



# Einführung in die Bibliometrie

Juan Gorraiz

Bibliometrie und  
Publikationsstrategien

## Agenda

- Hintergrund: Was ist Bibliometrie?
- Anwendungen der Bibliometrie
- Bibliometrische Datenquellen
- Die unerträgliche Bedeutung der Zitierungen im „Publish or Perish“ Dilemma und die multidimensionale Welt der Indikatoren
- Neues Zeitalter, neue Metriken, neue Herausforderungen

## Bibliometrie: die Qual der Definition

- Anwendung **quantitativer** Methoden zur **Beschreibung** und **Erklärung** der Prozesse schriftlicher Mitteilungen und ihrer Rezeption ("**Impact**"), sowie der Natur und der Entwicklung eines Wissenschaftsgebietes, durch **Zählung** und **Analyse** ihrer verschiedenen Aspekte
- **Vermessung** der Wissenschaft (= Szientometrie)
- **Metrik** der Kommunikations-/Informationsprozesse (= Informetrie)

## Bibliometrie: „back to the roots“

### Francis Galton (1822 – 1911)

Daktyloskopie, Eugenik, Antizyklon, Composite Portraits...  
der letzte „Gentleman-Wissenschaftler“,  
„a universal genius“

### Korrelation & Regression

### Percentiles

„English men of Science“

Historiometry

obsessives Messen und Zählen!



## Missbräuchliche Verwendung der Bibliometrie

Sir Francis Galton (1822 – 1911)



# Pinselstriche

=



# Maschen

## Bibliometrische Pioniere

### Paul Otlet (1868-1944)

Gründer der Dokumentationswissenschaft

Gründer des Universal Bibliographic  
Repertory (15 Millionen Karteikarten) (1895)

1896 hat Otlet ein gebührenpflichtiges Service  
eingerrichtet: Fragen konnten per Post gestellt  
werden.

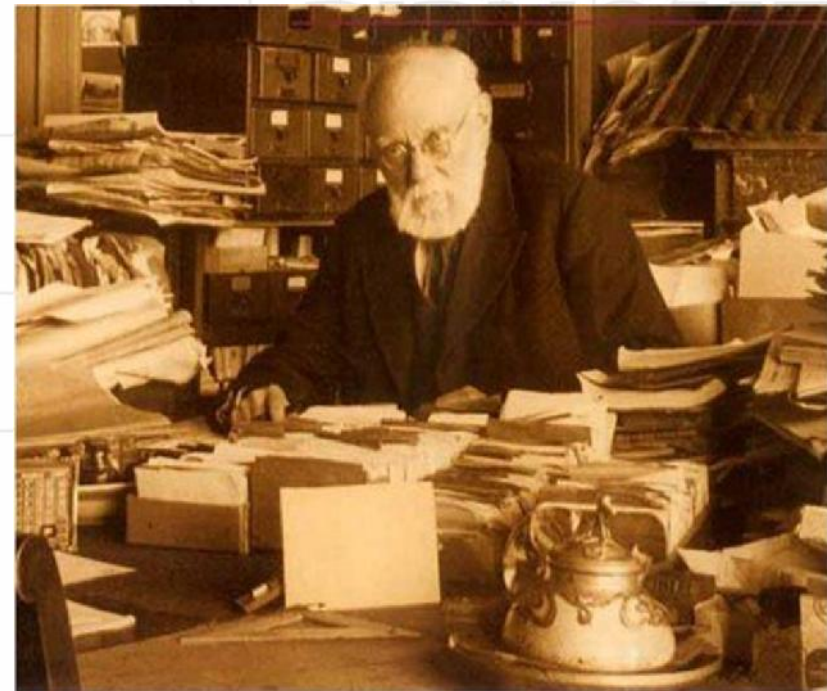
Als Antwort bekamen die Kunden die für die  
Fragestellung relevanten Karteikarten zugesandt

Gründer der "Universal Decimal Classification",  
eines der bedeutendsten Klassifikationen  
Beispiel für eine facettierte Klassifikation.

Speicherung von bibliographischen Daten auf  
Mikrofilm

In Richtung einer „city of knowledge“: „Palais  
Mondial“ or „The Mundaneum“

„Traité de Documentation“: le livre sur le livre“  
Bibliometrie: „les mesures relatives au livre et au  
document“



## Bibliometrische Pioniere

### Derek de Solla Price (1922- 1983)

Mitbegründer der Szientometrie

Seit Mitte des 17. Jhdt. verdoppelt sich das Wissen etwa alle 15 Jahre.

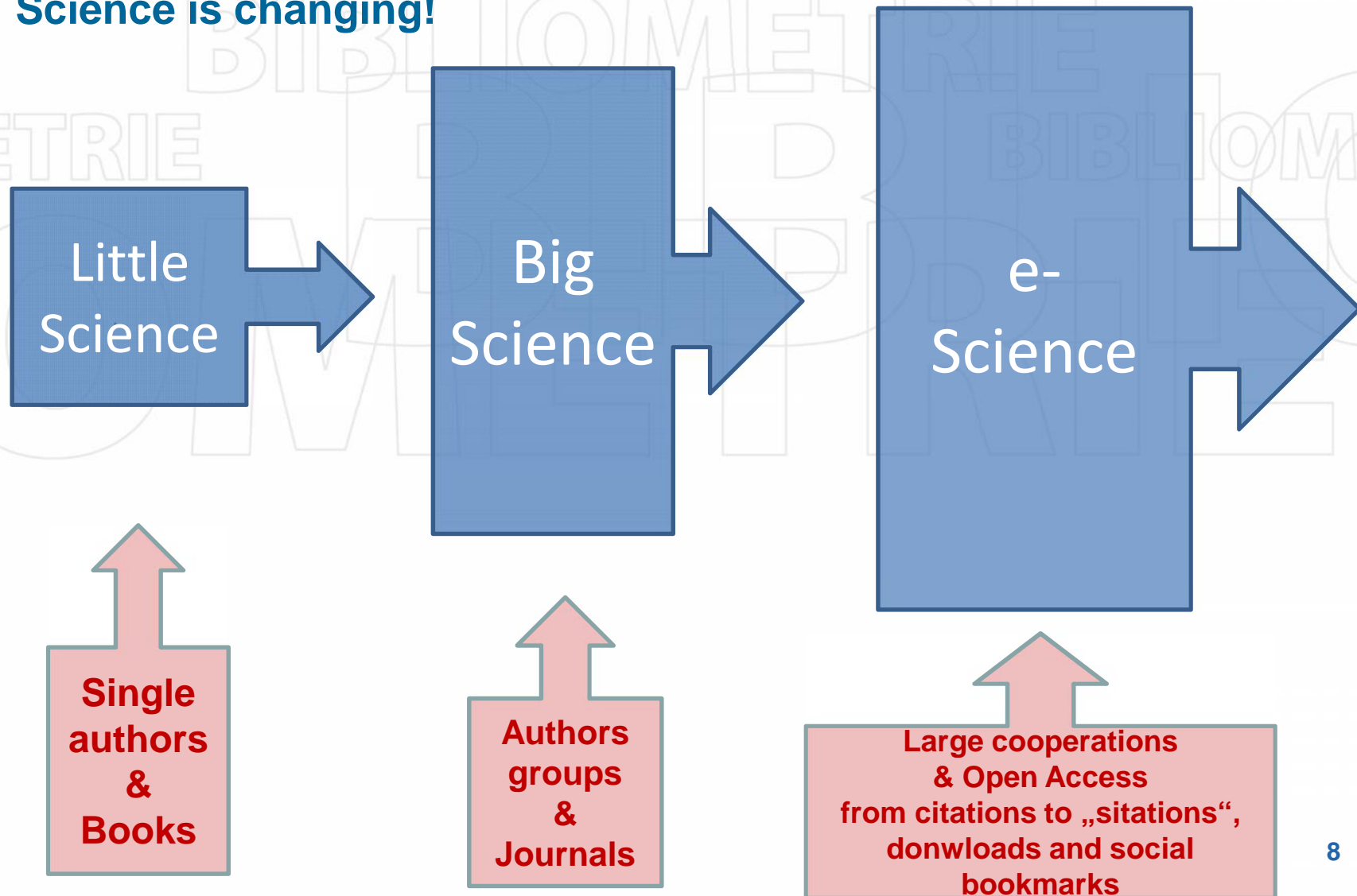
Man schätzt, dass 80 bis 90 Prozent aller WissenschaftlerInnen, die jemals gelebt haben, unsere Zeitgenossen sind.

Autor des Buches:

„Little science, big science“



Science is changing!



## Bibliometrische Pioniere

**Robert K Merton (1910-2003)**  
„the mertonian paradigm“

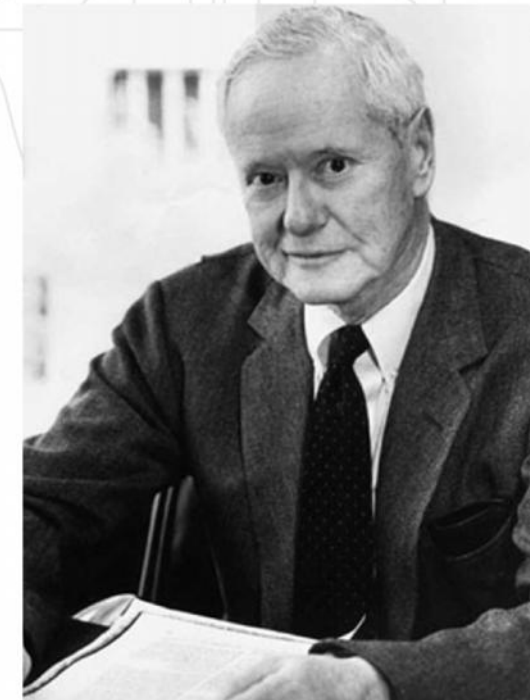
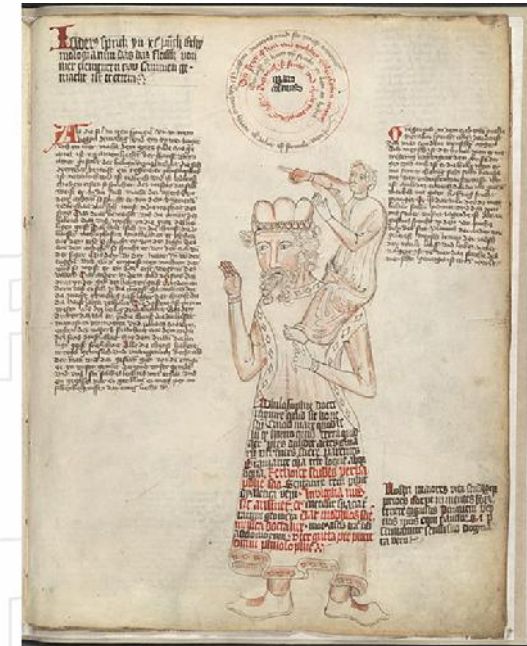
„Matthew effect“: positive Rückkopplung:  
bekannte AutorInnen werden häufiger zitiert als unbekannte und  
werden dadurch noch bekannter („success breeds success“).

„On the shoulders of giants“:  
„Ein Zwerg, der auf den Schultern eines Riesen steht, kann  
weiter sehen als der Riese selbst.“

Gemäß Merton beruht der wissenschaftliche Wert von Zitationen auf  
dem Handeln der WissenschaftlerInnen nach einem Verhaltenskodex :  
CUDO

- ✓ Communism (Kommunismus)
- ✓ Universalism (Vielseitigkeit)
- ✓ Disinterestedness (Uneigennützigkeit)
- ✓ Organized Skepticism (gesundes Misstrauen)

**Bibliografische Zitate sind gültige Indikatoren!**  
**Are you a „mertonian“?**



## „Epistemological (mertonian) model“

Career advancements,  
prizes, **“Esteem  
indicators”**

Peer review, **Citations,**  
**Downloads, Altmetrics**

Papers, Articles  
Reviews, Reports

Mental & expe-  
rimental  
stage



## Bibliometrische Pioniere

### Thomas Kuhn's (1922- 1996)

#### Theory of scientific progress

- Wissenschaft als Wechselspiel zwischen Phasen der Normalwissenschaft und der wissenschaftlichen Revolutionen
- Ein wichtiges Konzept ist hierbei das des [Paradigmas](#). Paradigmen sind „*konkrete Problemlösungen („puzzle solving“) die die Fachwelt akzeptiert hat*“
- Eine Revolution ist nach Kuhn stets mit einem [Paradigmenwechsel](#) verbunden. Paradigmen von Theorien, die durch eine Revolution getrennt sind, bezeichnet Kuhn als [inkommensurabel](#), nicht mit gleichem Maß messbar.
- WissenschaftlerInnen sind keine CUDOS-Wesen (Merton), sondern **konservative** Individuen, die tendenziell dazu geneigt sind, **Neues zu unterdrücken** und in ihrem Schaffen hauptsächlich an alltäglichen konkreten Problemlösungen ausgerichtet sind (normale Wissenschaft).



## Bibliometrische Pioniere: Eugene Garfield (1925,-)

Pionier der empirischen Informationswissenschaft  
Gründer von ISI (Institute for Scientific Information) (1960)

1955: *Current Contents*  
Idee in "Science": Erfassung der Zitationen

1963: *Science Citation Index (600 Journals)*  
Selection criteria for "Science Citation Index"  
(Bradford's concentration law)

1972: Idee in "Science": "Citation Analysis as a tool  
in journal evaluation"

1975: First Edition of Journal Citation Reports (IFs)

1997: ISI launches the Web of Science (7.500 Journals)

2004: ISI changes to Thomson Scientific

2012: Thomson Reuters: ISI Web of Science (>10.000 Journals)



## Wie arbeitet die Bibliometrie?

- **Quantifizieren** → Messen (Publikationen, Zitaten, etc.)
- **Vergleichen** → nur „like with like“, benchmarking
- **Visualisieren** → neue Aspekte und Verknüpfungen entdecken:  
*Mapping Science*
- **Interpretieren** → Validierung der Proxys,  
Limits einschätzen, etc

*„Die Welt der Bibliometrie ist multidimensional... und asymmetrisch“*

**→ Verwenden Sie niemals einen Einzelindikator!**

## Die drei Bausteine (“bibliometric laws”)

Sie zeigen die empirische Beziehung zwischen Quellen und die von ihnen produzierten Einheiten nach drei verschiedenen Aspekten:

- *a few authors turn out to be responsible for the largest portion of scientific literature in a given research field (Lotka)*
- *a few scientific journals seem to concentrate the literature required to satisfy their needs (Bradford)*
- *a relatively small number of recurrent word units govern their (and not only their) linguistic habits (Zipf)*

und stellen die Ungleichmäßigkeit in der Struktur der Informationsprozesse dar!

## Bibliometrie und Szientometrie

- Erste bibliometrische Untersuchungen von Bibliothekaren für bibliothekarische Zwecke (1927, Gross and Gross)
- Entwicklung der Bibliometrie zu einem Instrument für Evaluation und Benchmarking
- Neue Herausforderung als Messung der Wissenschaft: die Entstehung der Szientometrie

## Anwendungen der Bibliometrie

- Instrument zur Unterstützung und Analyse bibliothekarischer Entscheidungen
- Beschreibung und Erklärung der Natur und Entwicklung eines Wissenschaftsgebietes bzw. der Wissenschaft: „Emergence and Innovation“ → „Scientometrics“
- Kompass für WissenschaftlerInnen im Informationsdschungel
- Entwicklung der Bibliometrie zu einem Instrument für Evaluation und Benchmarking: Quantitative Analysen des wissenschaftlichen Outputs von Personen, Institutionen und Ländern → Academic Evaluation + Science Policy

## Bedeutung als Ergänzung von „Peer Review“

Moderne akademische Evaluation stützt sich auf eine Mischung (Balance) von **Peer Review** (qualitativ und subjektiv) und **bibliometrischen Methoden** (quantitativ und objektiv)

Krise der Peer Review Methode:

- a) teuer und sehr zeitaufwändig
- b) immer weniger Peers
- c) subjektive Natur => “old boys” Theorie
- d) Multidisziplinarität erschwert die Auswahl von Peers

→ Verbesserung und Demokratisierung des Peer Reviews  
(Delphi Method, Global Brain, etc.)

## Bibliometrische Analysen

auf 3 Ebenen:

- Mikro (Scientists, Research groups)
- Meso (Institutions, Universities, etc.)
- Makro (Countries, World-Trends)

mit besonderer Berücksichtigung von:

- Vollständigkeit und Zuverlässigkeit der bibliometrischen Daten
- Abdeckung des analysierten Fachgebietes in den Datenbanken
- Verwendung der geeigneten Proxys
- Statistische Probleme

## Bedingungen für bibliometrische Analysen (VR<sup>4</sup>)

- ❑ **Validity** – measuring what is intended to measure
- ❑ **Relevance** – results must be meaningful in the context of what they are produced for
- ❑ **Reliability** – assessability of correctness of measures and errors
- ❑ **Reproducibility** – Under identical conditions results should be reproducible
- ❑ **Robustness** – results may not be sensitive to marginal fluctuations or minor corrections in the underlying data sources

## Bibliometric data sources

3 Standards:

- Science Citation Index → Web of Science
- Scopus (2004)
- Google Scholar\*

Alle sind multidisziplinäre Datenbanken und enthalten die Affiliations aller Autoren

Alternativen:

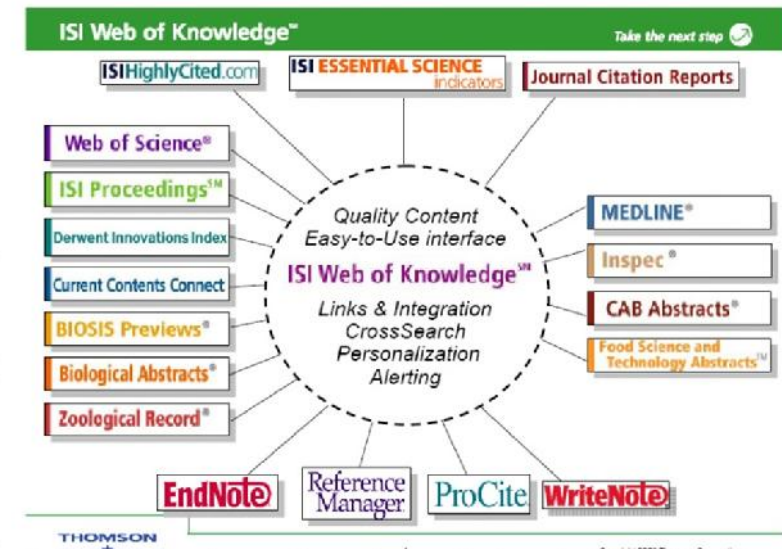
- arXiv.org / SPIRES
- Citebase
- Citeseer
- Chemical Abstracts
- MathSciNet
- ProQuest

## Warum internationale multidisziplinäre Zitationsdatenbanken?

- Revolutionierung des Information Retrieval:  
Zitationszusammenhänge komplementär zu kontrolliertem Vokabular  
(Garfield's Idee mit Zitierungen → Brin & Page mit Hyperlinks → Google  
Page Algorithmus)
- Suche nicht nur in den indizierten Datenquellen, sondern auch in den  
zitierten Referenzen
- Netzwerkanalysen aufgrund vollständiger Affiliationsangaben möglich
- Grundlage der wichtigsten Universitätsrankings → Erhöhung der  
Sichtbarkeit der Institution
- Kompass für WissenschaftlerInnen im Informationsdschungel und als  
Optimierung der wissenschaftlichen Karriere

## Web of Science (WoS) : Vorteile

- Qualitätskriterium:  
nur ausgewählte Zeitschriften (Bradford-Garfield Gesetz)
- Beeindruckender Umfang (ca. 13.000 Zeitschriften, viele zurück bis 1900, „Century of Science“ )
- Personalisierung (Researcher ID)
- Alert Services (Recherchen, „Citation alert“)
- ausgefeilte Zitations-Metrik (standardisierter Impact Factor - über JCR)
- Hilfstool für wissenschaftliche Administration (Evaluation)
- Mehrere Features in der Funktion als „Kompass“ für WissenschaftlerInnen



*Eine Publikation in WoS hat „mehr“ Wert als eine Publikation in anderen Datenbanken!*

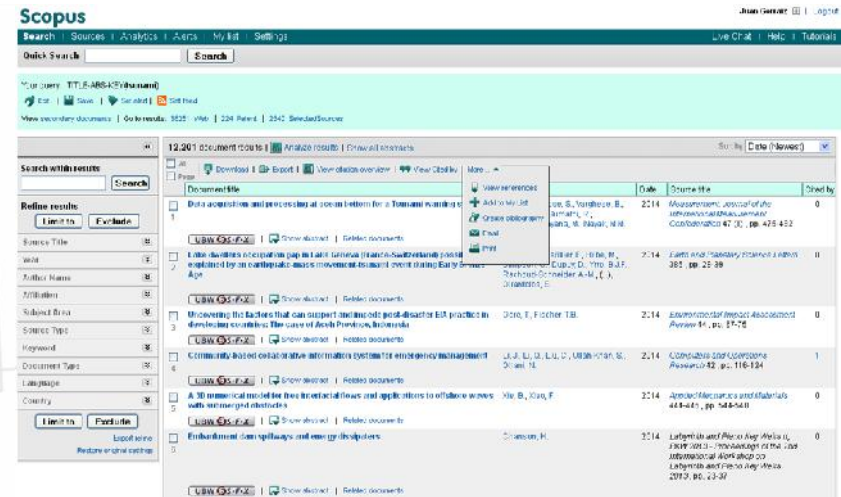
## Scopus (Elsevier) seit 2004

### Vorteile:

- Imposantes Volumen: größte Datenbank: fast doppelt so viele Journals als die Konkurrenz (Web of Science)
- Nicht englischlastig, europäische Zeitschriften stärker vertreten
- Inkludiert auch "Articles-in-Press"
- Volle Abdeckung von ca. 3800 Journals (Elsevier, Springer, Nature)
- Innovativ - State of the Art-Oberfläche (Filter, Citation Overview, getrennte Author bzw. Affiliation Search, etc. → von der Konkurrenz nachgeahmt)
- Eigene Zeitschriftenmetrik: SJR & SNIP
- Journal Analyzer ermöglicht direkten Vergleich von Journals
- Download von Volltexten integriert
- Inkludiert auch Patente, Repositorien, etc. (SciVerse)
- Seit kurzem altmetrics!

### Nachteile:

- Keine strengen und konsequenten Selektionskriterien
- vor 1996 lückenhaft
- (noch) nicht so renommiert wie „Web of Science“



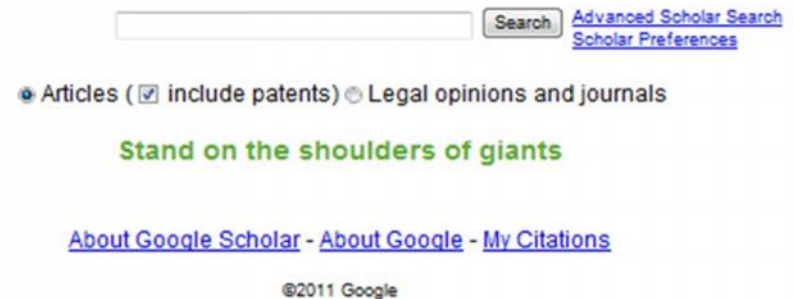
## Google Scholar

### Vorteile:

- Riesenumfang: akademische Verlage, Repositorien, OA Journals, etc. (<http://scholar.google.at/intl/de/scholar/about.html>)
- frei verfügbar
- keine Einschränkung auf Dokumenttypen
- Verbindung zu OCLC WorldCat und Google Books
- Volltext, Abstract, Zitation
- Relevanz nach Ranking-Technologie von Google
- eigene Metrik: „Publish or Perish“

### Nachteile:

- geringe Zuverlässigkeit und Reproduzierbarkeit
- geringe Inhaltskonsistenz → blackbox → keine Information über „coverage“
- keine Datenbankstruktur → nur basale Such- und Filterwerkzeuge



## Die unerträgliche Bedeutung der Zitierungen im „Publish or Perish“ Dilemma und die multidimensionale Welt der Indikatoren

Wie wird der wissenschaftliche Output gemessen?

- Produktivität
- Sichtbarkeit
- Impact
- Vernetzung
- Knowledge Base

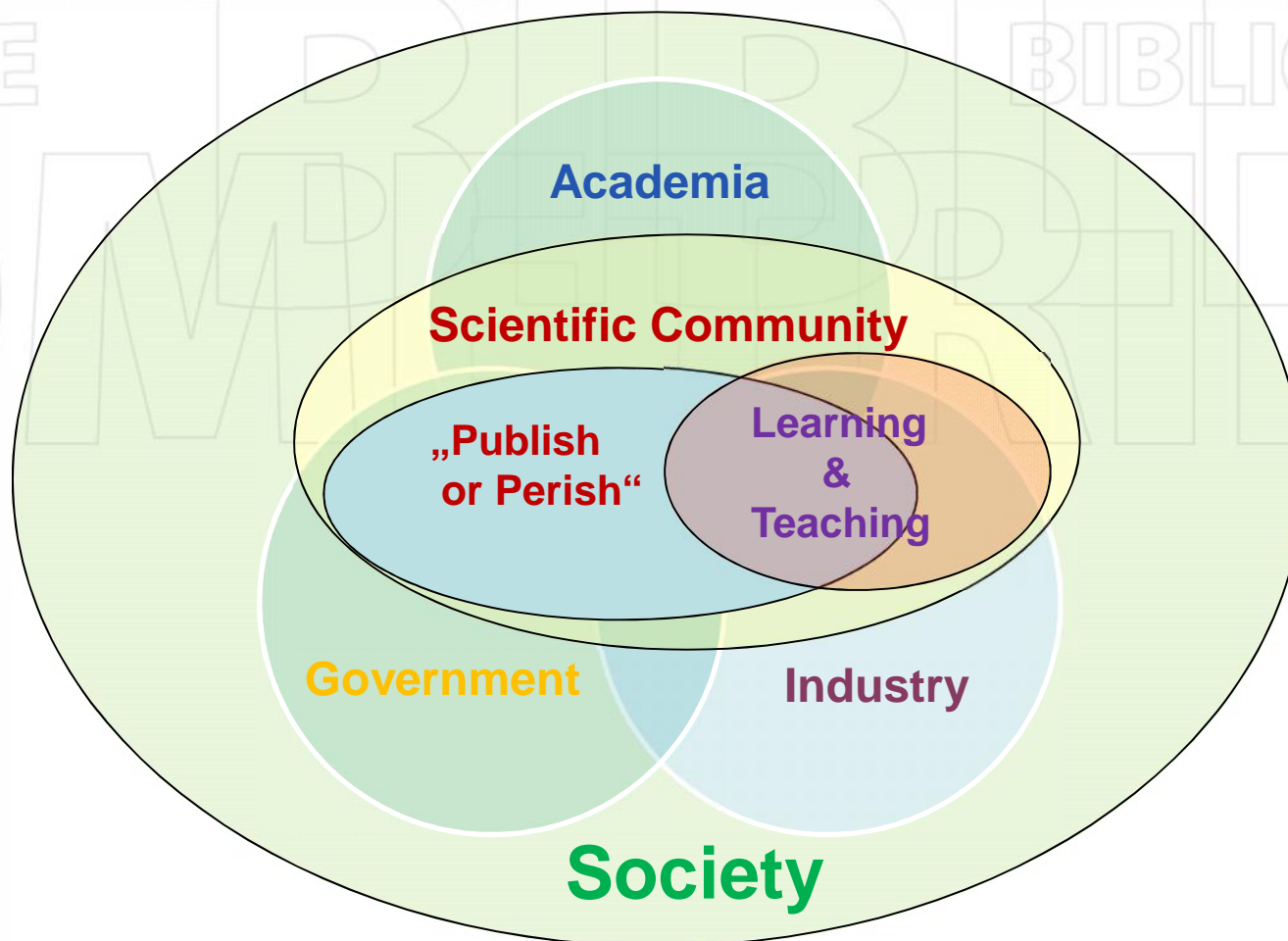


## The shortcomings of citations

- **highly controversial:** different reasons why to cite whom
- **only indirect measure of impact:** not for quality!
- **always delayed:** significance varies according to discipline
- **insufficient:** in some fields; surprisingly low mean citation rates for vast majority of published scientific literature
- **skew distribution:** 20/80 rule (see percentiles)



Target groups: Societal Impact



## Neues Zeitalter, neue Metriken, neue Herausforderungen

Internet Zeitalter →

Hyperlinks statt citations →  
from citations to “sitations” → webometrics  
prestige metrics → Google PageRank

e-Media (e-journals) Zeitalter →

Usage metrics (downloads, views)

Web.2.0 (social media) Zeitalter →

Altmetrics based on data from the social web: readerships, visits,  
discussions, etc.

Not only a metric but also a „movement“ with a manifesto (Priem et al. 2010)



## What really defines impact? How can it be measured?



*"My question is: Are we making an impact?"*

[http://www.condenaststore.com/sp/My-question-is-Are-we-making-an-impact-New-Yorker-Cartoon-Prints\\_j8562874\\_.htm](http://www.condenaststore.com/sp/My-question-is-Are-we-making-an-impact-New-Yorker-Cartoon-Prints_j8562874_.htm)



**Danke für Ihre Aufmerksamkeit!**