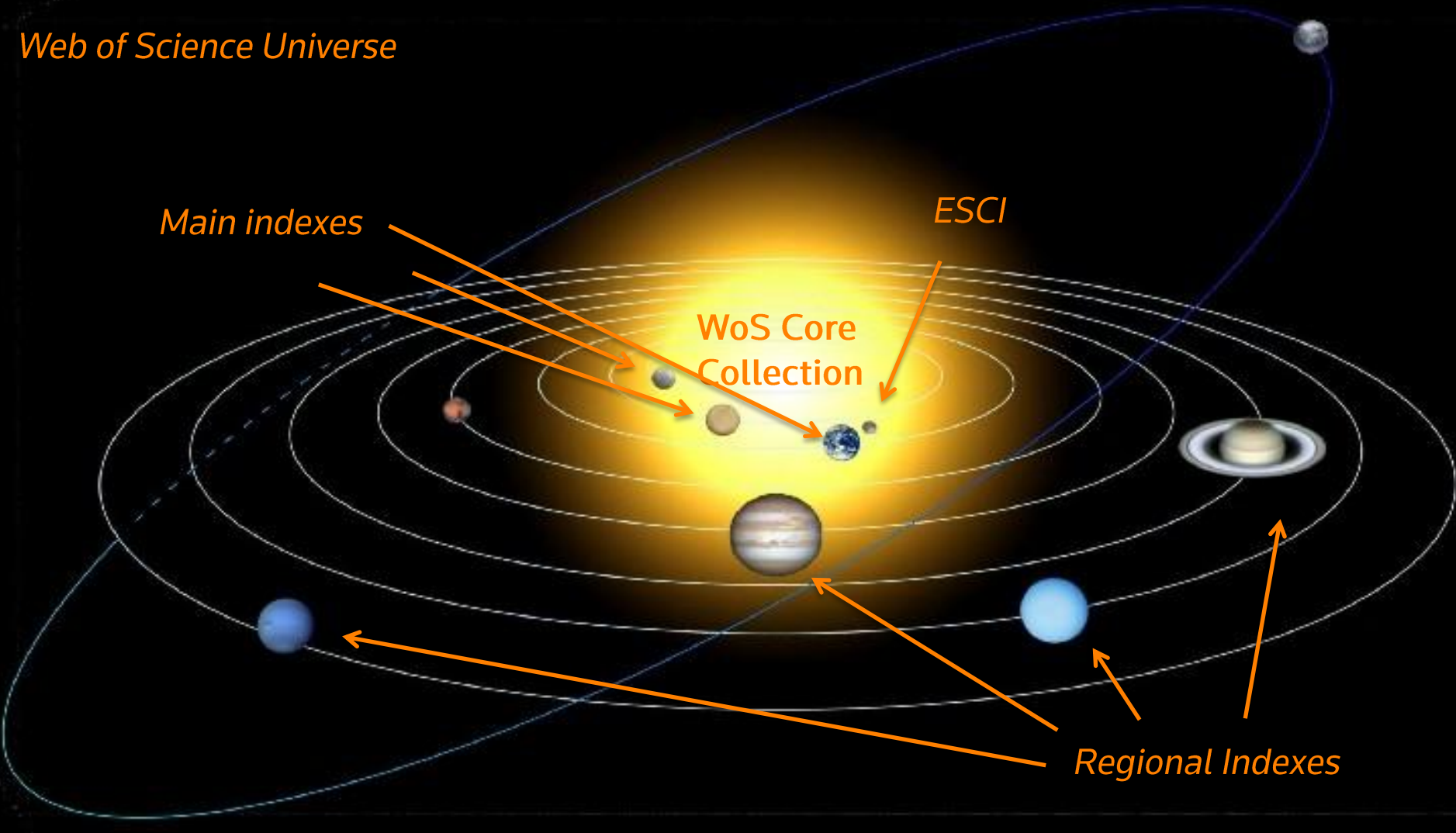**Book Citation Index (BkCI) 2005**

Emerging Sources Citation Index (ESCI) 2015

3,750 indexed journals

75k books ~1 million bk chapters

THOMSON REUTERS

Coverage. Building and expanding the Core/Platform

Web of Science Universe

Main indexes

ESCI

WoS Core Collection

Regional Indexes
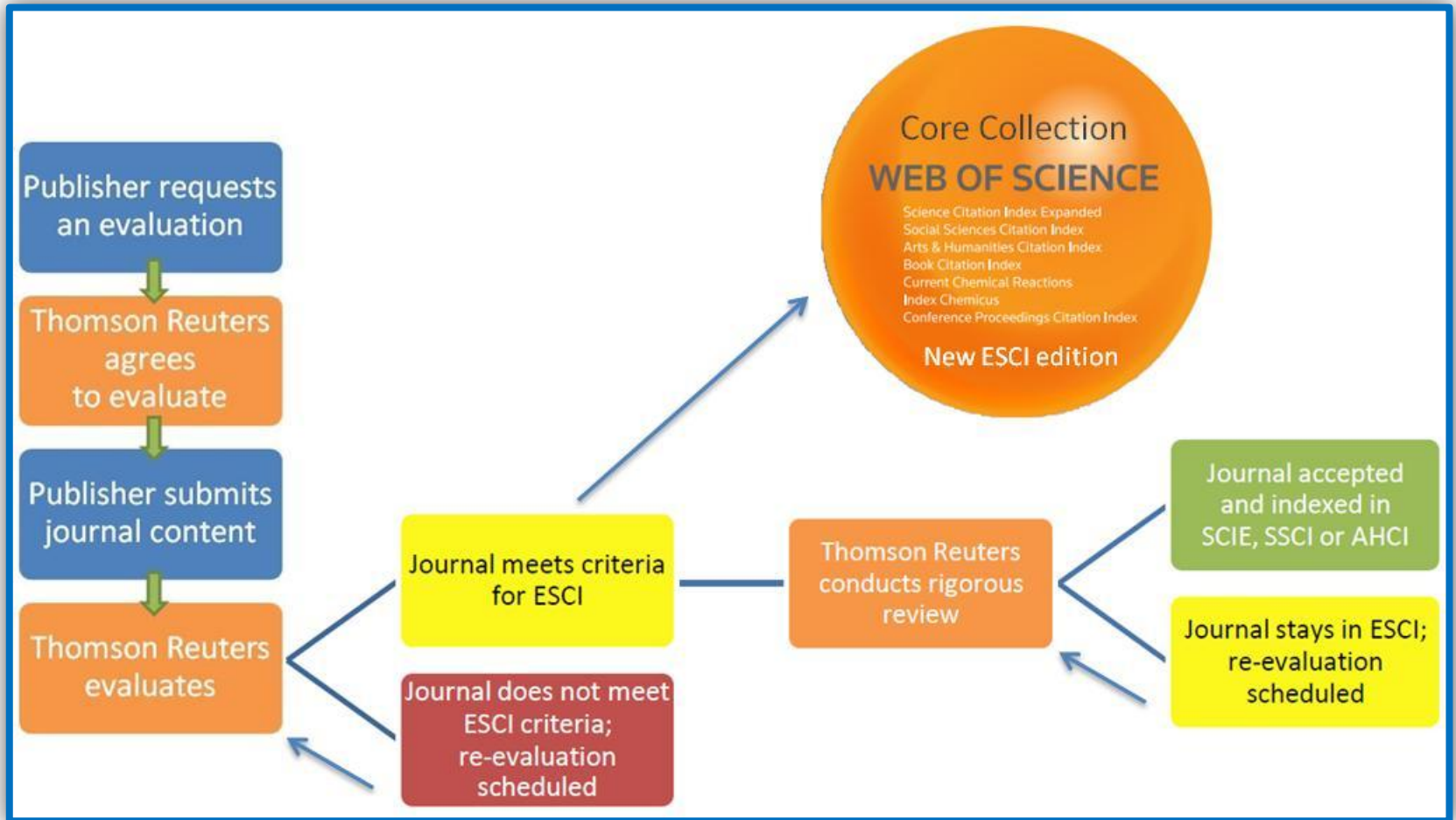
THOMSON REUTERS

Coverage. Emerging Sources Citation Index

Creation of a new edition within Web of Science Core Collection with some limitations

- All content must meet key selection criteria
- Same features and rigorous indexing standards as other Core Collection editions
- Cover-to-cover indexing (authors, institutions, etc.)
- True Citation Index, capturing times cited and cited references
- Content begins in 2015

Evidence of peer review

Adherence to ethical publishing practices

Content requested by or of high interest to subscribers

Preferred electronic formats (XML / PDF)

THOMSON REUTERS

Coverage. New Process

Coverage. New Selection Criteria

Journal Publishing Standards → **Editorial Content** → **International Focus** → **Citation Analysis**

Journal Publishing Standards

- **Peer review**
- **Ethical publishing practices**
- **Meets technical requirements (XML / PDF)**
- Timeliness of publication
- **International editorial conventions**
- **English-language bibliographic information**

Editorial Content

- **Has a scholarly audience searched for or requested this content?**
- **How does this journal compare with covered journals of similar scope?**
- **Is this subject already well covered?**
- **Will this journal enrich WoS with novel content?**

International Focus

- **Does this journal target an international audience or specifically a regional audience?**
- **Is international representation among authors and board members at an appropriate level for such a journal?**

Citation Analysis

- Total citations
- Recent citation activity
- Author and editorial board members' citations in the literature
- Integration of the journal into the literature over time

Black = SCIE/SSci/AHCI selection

Orange = ESCI selection requirements

THOMSON REUTERS

Coverage. Regionalization

Creation of partnerships with local community to host their curated datasets on WoS regional indexes

- Russian Science Citation Index
- eLIBRARY.RU
- > 500 journals

- Chinese Science Citation Database
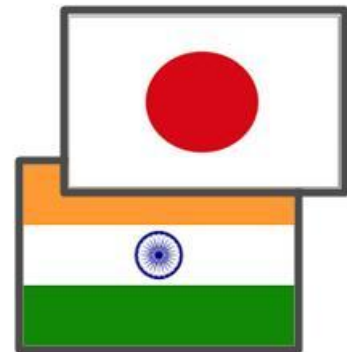- Chinese Academy of Science
- ~ 1,200 journals

- SciELO Citation Database
- FAPESP-BIREME
- ~ 650 journals

- Korean Journal Database
- National Research Foundation of Korea
- ~ 1500 journals

COMING

THOMSON REUTERS

Web of Science – other databases

Results: 136,478,965
(from All Databases)

You searched for: YEAR
PUBLISHED: (1800-2016) ...More

Databases | Refine | Exclude | Cancel | Sort these by: **Record Count** ▾

The first 100 Databases (by record count) are shown. Analyze Results feature not available for over 50,000

- [ ] Web of Science ™ Core Collection (62,507,114)
- [ ] Derwent Innovations Index ℠ (30,474,699)
- [ ] MEDLINE® (25,881,436)
- [ ] BIOSIS Previews® (25,245,640)
- [ ] BIOSIS Citation Index ℠ (25,241,317)
- [ ] Current Contents Connect® (22,998,113)

- [ ] Inspec® (17,018,345)
- [ ] Biological Abstracts® (16,580,859)
- [ ] CABI (10,402,749)
- [ ] Data Citation Index ℠ (6,577,877)
- [ ] Chinese Science Citation Database ℠ (4,510,701)

- [ ] Zoological Record® (4,153,506)
- [ ] FSTA® - the food science resource (1,248,406)
- [ ] KCI-Korean Journal Database (1,131,005)
- [ ] SciELO Citation Index (511,582)
- [ ] Russian Science Citation Index (374,707)

Results: 108,314,376
(from All Databases)

You searched for: YEAR
PUBLISHED: (1800-2016) ...More

→ **All content excluding hosted databases**

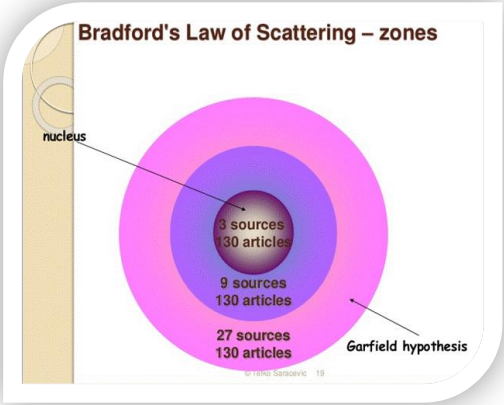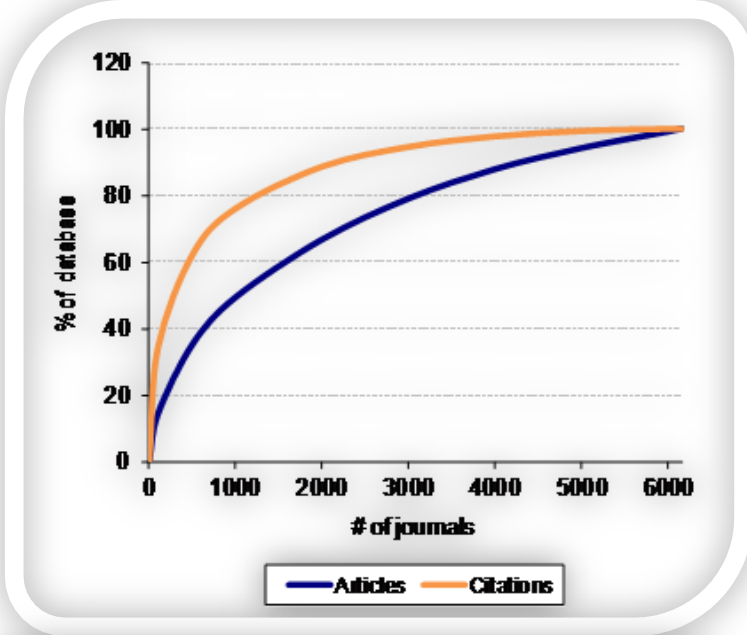Results: 72,101,612
(from All Databases)

You searched for: YEAR
PUBLISHED: (1800-2016) ...More

→ **All free content bundled with WoS-CC**

THOMSON REUTERS

Selectivity. Scientific Theory

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

	Journals	Aggregate IF
2003	2,000	~7,000
	~25%	~80%
2013	3,000	~14,000
	~25%	~70%

Bradford's Law of Scattering – zones

nucleus
3 sources 130 articles
9 sources 130 articles
27 sources 130 articles
Garfield hypothesis

*As a result, the core literature of **all** scientific disciplines involved a group of not more than 1,000 journals and may even possibly be reduced to as few as 500 journals (1972). A relatively small number of journals account for the bulk of what is published and what is cited.*
- E. Garfield

THOMSON REUTERS

Selectivity. Practical Reason

There are too many journals and digital publishing allowed market access to low level journals

N° of Peer Reviewed Journals	**72,876**
	Ulrich's- March 2016
Number of Predatory Journals	**975**
	Jeffrey Beall -March 2016
Number of Fraudolent Papers	**25,500***

*How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data Fanelli – PlosOne 2009

* Ulrich's web

THOMSON REUTERS

Selectivity. Who are the editors

Uniformity of judgement
Continuous check during different process steps. **Monitoring current content**

Around 150 years of experience in the role. 12 Master's degree. **Full time job**

18 Full Time Editors

Bi-weekly meetings

Thomson Reuters employees

EDITORIAL TEAM

No one of the editors edit a journal

12 main languages covered with fluency

No one of the editors publish

No **conflict of interest**

No national boards. Regionalism covered with a different strategy. **See REGIONAL INDEXES**

THOMSON REUTERS

Selectivity. The Results

2015 WoS CC Adds/Deletes by Edition

	SCIE	SSCI	AHCI	Total WoS CC Unique
# Adds	156	67	27	227
# Deletes	32	10	4	43

THOMSON REUTERS
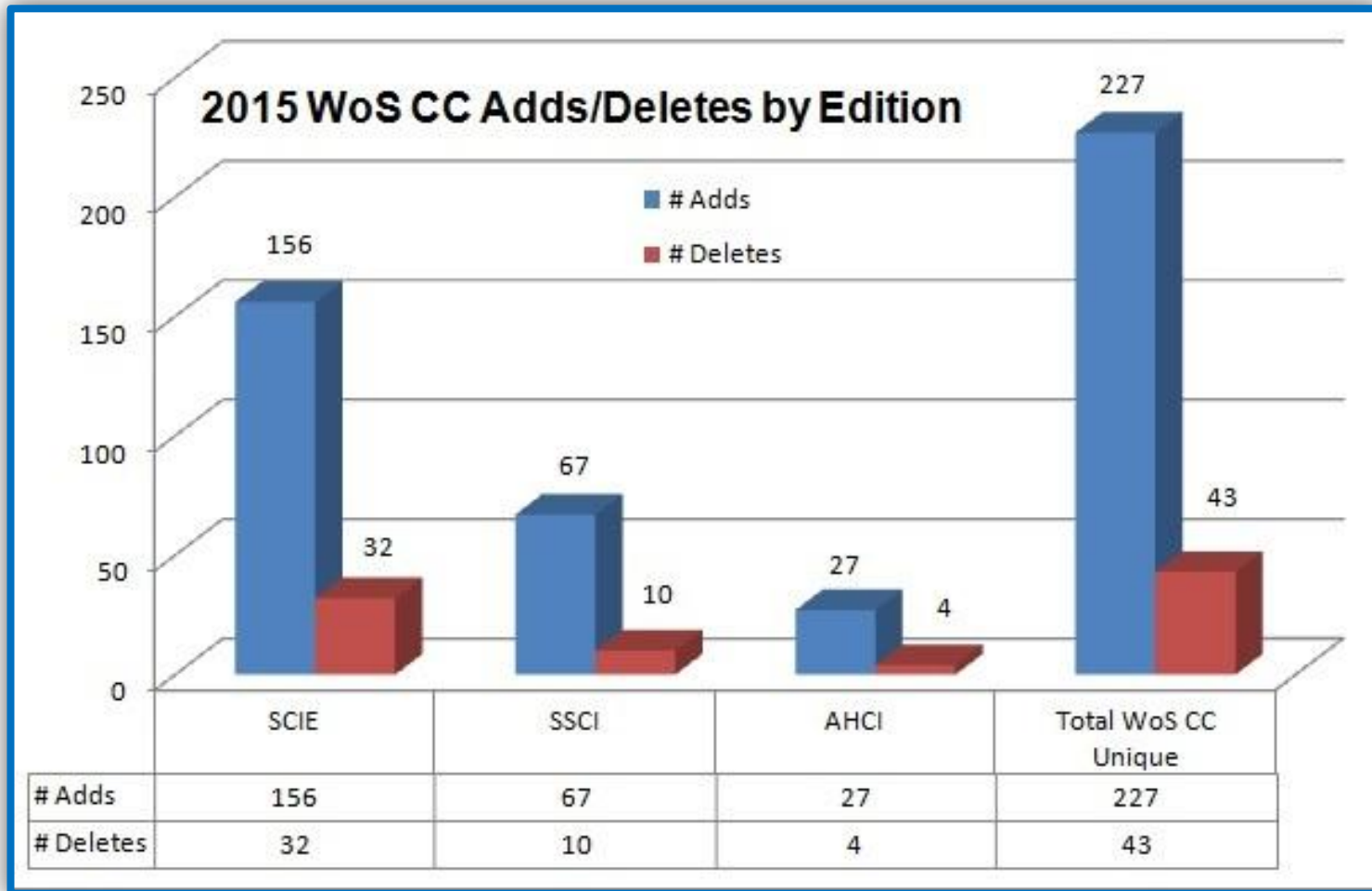
Jeffrey Beall's list

African Journal of Traditional Complementary and Alternative Medicines

International Journal of Electrochemical Science

Journal of Environmental Biology

Journal of Physical Therapy Science

6

Oncotarget

Romanian Biotechnological Letters

Value of the data: the four pillars

Indexing		
Cover to cover	All authors & affiliations	All Cited References
Funding Data	Document Types	Classification Schemes

Consistency is the key

A different level of metadata capture

Selectivity	Coverage	Process

THOMSON REUTERS

Process. Authors
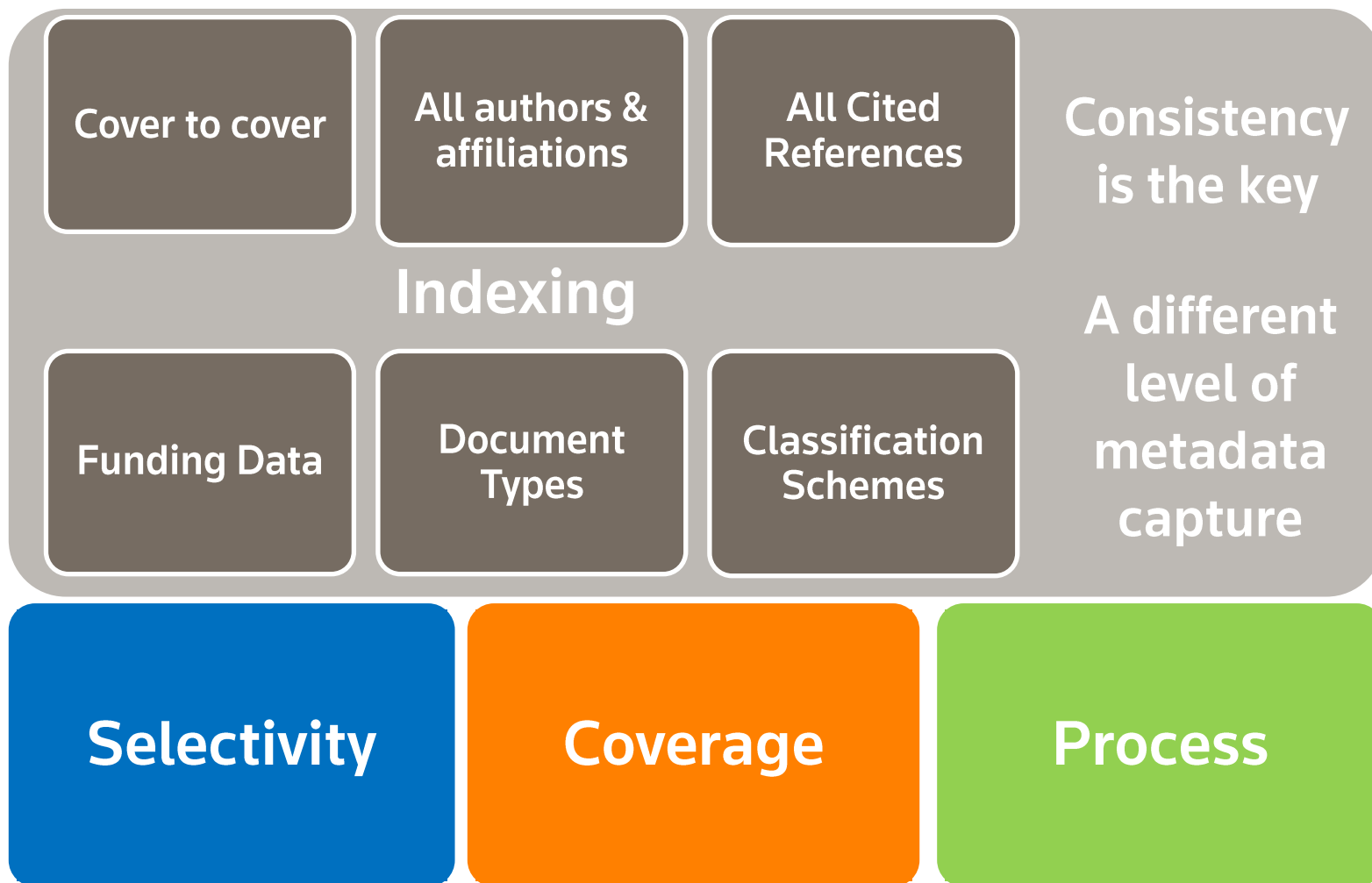
Author Name Disambiguation

In 2006 "Full Author Names" have been introduced in lieu of "Block Names" (last name + first name initial)

Authors and their address affiliations are linked together since 2008

The correct way to write an author name, is the following format: **Last name, First name / First initial**

Names are ambiguous (Eastern countries, Spain, etc...)

	Last Name	First Initial	Frequency	Region
1	wang	y	41,849	China
2	lee	j	41,548	Korea
3	lee	s	41,479	Korea
4	wang	j	37,909	China
5	kim	j	37,848	Korea
6	kim	s	34,861	Korea
7	zhang	y	32,943	China
8	wang	x	32,512	China
9	li	y	32,346	China
10	chen	y	31,455	China
11	kim	h	28,568	Korea
12	zhang	j	28,520	China
13	chen	j	28,388	China
14	li	j	28,148	China
15	liu	y	28,122	China

Almost 200 million authorships in WOS Core Collection.

Need of methods able to scale up millions of papers.

Author Name Disambiguation

Dis-algorithm

Auth-identifiers

The solution is, once again, a trade-off between algorithmic research of papers and manual curation through **author identifiers**.

Historically Thomson Reuters has developed a proprietary algorithm called **DAIS (Distinct Author Identification System)** of whom, recently, has been released a new version **(NG–Next Generation)**. The algorithm combines data point to collect authors names in clusters.

THOMSON REUTERS

Author Name Disambiguation

1
INITIAL CLUSTERING
Starting from scratch

2
ONGOING AUTHOR CREATION
Assign new author-Ids as new data come in

3
INCREMENTAL CHANGES
Integrates manually created (RID/ORCID) publication lists

4
PERIODIC RE-EVALUATION
Fresh full clustering

THOMSON REUTERS

Author Name Disambiguation

More data points. A machine learning (assigning weights) algorithm

THOMSON REUTERS

Author Name Disambiguation

Results: 53
(from Web of Science Core Collection)

👤 *Select articles grouped for author name* ℹ️ : *CARLONI M*

You searched for: AUTHOR:
(CARLONI M) **...More**

🔔 **Create Alert**

Distinct Author Sets

👤 Records grouped together are likely written by the same person. (Tell me more.)
If your papers appear in multiple article groups use the I wrote these publications button to add them to your publication list in ResearcherID and have them grouped together.

| → View Records | Add to ResearcherID - I wrote these | Clear All |

	Author Names	Last Known Organization ℹ️	Research Areas	Publication Years
1. ☐	CARLONI M Also published as: CARLONI MANUEL	University of Camerino	TOXICOLOGY (3) ENVIRONMENTAL SCIENCES ECOLOGY (3) NEUROSCIENCES NEUROLOGY (2) GERIATRICS GERONTOLOGY (2) BIOCHEMISTRY MOLECULAR BIOLOGY (2)	2010 - 2015

Records: 12
+ **A Sampling of Publications by this Author:**

4. ☐	CARLONI MT Also published as: CARLONI MARIA TERESA CARLONI M T CARLONI M	Undetermined	HEMATOLOGY (3)	2008 - 2013

Records: 3
+ **A Sampling of Publications by this Author:**

Integrated data

New features at a glance

Item Level Usage Metrics (ILUM)

Sort by Usage Counts. Counts display when sort option is chosen.

Full Usage Count view can be opened and closed.

New features at a glance

Saved Marked List

Marked List   ⊗

Marked lists can now be saved to be used in later sessions

Marked List 501 records | **View Derwent Compounds Marked List: 0 compounds**

Save | Open/Manage | ✕ Clear

501 total records on the Marked List
Output author, title, source, abstract, and times cited for all records in the Marked List.

501 records from *Web of Science™ Core Collection*
Output complete data from this product for these records.

Up to 50 Lists. Each of them with 5k records

Open/Manage Saved Marked Lists
<< Back to previous page

☐ Select All   ✕ Delete

Name	Description	Count	Last Modified	Edit
☐ Prova	Marked List di prova	501	01-27-2016	Edit

☐ Select All   ✕ Delete

THOMSON REUTERS

New features at a glance

Manuscript Matcher

New features at a glance

Landing Page

From "all databases" to Web of Science Core Collection

All databases features - discoverability

Multiply the chances to discover valuable piece of literature exploiting connection and analysis tools

All databases features – enriched information

Adolescent exposure to THC in female rats disrupts developmental changes in the prefrontal cortex

By: Rubino, T (Rubino, Tiziana)[1,2]; Prini, P (Prini, Pamela)[1,2]; Piscitelli, F (Piscitelli, Fabiana)[3]; Zamberletti, E (Zamberletti, Erica)[1,2]; Trusel, M (Trusel, Massimo)[4]; Melis, M (Melis, Miriam)[5]; Sagheddu, C (Sagheddu, Claudia)[5]; Ligresti, A (Ligresti, Alessia)[3]; Tonini, R (Tonini, Raffaella)[4]; Di Marzo, V (Di Marzo, Vincenzo)[3]; Parolaro, D (Parolaro, Daniela)[1,2,6] ...Less

NEUROBIOLOGY OF DISEASE
Volume: 73 Pages: 60-69
DOI: 10.1016/j.nbd.2014.09.015
Published: JAN 2015
View Journal Information

This record is from:
Web of Science™ Core Collection

View Record in Other Databases:
View **most recent data** (in Current Contents Connect®)

View **biological data** (in BIOSIS Citation Index SM)

View **biological data** (in BIOSIS Previews®)

View **biological data** (in Biological Abstracts®)

View **medical data** (in MEDLINE®)

SUPER TAXA	TAXA NOTES	Organism Classifier	Orga Name	
Rodentia, Mammalia, Vertebrata, Chordata, Animalia	Animals, Chordates, Mammals, Nonhuman Vertebrates, Nonhuman Mammals, Rodents, Vertebrates	Muridae [86375]	rat	mature; female; strain-Sprague-Dawley

Different level of metadata capture (several information layers)
The SUPER article concept

THOMSON REUTERS

All databases features – subject indexing exploitation

Basic Search

Example: oil spill* mediterranean   ⊗ | Topic   ⌄ | **Search**

+ Add Another Field

Topic Search Mapping

Web of Science Core collection Title, Abstract, Author Keywords, KeyWords Plus®	**BIOSIS Citation Index Biological Abstracts** Title, Abstract Major Concepts, Concept Code(s) Taxonomic Data, Disease Data, Chemical Data, …	**Derwent Innovations Index** Title. Abstract, Equivalent abstracts, International patent classification, Derwent Class codes, Derwent Manual codes	**Inspec** Title, Abstract, Controlled Indexing, Uncontrolled Indexing, Original Indexing Classification Code(s)
Zoological Records Title, Abstract, Broad Terms Descriptors Data, Super Taxa, Taxa Notes	**MEDLINE** Title, Abstract, MeSH Terms Keyword List, Chemical, Gene Symbol, Subject,…	**CABI** Title, Abstract, Descriptors, Broad Descriptors, Organism Descriptors, Geographic Location, CABICODE Names	**Data Citation Index** Titles, Abstracts, Repository Name, Data Study, Data Set
Food Science and Technology abstracts Title, Abstract, FSTA Thesaurus, MeSH Thesaurus	**Current Contents Connect** Title, Abstract, Author Keywords KeyWords Plus®	**Chinese Science Citation Database** Title, Abstract, Author Keywords	**SciELO Citation Index** Title, Abstract, Author Keywords

JCR 2016 – what's new

11,365 total journals
239 new journals

10%: the average increase in JIF across all journals

234 disciplines (1 new category: Green & Sustainable Science & Technology)

81 countries

18 journals suppressed:
16 self-citing
2 citation stacking
1 suppressed previous year

JCR 2015 – Innovative Metrics Development

- ## Journal Impact Factor Percentile
  - Allows for meaningful comparison of journals in different categories

$$\frac{(n - r + .5)}{n}$$

- ## % Articles in Citable Items
  - Calculates the proportion of articles in the total number of citable items for any given year

- ## Normalized Eigenfactor
  - Improved usability for understanding journal influence

Before

Rank	Journal	Eigenfactor
1	Proceedings of the Royal Society B – Biological Sciences	0.091
2	Philosophical Transactions of the Royal Society B	0.078
3	Ecology	0.073
4	Molecular Ecology	0.072
5	Ecology Letters	0.069
6	New Phytologist	0.065
7	Global Change Biology	0.061
8	Trends in Ecology and Evolution	0.052
9	Marine Ecology – Progress Series	0.051
10	Evolution	0.048

After

Rank	Journal	Eigenfactor
1	Proceedings of the Royal Society B – Biological Sciences	9.93
2	Philosophical Transactions of the Royal Society B	8.45
3	Ecology	7.88
4	Molecular Ecology	7.79
5	Ecology Letters	7.52
6	New Phytologist	7.08
7	Global Change Biology	6.62
8	Trends in Ecology and Evolution	5.68
9	Marine Ecology – Progress Series	5.52
10	Evolution	5.20

THOMSON REUTERS

now ESCI !

JCR us WoS

*3 years needed for IF
Arts & Humanities*

Name Changes

*Self-citations &
Stacking*

http://ipscience-help.thomsonreuters.com/incitesLiveJCR/JCRGroup/titleSuppressions.html

THOMSON REUTERS

THOMSON REUTERS™

webofscience.com

converis.thomsonreuters.com

my.endnote.com

researcherid.com

wokinfo.com

youtube.com/WoSTraining

massimiliano.carloni@thomsonreuters.com

linkedin.com/in/mcarloni