

# کارگاه آشنایی با جستجو تحلیل استنادی در علم سنجی

دکتر پیام کبیری

پزشک ، دکترای اپیدمیولوژی

رئیس کارگروه علم سنجی وزارت بهداشت ، درمان و آموزش پزشکی

رئیس مرکز توسعه و هماهنگی اطلاعات و انتشارات علمی وزارت بهداشت ، درمان و آموزش پزشکی

# برنامه کارگاه امروز !

■ اطلاع‌سنجی و علم‌سنجی ، مفاهیم و کاربرد

■ معرفی خلاصه شاخص‌های علم‌سنجی

■ آشنایی با نمایه‌نامه‌های استنادی

■ جستجو و تحلیل استنادی در ISI Web of Science و Scopus

■ ضریب تاثیر مجلات Impact Factor ، معرفی (JCR)

■ معرفی شاخص H-Index

■ آشنایی با شاخص‌های نوین علم‌سنجی

# Scientometrics (bibliometrics)

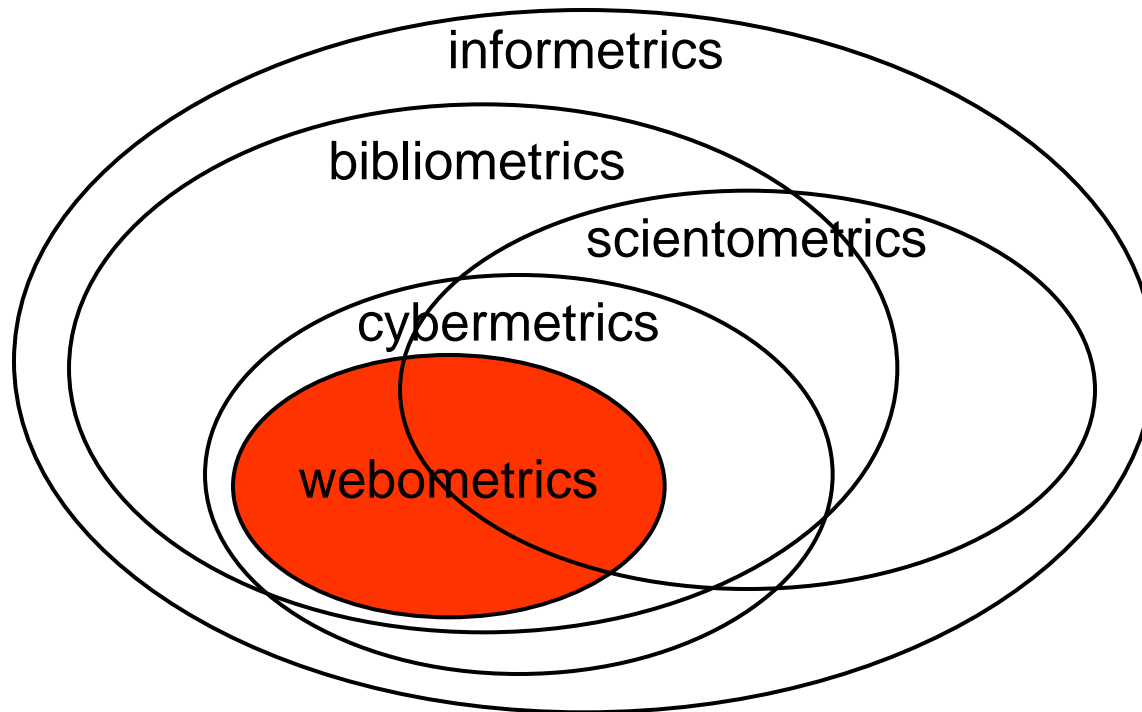
- **Scientometrics (bibliometrics)** - The measurement of scientific output activity through statistics on academic publications
- The scope of bibliometrics includes:
  - “all quantitative aspects and models of science communication, storage, dissemination and retrieval of scientific information”.*

## تعریف علم سنجی

■ آن دسته از روش‌های کمی را که به تحلیل علم بعنوان یک فرآیند اطلاعاتی تأکید دارند "علم سنجی" می‌نامند.

■ "علم سنجی" به تعبیری ساده تر عبارت است از دانش اندازه‌گیر علم.

# Scientometrics



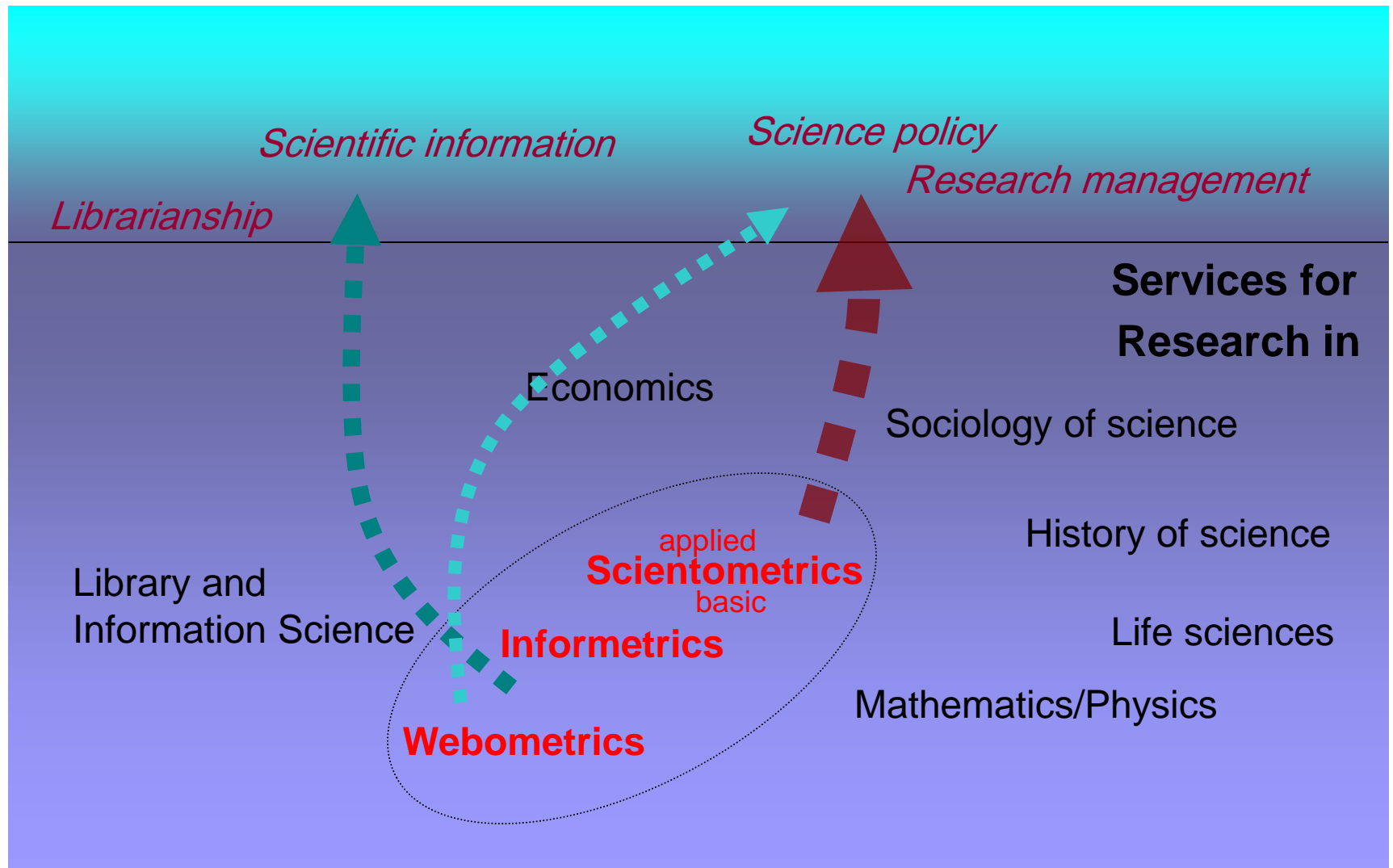
# Informetrics

- broader than bibliometrics and scientometrics
  - Informetrics examines patterns that show up not only in publications but also in many aspects of life as long as the patterns deal with information.
- e.g. Willis Law deals with the relationships between the age of a group of the animal and plant species and the habitation of these species
- its concepts, formulas applicable to publication patterns.

# Bibliometrics Data **Used for..**

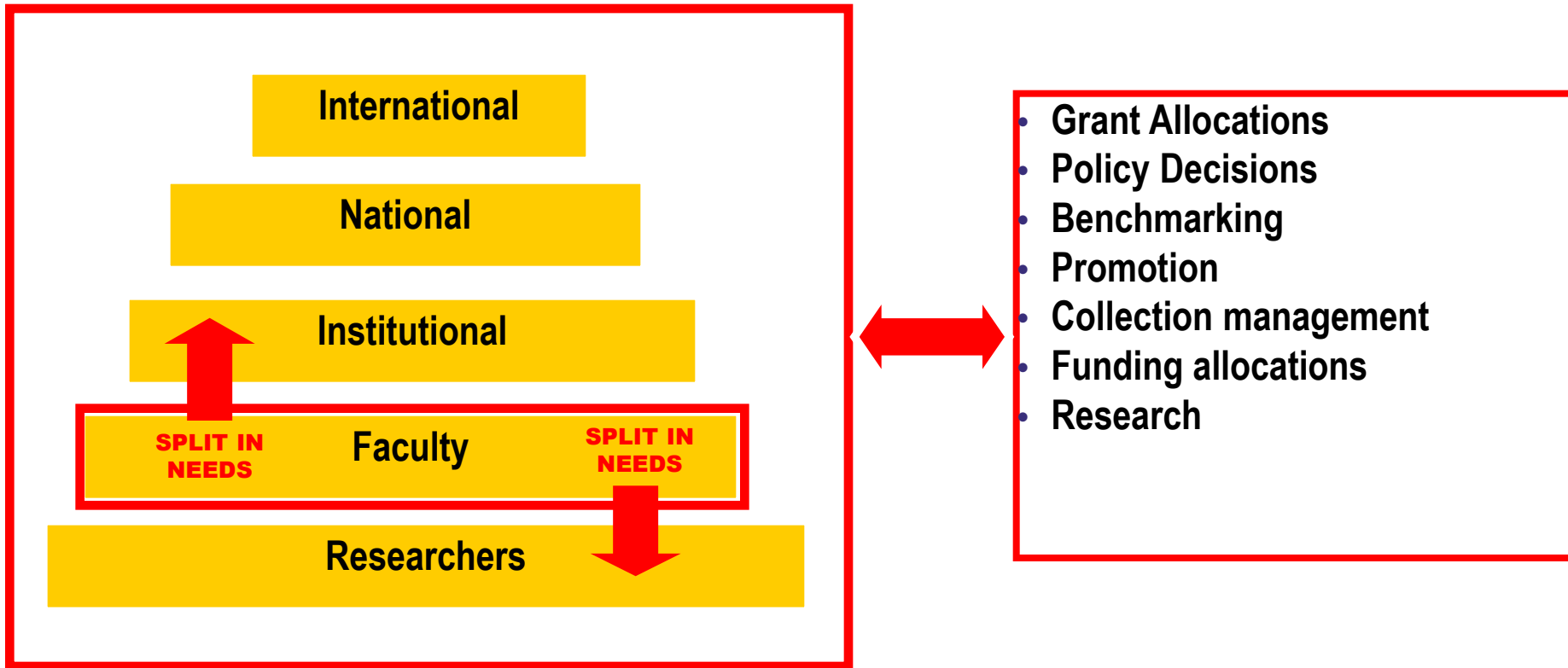
- Scientific output evaluation
  - Impact
  - Citations
- History of science
- Publication strategies
- Science policy; resource allocation
- Collection management
- Sociology of science
- Information organization
- Information management & utilization

# Links of Scientometrics with Related Research Fields and Application Services

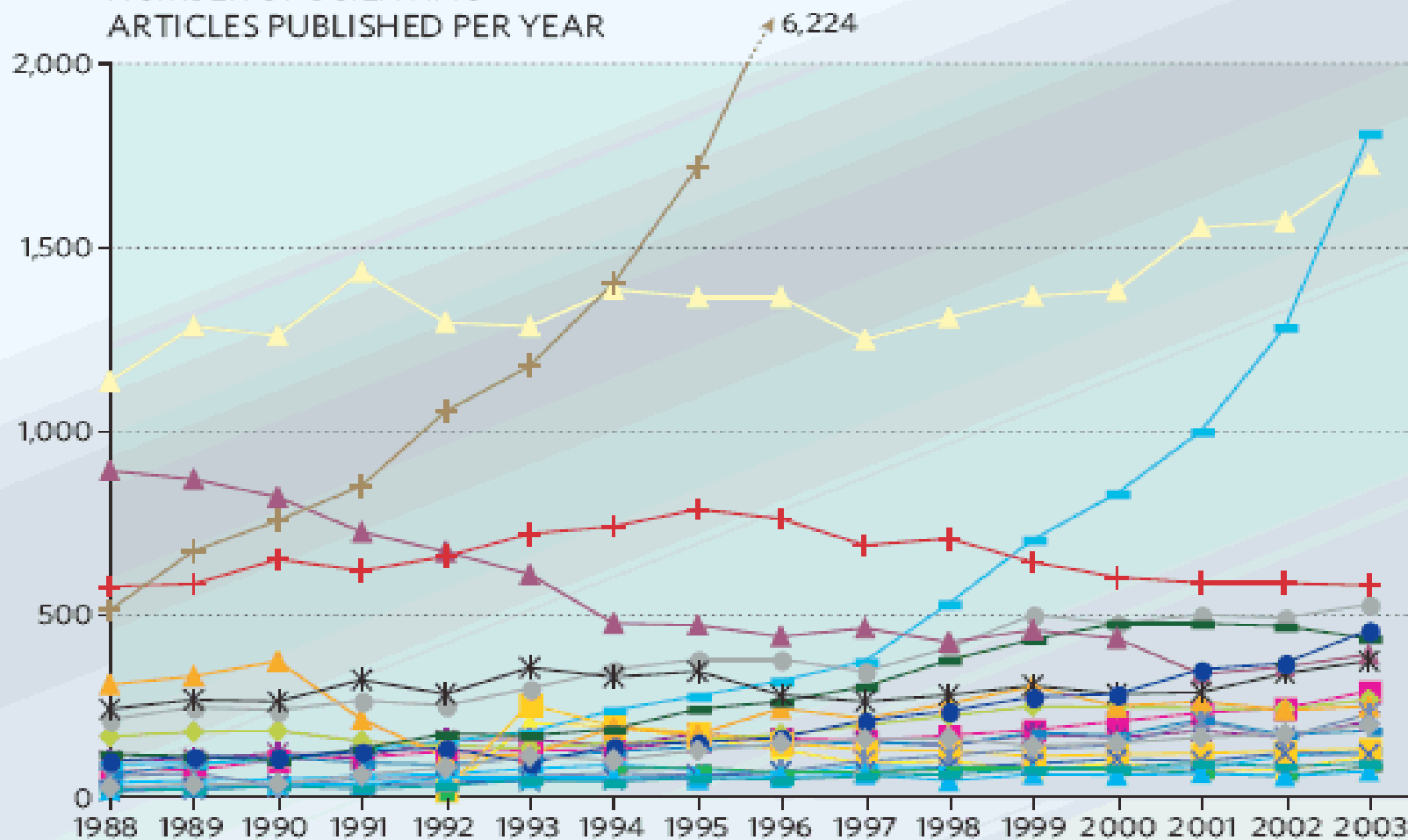




# Why do we evaluate scientific output



### NUMBER OF SCIENTIFIC ARTICLES PUBLISHED PER YEAR



## رشد 11 برابری شمار مقالات ایران در مقایسه با میانگین جهانی

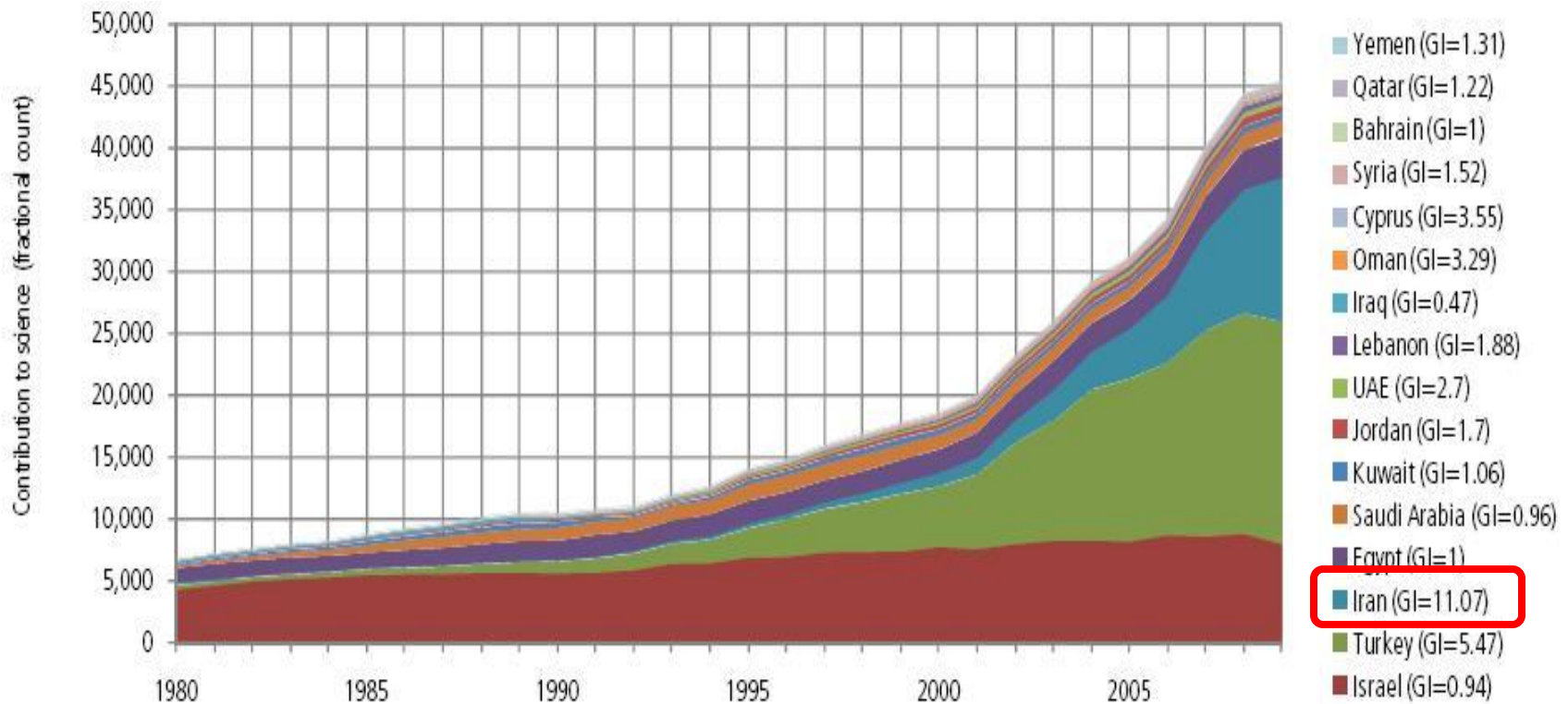


Figure 5

Contribution of Middle East countries to world science, 1980–2009

Source:

Calculated by Science-Metrix using the Web of Science (Thomson Reuters)



مقایسه فراوانی مقالات و ارجاعات صورت گرفته به مقالات منتشر شده دانشگاه علوم پزشکی تهران و دانشگاه تهران با دانشگاه‌های رتبه یک‌صدم و دویستم دنیا در مجموعه Scopus در رتبه‌بندی سال 2009 تایمز THS-QS

رتبه دانشگاه	نام دانشگاه	کشور	تعداد مقالات 5 ساله در Scopus	تعداد ارجاعات به مقالات 5 ساله Scopus	تعداد اعضای هیئت علمی	نسبت ارجاع به عضو هیئت علمی	نسبت ارجاع به مقاله
100	دانشگاه Rice	آمریکا	6254	66343	772	85/9	10/61
200	دانشگاه Twente	هلند	6193	37403	1000	37/4	6/04
201	دانشگاه اندونزی	اندونزی	549	2093	4321	0/48	3/81
368	دانشگاه علوم پزشکی تهران و دانشگاه تهران	ایران	11359	31030	3243	9/6	2/73



# Databases

- **General** Databases  
(Comprehensive OR Core Databases)
- **Specialized** Databases  
(Subjects Specified Databases)



# General Databases

## (Comprehensive OR Core Databases)

- **Medical Sciences**

- Medline
- Embase

- **All Sciences**

- Web of Sciences
- Scopus



# Specialized Databases

## (Subjects Specified Databases)

- Biological Abstracts
- International Pharmaceutical Abstract
- PsychInfo
- CINAHL
- Chemical Abstracts
- Agricola
- Econlite



# Citation Databases

- Web of Science
- Scopus
- Google Scholar
  - (<http://scholar.google.com>)

# 3 Types of Citation Data Indexes

## Articles

- Citation Impact

## Authors

- Number of Papers (**Quantity**)
- Number of Citations (**Quality**)
- Average number of **Citations/Paper**
- *h*-index & *g*-index (Quantity & Quality Both)

## Journals

- Journal Impact Factor
- *h*-index

# Scopus

- Positioning itself as an alternative to ISI
- More journals from smaller publishers and open access (+21,000 journals; +1000 conf proceedings)
- Source data back to 1960.
- Excellent for physical and biological sciences; poor for social sciences; does not cover humanities or arts.
- Better international coverage (60% of titles are non-US)
- Back to 1996 ! (e.g. citation data for the last decade only)
- Easy to use in searching for source publications; clumsy in searching cited publications.
- Citation tracker works up to 1000 records only.

# Google Scholar

- Better coverage for all citations as it retrieve web !
- More coverage of references also gray literature !
- Coverage and scope?
- Inclusion criteria?
- Very limited search options
- No separate cited author search
- Back to 1990 NOT more !
- Free!



# Web of Science

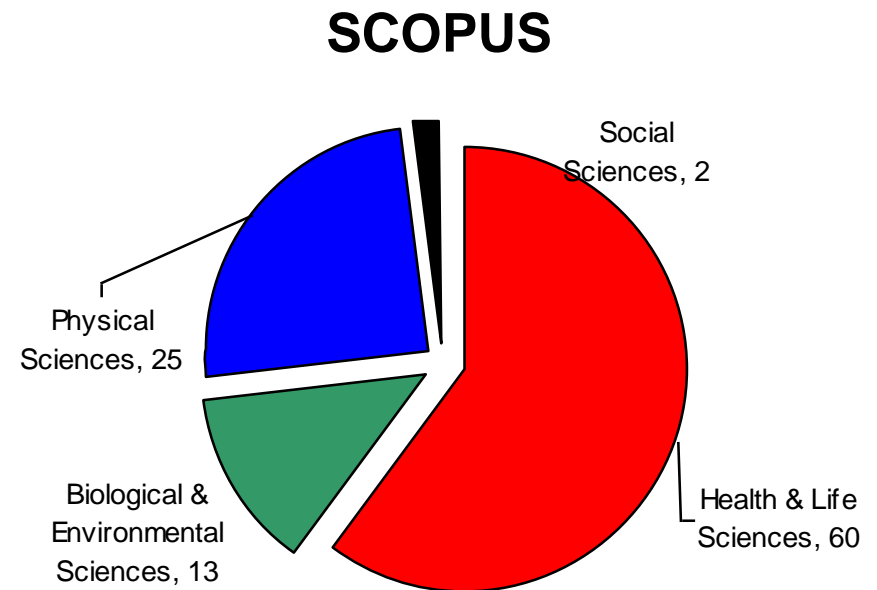
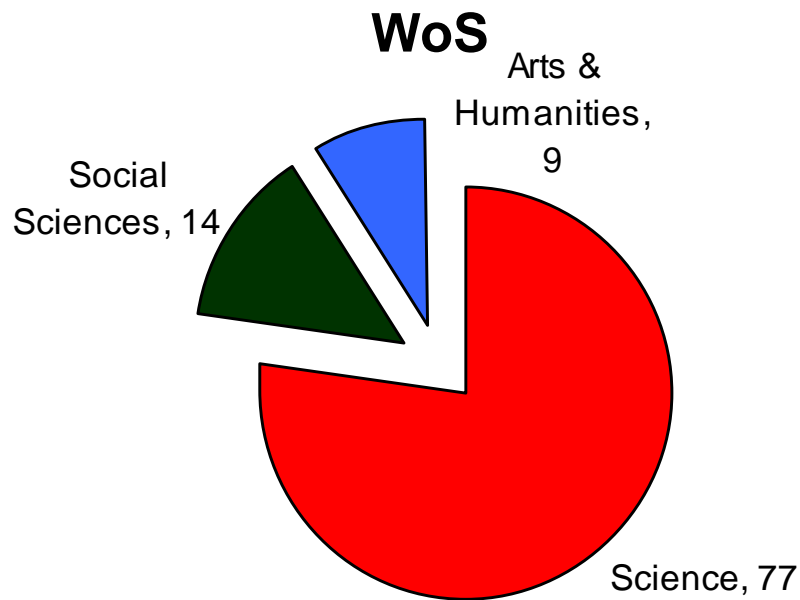
Ex **Thomson Reuter** ISI,

Now **Clarivate Analytics**

- Covers around 11,000 journal titles and 110,000 conference proceedings & also 200 book series divided between SCI, SSCI and A&HCI.
- Over 40 million records!
- Electronic back files available to 1900 for SCI and mid- 50s for SSCI and mid-70s for A&HCI.
- Very good coverage of sciences; “softer” sciences, social sciences humanities.
- Full coverage of citations.



# WoS and Scopus: Subject Coverage (% of total records)



Google Scholar ?

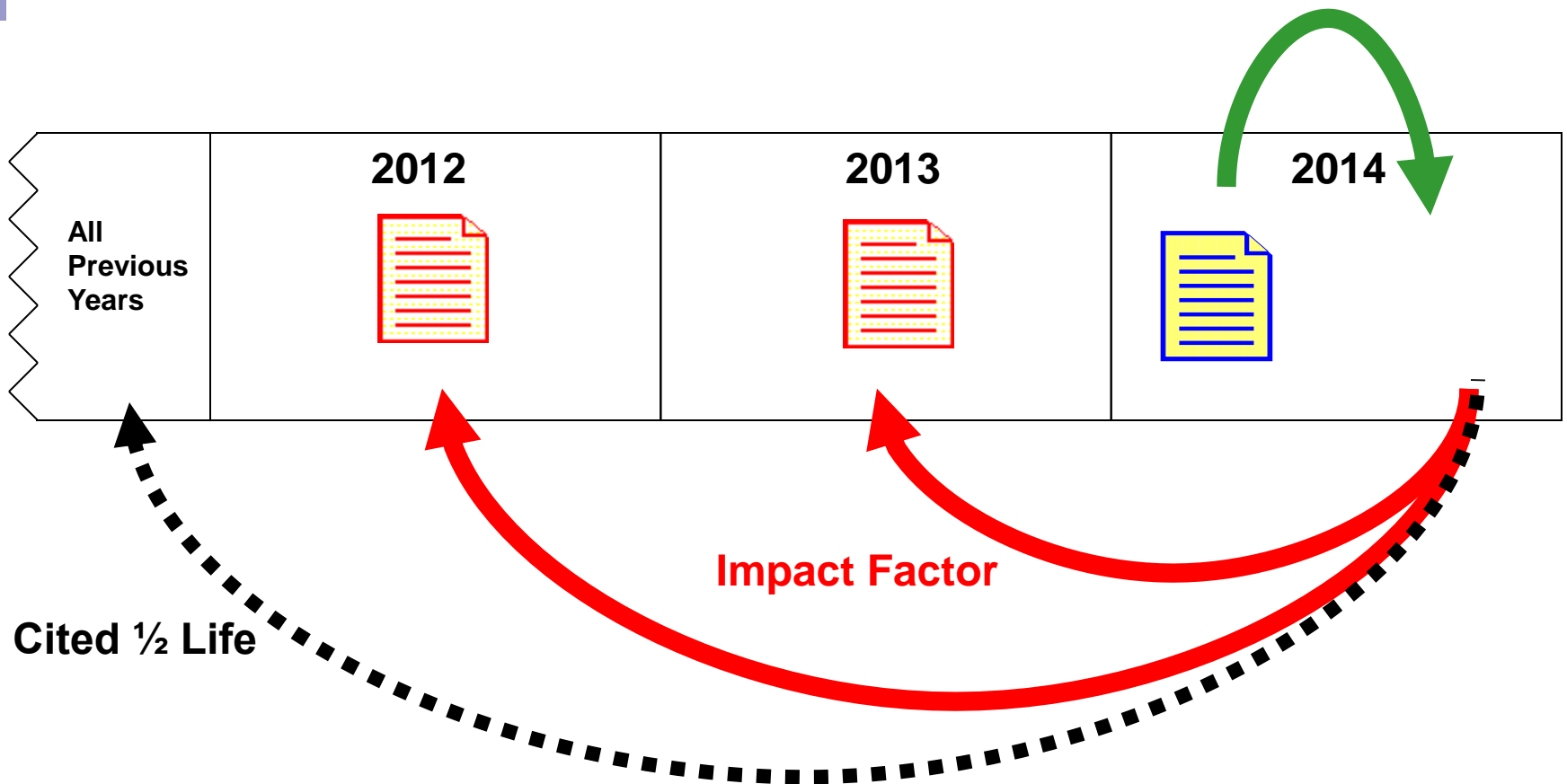
# ISI Web of Science

- If a journal is in **Web of Science**, it means it should have **Impact Factor** & vice versa.
- Don't confuse **Web of Sciences Journals** with **ISI Master list Journals** !
- Web of Science indexed about 11,000 Journals while ISI Master List 14,000 Journals.

# What is **Journal Impact Factor**?

- It is primarily invented to screen those **high quality** journals into be included in **ISI Web of Science** database.
- The **average citation frequency for articles published in a journal**, or **how many times**, on average, during the study year the articles that appeared in the 2 preceding years of that journal received citations in other **(ISI)** indexed journals only.





Cited  $\frac{1}{2}$  Life

Impact Factor



Citation



Source paper – published in 2014

Cited reference – published in 2014

Cited reference – published in 2012 or 2013

# Impact Factor **Formula** & **Calculations**

- Suppose there is a journal, published some papers in 2012 & 2013:

Item	2012	2013	Total
Total Papers Published	130	170	300
Citations Achieved in 2014	360	240	600

$$\text{Impact Factor} = \frac{\text{Citations}}{\text{Papers}} = \frac{600}{300} = 2 \quad \text{IF}$$

# Impact Factor Calculation

- Citations in the current JCR year to articles published in the previous **two years** divided by **the number of articles** published in the previous two years.

**Citations in 2014 to articles published in 2012 + 2013**

$$\text{IF} = \frac{\text{Citations in 2014 to articles published in 2012 + 2013}}{\text{Total 2012 + 2013 Papers}}$$

# The Journal **Impact Factor**

- The **Journal Impact Factor** is calculated for those journals only which are indexed & included in **Web of Sciences** Databases, NOT more, So

ISI Indexing  $\longleftrightarrow$  Having IF

- Except .....

# How to Find Impact Factors?

- The database which contain the Journal Impact Factors is **Journal Citation Report** abbreviated **JCR**. [Click here](#)
- Directly from **Journal Citation Report** Database through ISI.
- Using the Excel Datasheets, prepared especially for this matter.

# 2013 Impact Factors

Mark	Rank	Abbreviated Journal Title <i>(linked to journal information)</i>	ISSN	JCR Data <sup>j</sup>						Eigenfactor <sup>®</sup> Metrics <sup>j</sup>	
				Total Cites	Impact Factor	5-Year Impact Factor	Immediacy Index	Articles	Cited Half-life	Eigenfactor <sup>®</sup> Score	Article Influence <sup>®</sup> Score
<input type="checkbox"/>	1	<a href="#">CA-CANCER J CLIN</a>	0007-9235	16130	162.500	107.740	27.760	25	3.1	0.06030	34.798
<input type="checkbox"/>	2	<a href="#">NEW ENGL J MED</a>	0028-4793	257469	54.420	52.426	14.747	348	8.2	0.65797	22.412
<input type="checkbox"/>	3	<a href="#">CHEM REV</a>	0009-2665	124463	45.661	48.832	7.101	207	8.1	0.21741	14.261
<input type="checkbox"/>	4	<a href="#">REV MOD PHYS</a>	0034-6861	37647	42.860	52.577	8.333	45	>10.0	0.12864	32.037
<input type="checkbox"/>	5	<a href="#">NATURE</a>	0028-0836	590324	42.351	40.783	8.457	857	9.8	1.60305	22.184
<input type="checkbox"/>	6	<a href="#">ANNU REV IMMUNOL</a>	0732-0582	16653	41.392	46.174	9.833	24	8.7	0.04785	23.737
<input type="checkbox"/>	7	<a href="#">NAT REV GENET</a>	1471-0056	26358	39.794	40.274	6.191	68	5.2	0.12696	20.650
<input type="checkbox"/>	8	<a href="#">LANCET</a>	0140-6736	176528	39.207	39.315	12.649	276	9.0	0.38061	15.986
<input type="checkbox"/>	9	<a href="#">NAT BIOTECHNOL</a>	1087-0156	42156	39.080	35.620	12.043	93	6.8	0.14395	17.560
<input type="checkbox"/>	10	<a href="#">NAT REV CANCER</a>	1474-175X	36052	37.912	41.706	4.818	66	6.9	0.11207	18.684

# How to **Find** Impact Factors?

- Directly from **Journal Citation Report** Database through ISI.
- Using the Excel & PDF Datasheets, prepared especially for this matter.
  - [www.isc.gov.ir](http://www.isc.gov.ir)
  - [www.sid.ir](http://www.sid.ir)

# The *h*-index

■ شاخص جدیدی از شاخص‌های علم‌سنجی است. این شاخص در سال 2005 میلادی توسط **Jorge Hirsch** در دانشگاه کالیفرنیا ابداع شد. این شاخص در واقع با هدف ارزیابی کیفی اثر و ارزیابی کمی برون‌داد پژوهشی محققین ابداع شده است.



# The *h*-index

■ مفهوم **H-Index** عبارت است از تعداد مقالات نویسنده که تعداد ارجاعات برابر با **h** و یا کمتر از آن دارند. مثلاً چنانچه **H-Index** محققی 5 باشد، مفهوم آن این است که این محقق 5 مقاله منتشر شده دارد که هرکدام حداقل 5 استناد یا **Citation** دارند. به عبارت دیگر مفهوم آن این است که سایر مقالات این محقق کمتر از 5 استناد دارند.

■ امروزه این شاخص معادل **Impact Factor** برای محققین محسوب می شود.

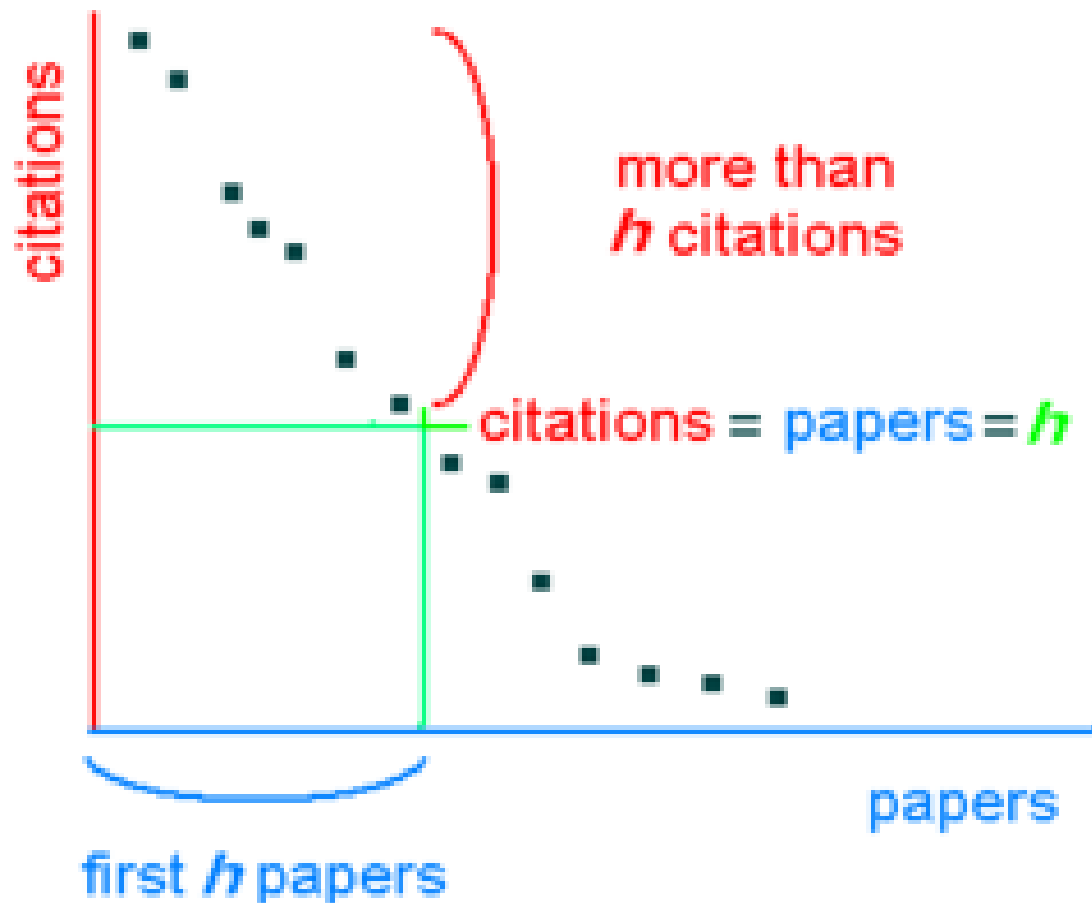
# H-index was born !

- We need an Index both to include quantity & also quality of an authors' paper
  - ✓ Productivity
  - ✓ Impact
  - ✓ Not affected by “big hits”
  - ✓ Not affected by “noise”

# The H-index: a definition

- ‘The H-index is the highest number of papers a scientist has that have at least that number of citations.’ *Nature* (2005)

# H-index Concept through its Graph



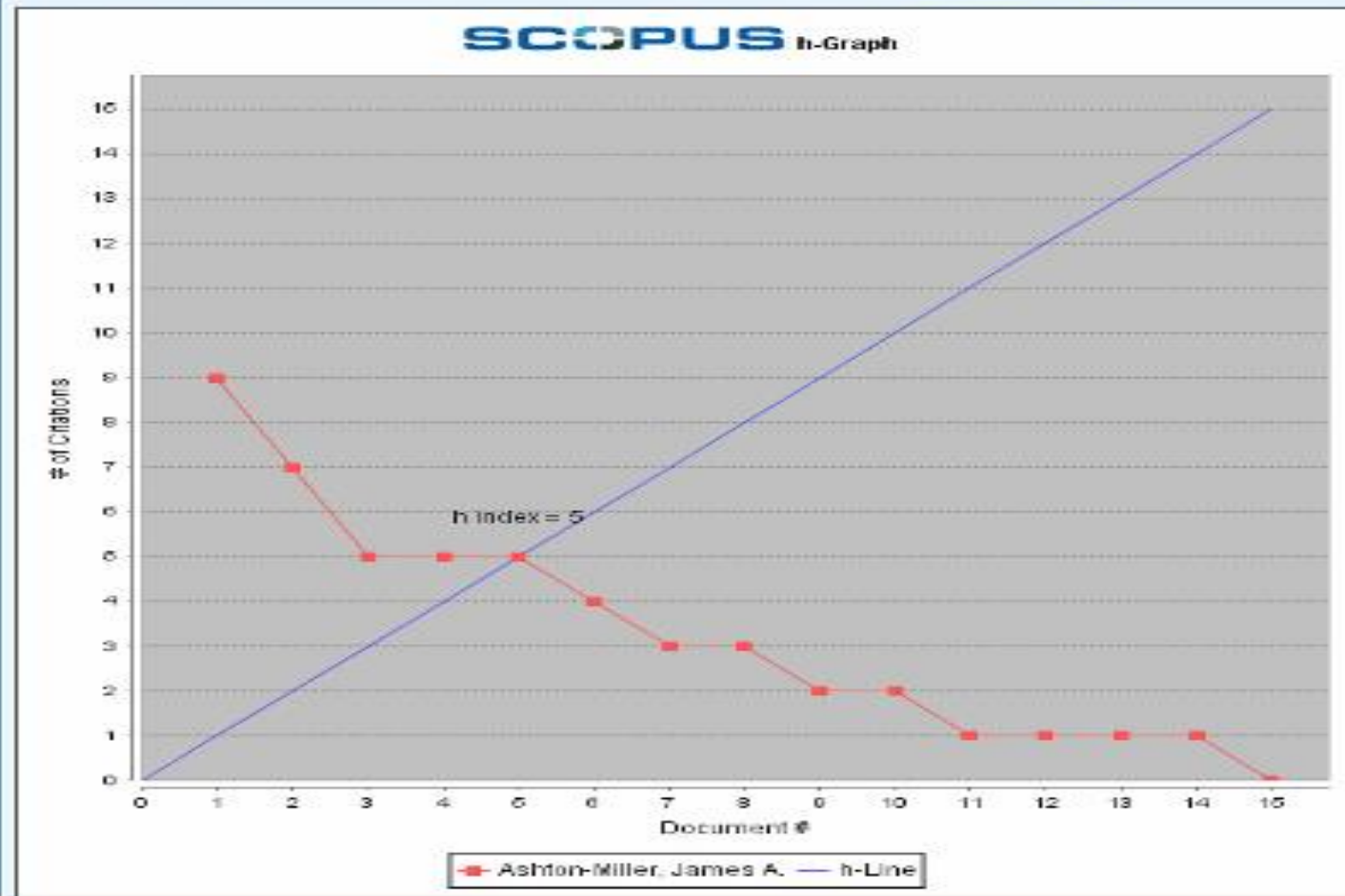
# The $h$ - Graph

**$h$ -Graph** Articles Published Citations

Only consider documents published from: 1996 to 2007 [Update graph](#)

$h$  index = 5 (5 out of 15 documents have each been cited at least 5 times)

Notes: The  $h$  Index considers Scopus documents published after 1995. [More Information](#)



# How to calculate the h-index?

- You can calculate the h-index through the following citation databases:
  - Web of Science
  - Scopus
  - Google Scholar
    - (<http://scholar.google.com>)

# h-index Calculation through Google Scholar

- To calculate h-index through Google Scholar, you should use the 3<sup>rd</sup> parties services.
- Some softwares has been developed for this matter, like: **Publish or Perish**

# Publish or Perish Software

- You can download from:  
<http://www.harzing.com/>
- It's a free software which can be downloaded directly at:  
<http://www.harzing.com/pop.htm>



# Publish or Perish Software

The screenshot shows the 'Harzing's Publish or Perish' application window. The interface includes a menu bar (File, Edit, View, Tools, Help) and a toolbar with tabs for 'Author impact', 'Journal impact', 'General citations', 'Multi-query center', and 'Web Browser'. The 'Web Browser' tab is active, displaying a page titled 'Publish or Perish start page' at the address 'C:\Program Files\Harzing's Publish or Perish 3\Help\start.htm'. The page content includes a purple header 'Welcome to Publish or Perish', a paragraph describing the software's functionality, a list of links (Getting started, Documentation, The Publish or Perish Book), a copyright notice, a quote about the 'The Publish or Perish Book', and a 'Getting started' section with three numbered steps. A sidebar on the left contains navigation links for 'Citation analysis', 'Program maintenance', and 'Help resources', along with a section for 'The Publish or Perish Book'. The bottom status bar shows the version '3.4.4351' and the date 'الثبت، ١٧ ديسمير، ٢٠١١'.

Harzing's Publish or Perish

File Edit View Tools Help

Author impact Journal impact General citations Multi-query center Web Browser

Web browser - Perform general web-oriented tasks using Internet Explorer

Address: C:\Program Files\Harzing's Publish or Perish 3\Help\start.htm Go Help

Page title: Publish or Perish start page << Back Next >>

## Welcome to Publish or Perish

Publish or Perish is a software program that retrieves and analyzes academic citations. It uses [Google Scholar](#) to obtain the raw citations, then analyzes these and calculates a series of citation metrics. The results are available on-screen and can also be copied to the Windows clipboard (for pasting into other applications) or saved to a text file (for future reference or further analysis).

- [Getting started](#)
- [Documentation](#)
- [The Publish or Perish Book](#)

The Publish or Perish software and documentation are copyright © 1990-2011 Tarma Software Research Pty Ltd. All rights reserved.

*The Publish or Perish Book is more than just a guide to the software. It is a tremendously accessible introduction to bibliometrics and citation analysis, which many academics and graduate students find a pretty arcane and impenetrable topic.*

[Read more Amazon reviews about The Publish or Perish book...](#)

### Getting started

To get started with Publish or Perish:

1. Go to one of the analysis tabs (**Author analysis**, **Journal impact analysis**, **General citation search**),
2. Enter your search terms,
3. Click **Lookup**

The software will contact [Google Scholar](#), then parse the results and present them in an easy to use list along with the most important citation statistics.

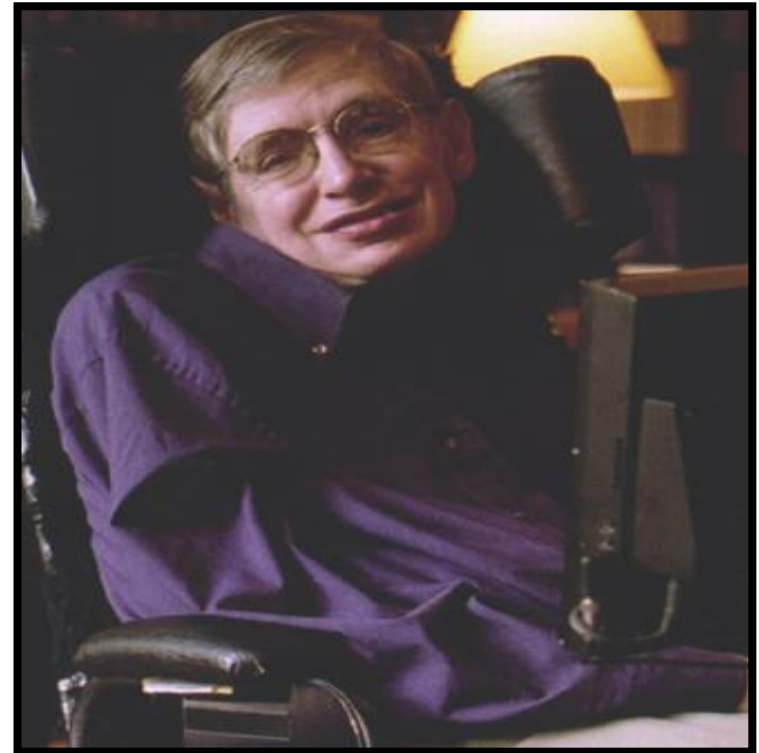
3.4.4351 الثبت، ١٧ ديسمير، ٢٠١١

*My h-index is bigger than yours!*

*But more people know who I am!*



*Edward Witten  
Physicist  
h=132*



*Stephen Hawking  
Physicist  
h=62*

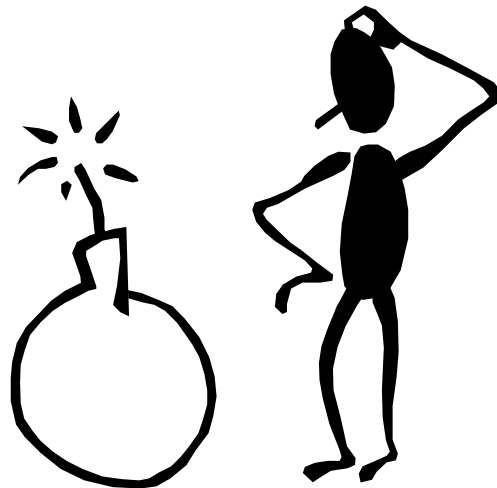


# **New Scientometrics Indices**

# An Example

Entity	Subject Area	Citation/Paper
Researcher A	Clinical Medicine	3
Researcher B	Mathematics	1

Which one's **Performance** is Better?



# A **New Metrics** is Required !

- We need a new Metrics to **fairly compare** the papers within their **similar publication group** in the universe.
- By similar publications, we mean:
  - Same Publication **Year**
  - Same Publication **Discipline**
  - Same Publication **Type**

## New Metrics:

# Field-Weighted Citation Impact

- Field-Weighted Citation Impact is a **Ratio** that takes into account the **differences in research behavior across disciplines**.
- Field-Weighted Citation Impact is the **ratio** of the **total citations** actually **received** by the denominator's output, and the **total citations that would be expected** based on the **average of the subject field, year & document type**.

# New Metrics:

## Field-Weighted Citation Impact

- Field-Weighted Citation Impact is the **Ratio** of the **total citations** actually **received** by the denominator's output, and the **total citations that would be expected** based on the **average of the subject field**.
- Field-Weighted Citation Impact takes into account the **differences in research behavior across disciplines**.

## New Metrics:

# Field-Weighted Citation Impact

- Field-Weighted Citation Impact in SciVal indicates how the number of citations received by an entity's publications compares with the average number of citations received by all other similar publications in the data universe: how do the citations received by this entity's publications compare with the world average?



# What we mean by **Similar Publications?**

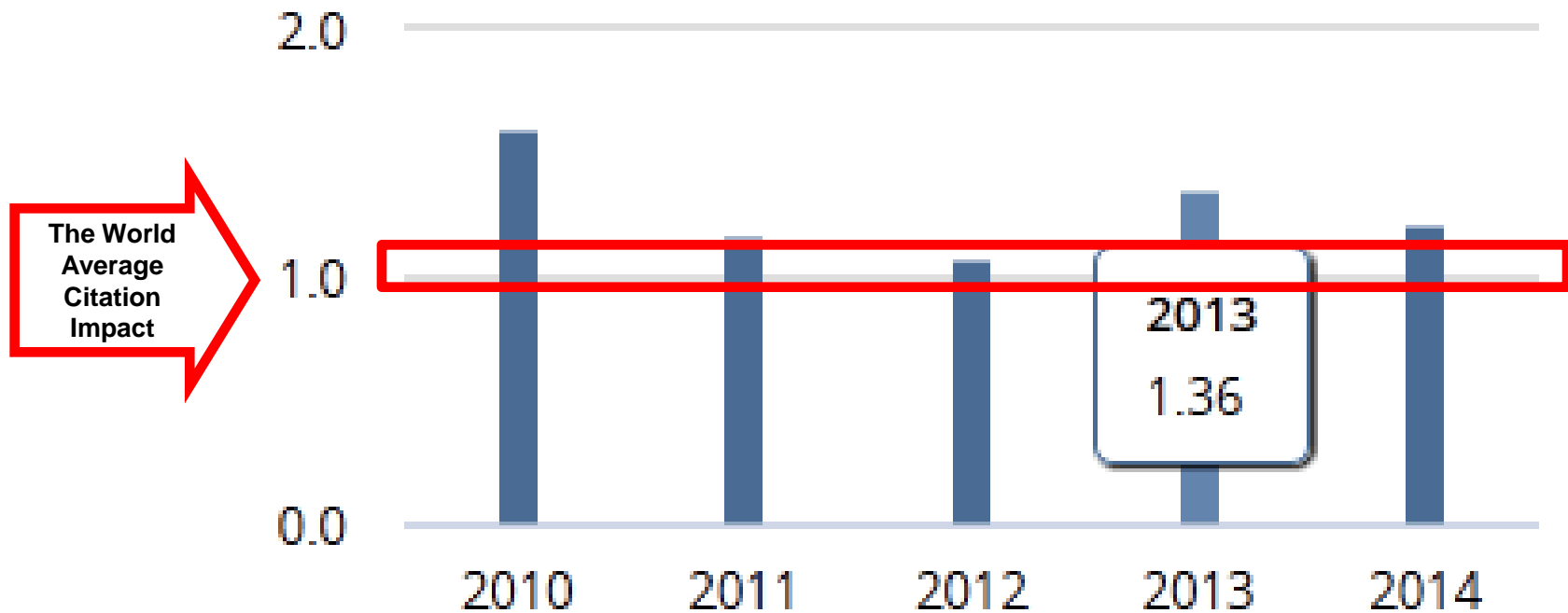
- By **similar** publications, we mean:
  - Same Publication **Year**
  - Same Publication **Discipline**
  - Same Publication **Type**
- So it would be **adjusted** for Papers **Disciplines, Age, & Type.**

# Field-Weighted Citation Impact

- $FWCI=1$  means that the output performs just as expected for the global average
- $FWCI>1$  means that the output is more cited than expected according to the global average; for example, 1.48 means 48% more cited than expected
- $FWCI<1$  means that the output is cited less than expected according to the global average.

# Field-Weighted Citation Impact

## Field-Weighted Citation Impact



# The **Example** Again

Entity	Subject Area	Citation/Paper
Researcher A	Clinical Medicine	3
Researcher B	Mathematics	1

Which one **Performance** is Better?

Entity	Subject Area	<b>Field-Weighted Citation Impact</b>
Researcher A	Clinical Medicine	1.5
Researcher B	Mathematics	3

# Mashhad University of Medical Sciences Research Performance

Publications

4,252 ▲

Citations

12,552

Authors

4,359 ▲

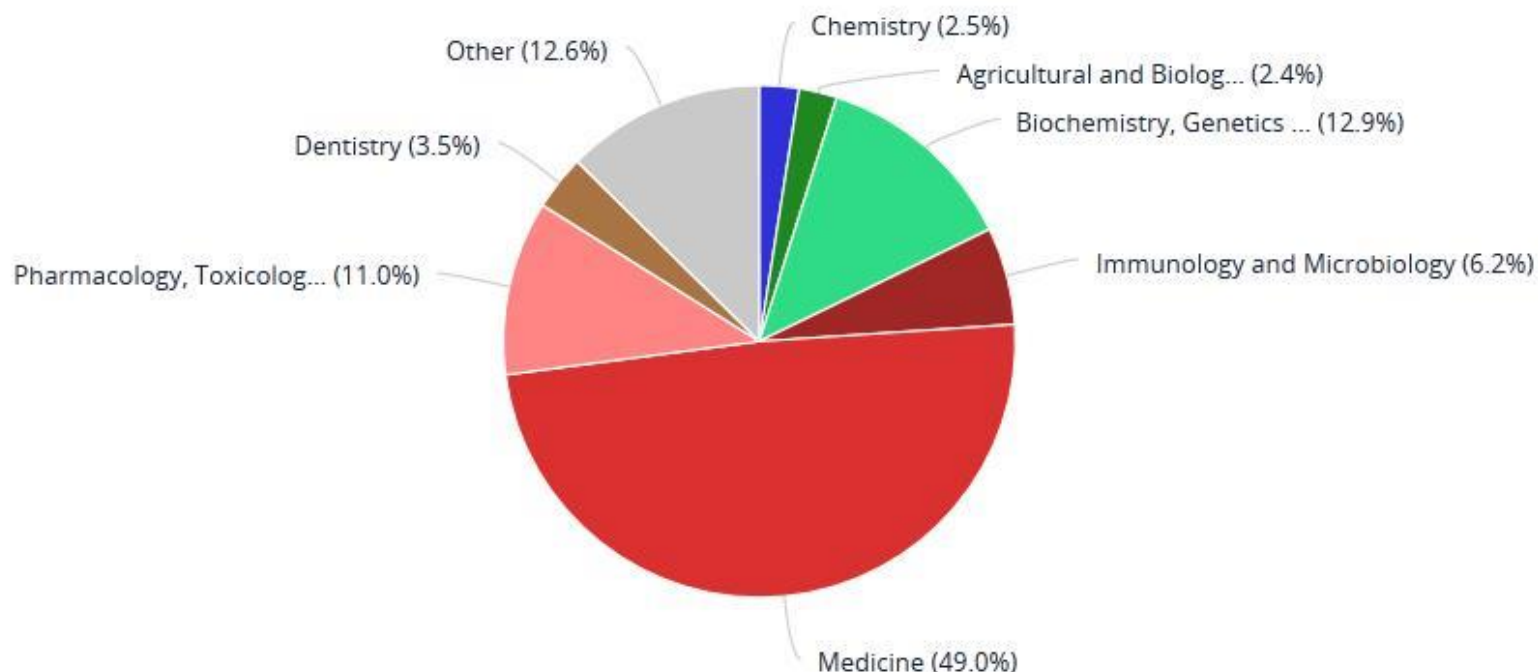
Field-Weighted Citation Impact

0.76

























Citations per Publication

3.0

[View list of publications](#)



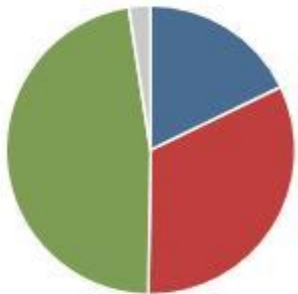
# شاخص FWCI دانشگاه علوم پزشکی مشهد

Journal Category	Publications 	Citations	Authors 	Field-Weighted Citation Impact
Rheumatology	21 	137	51 	0.73
Physiology (medical)	20	61	48 	0.38
Biochemistry (medical)	19 	92	43 	1.25
Transplantation	18 	16	29 	0.21
Emergency Medicine	12 	42	28 	0.66
Health Policy	12 	42	29 	0.66
Health Informatics	8 	9	13 	1.07
Anatomy	7 	10	27 	0.31
Rehabilitation	7 	17	16 	1.26
Histology	5 	12	13 	0.46
Embryology	2 	13	7 	1.57
Geriatrics and Gerontology	2 	17	8	2.37

# Top Authors Based on H-Index

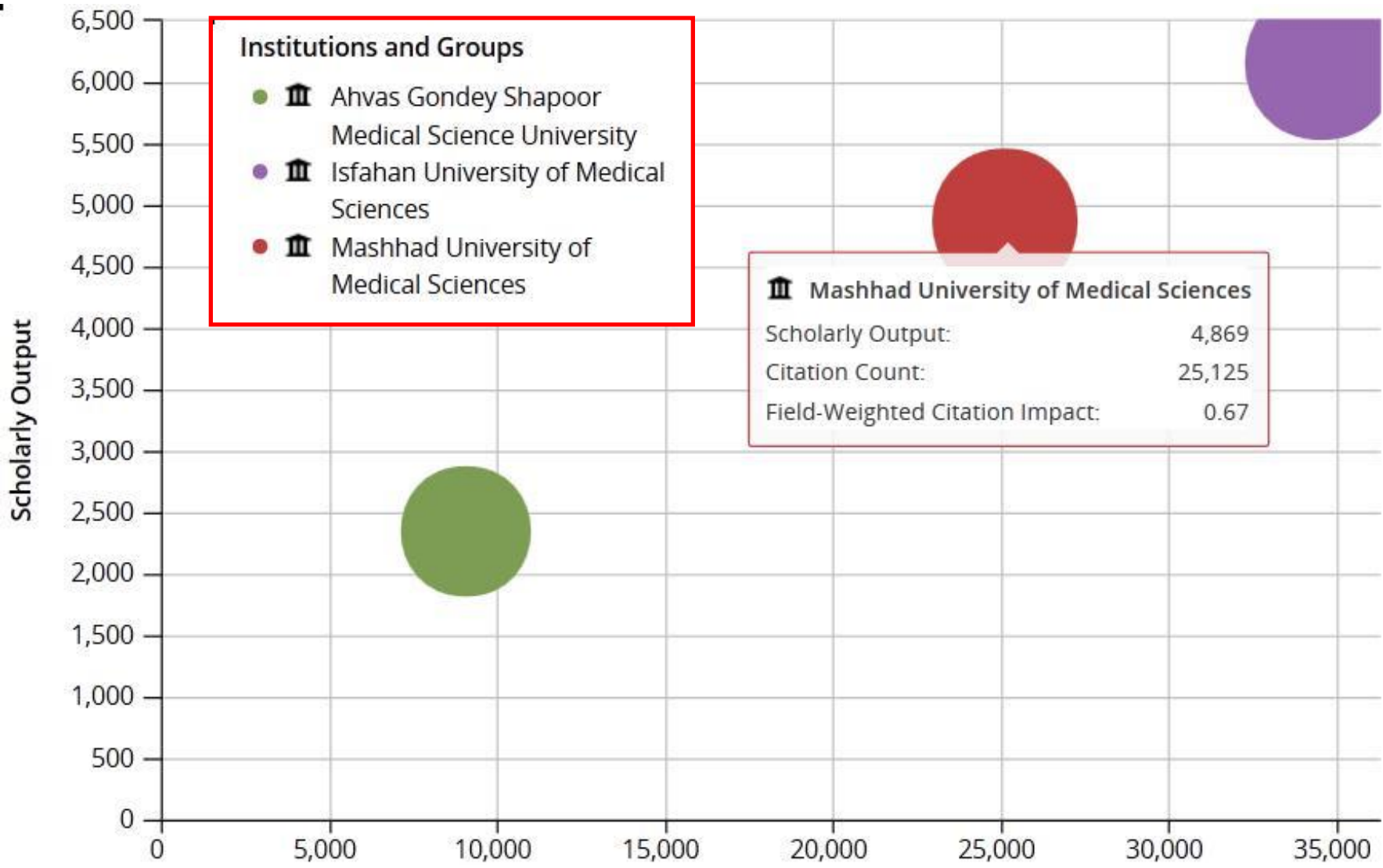
Name	Publications	Most recent publication	Field-Weighte	<i>h</i> -index
1. Hosseinzadeh, H.	83	2014	1.23	33
2. Boskabady, Mohammad Hossain	63	2014	0.93	22
3. Iranshahi, Mehrdad	59	2014	1.08	19
4. Ghayour-Mobarhan, Majid	101	2014	0.86	19
5. Darroudi, Majid	18	2014	2.12	18
6. Sahebkar, Amirhossein	138	2014	1.77	18
7. Jaafari, Mahmoud Reza Eza	34	2014	1.39	18
8. Mahmoudi, Mahmoud M.	59	2014	0.54	18
9. Ramezani, Mohammad	50	2014	0.89	17
10. Karimi, Gholamreza H.	23	2014	0.92	16

# Institutional, National & International Collaboration & Its Effect on Citation



Metric		Publications	Citations per Publication
International collaboration	17.8%	641	5.3
Only national collaboration	32.4%	1,168	3.5
Only institutional collaboration	47.3%	1,704	2.7
Single authorship (no collaboration)	2.4%	87	2.8





Scholarly Output 

Citation Count 

Field-Weighted Citation Impact 

# قابلیت شخصی سازی برای مجموعه ها

Name	Publications ▾	Most recent publication	Field-Weighte ▾	<i>h</i> -index
1. Malekzadeh, Reza	136	2014	7.65	47
2. Islami, Farhad	61	2014	1.73	27
3. Boffetta, Paolo A.	50	2014	1.75	92
4. Kamangar, Farin	44	2014	1.59	37
5. Pourshams, Akram	37	2014	1.33	24
6. Dawsey, Sanford M.	29	2014	1.52	49
7. Merat, Shahin	29	2014	0.90	23
8. Abnet, Christian C.	28	2014	1.57	49
9. Poustchi, Hossein	27	2014	0.95	11
10. Sotoudeh, Masoud	27	2014	1.11	24

اگر میل داشتید Email بزنید !

[kabiri@research.ac.ir](mailto:kabiri@research.ac.ir)

[kabiri@behdasht.gov.ir](mailto:kabiri@behdasht.gov.ir)