////// studie / articles //////////////////////////////////////

"NEEDLESS TO SAY MY PROPOSAL WAS TURNED DOWN." THE EARLY DAYS OF COMMERCIAL CITATION INDEXING, AN "ERROR-MAKING" (POPPER) ACTIVITY AND ITS REPERCUSSIONS TILL TODAY

Abstract: *Today university rankings and performance rankings (often based on JIFs, h-indexes) are believed to be indispensable to assure scientific "quality". Most of these performance rankings employ citation data provided by Thomson Reuters. TR's current influence on funding decisions, individual careers, institutions, disciplines and countries is immense and ambivalent. There is increasing resistance against "impactitis" and "evaluitis". Usually overseen: Trivial errors in TR's citation indexes (SCI, SSCI, AHCI) produce severe non-trivial effects: Their victims are authors, institutions, journals with names beyond the ASCII-code and scholars of humanities and social sciences. Based on the Joshua Lederberg Papers I claim: To overcome severe resistance Eugene Garfield and Joshua Lederberg had to foster overoptimistic attitudes and to downplay the severe problems connected to global and multidisciplinary citation indexing. The difficulties to handle different formats of references and footnotes, non-Anglo-American names, and of publications in non-English languages were known to the pioneers of citation indexing.*

Keywords: *evaluation; rankings; errors; scientometrics; critical science studies*

„Není třeba zmiňovat, že můj projekt byl zamítnut." Počátky komerčních citačních indexů, dělání chyb podle Poppera a jejich dnešní následky

Abstrakt: *Dnešní žebříčky univerzit a výkonnosti (často založené na JIF a h-indexech) jsou považovány za nepostradatelné pro zajištění vědecké „kvality". Většina z těchto žebříčků produktivity využívá citační údaje poskytnuté Thomson Reuters. Současný vliv TR na rozhodování o financování, na individuální kariéry, instituce, obory a země je ohromný a ambivalentní. Odpor vůči „impaktitidě" a „evaluatitidě" se zvyšuje. Obvykle je přehlížena skutečnost, že triviální chyby v citačních indexech TR (SCI, SSCI, AHCI) mají závažné, netriviální následky: jejich obětmi jsou autoři, instituce, časopisy vymykající se ASCII-kódu a akademici v humanitních a sociálních vědách. Na základě rozboru Joshua Lederberg Papers tvrdím, že aby překonali tvrdý odpor, Eugene Garfield a Joshua Lederberg museli protěžovat přehnaně optimistické postoje a zlehčovat vážné problémy spojené s globálními a multidisciplinárními citačními indexy. Obtíže plynoucí z různých formátů odkazů a poznámek, jiných než anglo-amerických jmen a publikací v jiných jazycích než v angličtině byly známy již průkopníkům citačních indexů.*

Klíčová slova: *evaluace; žebříčky hodnocení; chyby; scientometrie; kritická studia vědy*

TERJE TÜÜR-FRÖHLICH
Department of Philosophy and Philosophy of Science
Johannes Kepler University
4040 Linz /Austria
email / terje.tuur@jku.at // url / http://www.iwp.jku.at/tuur/

Introduction: Technical Terms Used

This paper uses several technical terms from bibliometrics and scientometrics, which will be explicated briefly in the following: (1) *Quantitative evaluation* of scientific achievements means the counting and analysis of scientific achievement in terms of input (funding), output (productivity) and impact (citations). (2) *Citation Indexing*: Common bibliographies or literature databases provide bibliographic information, keywords, and abstracts. Citation indexes provide (to be precise: *should* provide) the complete and error-free reference lists of all covered citing documents. (3) *SCI/SSCI/AHCI*: SCI – the Science Citation Index was the first one (1964), later followed by the SSCI (1973) – the Social Sciences Citation Index and at last by the AHCI – the Arts and Humanities Citation Index (1978). These indexes were originally launched as voluminous paper-based reference books by the Institute for Scientific Information (ISI), a private firm in Philadelphia/USA, leaded by its founder Eugene Garfield. Currently, these indexes are offered as licenced online databases and are owned by Thomson Reuters (in the following: TR). TR's citation databases are very selective and contain only a marginal share (currently approx. 16,000 journals) of all scientific journals worldwide (their total number is estimated about 50,000 to 100,000). (4) *Web of Science* is the web-based pay-for-content service by Thomson Reuters, offering SCI, SSCI, ACHI, also some newer but smaller citation indexes like conferences or books. (5) *Thomson Reuters*: Huge North American media corporation. (6) The *Journal Impact Factor (JIF)* is specified by co-inventor Eugene Garfield (the second was Irving Sher) in the following way: "A journal's impact factor is based on 2 elements: the *numerator*, which is the number of citations in the current year to items published in the previous 2 years, and the *denominator*, which is the number of substantive articles and reviews published in the same 2 years" (italics added by TTF).[1] A simple fictitious example: Any item of journal ABC which had been cited (in TR's citation databases) N = 300 times in total, in the years of 2010 and 2011. Journal ABC has published n = 30 "citable" articles in 2010 and 2011. The JIF of 2012 is 10. (7) The *index h*

This article is based on an earlier version (November 2013) of Terje TÜÜR-FRÖHLICH, *The Non-trivial effects of Trivial Errors in Scientific Communication and Evaluation*. Johannes Kepler University Linz, Doctoral Thesis, unpublished, 2014. I would like to thank to Volker Gadenne, Gerhard Fröhlich, Ingo Mörth and two anonymous referees for their critical feedback and valuable suggestions.

[1] Eugene GARFIELD, "The History and the Meaning of the Journal Impact Factor." *JAMA*, vol. 295, 2006, no. 1, p. 90 (90–93).

or h-index or Hirsch Index in the words of his inventor, J. E. Hirsch: "I would like to propose a single number, the 'h-index', as a particularly simple and useful way to characterize the scientific output of a researcher. A scientist has index h if h of his/her N_p papers have at least h citations each, and the other (N_p – h) papers have no more than h citations each."[2] Two simple fictitious examples: author A publishes n = 3 articles X, Y, Z (in co-authorship with 5 colleagues). X has been cited n = 5, Y n = 4, Z n = 3 times. The h-index of all co-authors based on these publications is 3. Author B publishes 3 articles (as single author) U has been cited n = 500, V n = 200, W n = 3 times. Her/his h-index is also 3.

1. DORA and Citation Indexing as "Error-Making Activities"

Today university rankings, quantitative evaluation of publications by JIF (Journal Impact Factor) or researchers by HI (Hirsch-Index) are believed to be indispensable instruments for "quality assurance" in the sciences – at least from the perspective of politicians**,** science administrators and science policy makers as well as many scientometricans.

1.1 DORA, References, Database Errors

But a growing number of learned societies, journals, scientific institutions and scientists/ scholars argue and campaign against the "almighty" *journal impact factor*, based on citation indexing (both produced by the media corporation Thomson Reuters). The most famous initiative of protest and recommendations is named DORA, *The San Francisco Declaration on Research Assessment*[3] (one of the first organizational signers: The Academy of Sciences of the Czech Republic). Worldwide more and more oppositional action groups of scientists / scholars, librarians, journals, universities, research funds and scientific associations stand up against *university rankings* and emphasize their negative effects on scientific personnel (especially early career scientists) and scientific development.

My point of criticism of commercial citation indexes is the *tremendous amount of trivial errors* in their database records, e.g. misspellings, typos,

[2] Jorge E. HIRSCH, "An Index to Quantify an Individual's Scientific Research Output." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, 2005, no. 46, pp. 16569–16572.

[3] *San Francisco Declaration on Research Assessment. Putting science into the assessment of research* [online]. 2013ff. Available at: <http://am.ascb.org/dora/> [cit. 3. 11. 2014].

mistakes, even mutations and mutilations of author, journal and institutions' names; misclassifications of documents; non-indexed references. These errors, inconsistencies and losses end in citation calculation losses. They negatively affect the evaluation scores of authors, journal, institutions and countries involved: The consequences of lower citation rates and lower positions in rankings provoke lower chances for funding, research topics, careers and visibility. Following Sir Karl Popper, I think that sciences are "error making activities". Scientific documentation, citation indexing, scientific evaluations are error-making activities, too.

1.2 Sciences As Error Making Activities (Popper)

The Austrian philosopher of science Sir Karl Popper conceptualizes sciences as *error making activities*:

> We are all fallible, and it is impossible for anybody to avoid all mistakes, even avoidable ones. The old idea that we must avoid them has to be revised. It is mistaken and has led to hypocrisy. Nevertheless, it remains our task to avoid errors. But to do so we must recognise the difficulty [...] Errors may lurk even in our best-tested theories. It is the responsibility of the professional to search for these errors [...] For all these reasons our attitude towards mistakes must change [...] The old attitude leads to the hiding of our mistakes and to forgetting them. Our new principle must be to learn from our mistakes so that we avoid them in future; this should take precedence even over the acquisition of new information. Hiding mistakes must be regarded as a deadly sin. It is therefore our task to search for our mistakes and to investigate them fully.[4]

In other words: Popper thinks that *to detect, to (publicly) correct and to retract errors* is important for the progress of knowledge accumulation. Some followers of Popper claim: "Yes, Popper demands the correction of errors, but he means only the 'important, theoretical errors' i.e. errors in theories." But the above mentioned quotation is from Popper's paper co-authored by the medical ethics expert Neil McIntyre, titled "The critical attitude in medicine: the need for a new ethics." This article discusses banal medical errors – for example forgotten operation instruments in patients' bodies. Therefore I think that Popper would have recommended to learn

[4] Neil McINTYRE – Karl POPPER, "The Critical Attitude in Medicine: The Need for a New Ethics." *British Medical Journal*, vol. 287, 1983, p. 1920 (1919–1923).

from *any* kind of errors – including the trivial errors – and to criticize all errors publicly in order to learn from them.

1.3 Research Theses and Methods

Complementary to Popper's standpoint the research theses of my dissertation are:

> (T1) *Trivial errors are of high relevance in the evaluation context*. Under today's evaluation pressure, the not detected, not publicly eliminated or retracted errors can be important for the "sake" of the careers of the scientists and their institutions.
> (T2) Trivial errors are associated with biases by power structure and symbolic capital (prominence, reputation, "impact"). These Matthew and Matilda Effects – the rich get richer, the poor get poorer[5] – impinge on authors, journals, institutions, scientific disciplines and fields, countries. They are linked with language biases and gender inequality in sciences. These errors and biases tend to persist, to interact with each other and to exaggerate.

Research theses T1-T2 were the starting points for my investigations. Due to the huge influence of Thomson Reuters' global citation databases on the evaluation of research productivity and impact I decided to conduct case studies on the data quality of Thomson Reuters' Social Sciences Citation Index (SSCI). Due to TR's non-transparent reference indexing and data quality procedures it was more than necessary to examine the historical aspect of commercial citation indexing. Therefore I added research thesis T3:

> (T3) The difficulties to handle different formats of references and footnotes, non-Anglo-American names, and of publications in non-English languages were known to the pioneers of commercial citation indexing. The blunt ignorance of lingual, disciplinary and cultural differences have led to errors and to the underestimation of errors, in other words: "The tomato (i.e. the first citation index SCI) was rotten from the beginning".

This article is concentrating on research thesis T1 and T3.

The investigation employs the following *non-reactive methods*: Systematic literature search and critical overview; critical investigation of the struc-

[5] Robert K. MERTON, "The Matthew Effect in Science." *Science*, vol. 159, 1968, no. 3810, pp. 56–83; Margaret W. ROSSITER, "The Matthew Matilda Effect in Science." *Social Studies of Science*, vol. 23, 1993, no. 2, pp. 325–341.

tures of Thomson Reuters' *Social Sciences Citation Index's* data; qualitative and quantitative error analyses of SSCI record; content analysis of the *Joshua Lederberg Papers* (provided by the National Library of Medicine).

2. Limitations of Errors Research

Generally, the science and social sciences publications have discussed at least two types of errors in scientific practice: there are widely "acknowledged" errors such as errors of "type I" (tests reject the true null hypothesis) and errors of "type II" (tests fail to reject the false null hypothesis), or errors of measurement and observational errors.

The so-called *trivial errors* are e.g. typing errors, misspellings or misprints of author names or initials, journal titles, names of scientific institutions; misclassification of documents; missing entries. The general opinion of scientists is that trivial errors are of low relevance. Many scientific communication experts, especially scientometricians and database providers[6] believe that errors in scientific publications and data banks are of less importance: There are many errors, yes – *but they would counterbalance each other*. Contrary to this widespread opinion I have formulated my research these T2 – shortly: errors are not distributed randomly, but associated with strong biases (e.g. language biases) and tend to persist, to interact with each other and to exaggerate.

In the following I present a short critical overview on the error literature:

(1) The systematic literature search of *error detection, error reporting and error management literature* provides mainly psychology and management science literature, dealing with catastrophes like Chernobyl – due to human erring and disregarding the established security rules. Shortly, these results were interesting but not useful for my research. The error typologies found in this literature are interesting, but unfortunately useless in the context of my investigation.

(2) "*Typos" and "accuracy of references" studies* are found mainly in medical, nursing, library and information sciences journals. Following generalizations can be drawn from this literature: The majority of studies have classified errors either as minor or major. But there are no generally

[6] E.g. Eugene GARFIELD, *The Agony and the Ecstasy –The History and Meaning of the Journal Impact Factor* [online]. 2005. Available at: <http://garfield.library.upenn.edu/papers/jifchicago2005.pdf> [cit. 15. 10. 2013].

applicable definitions. Often an error was considered as *minor* if an article still could be located with less effort despite the erroneous information in the reference. An error was considered *major* if it would inhibit the article from being found at all. If studies consider typographical errors / misspellings at all, these "typos" are classified as minor errors. Usually the *author(s)* of the publications are blamed for error making. According to Unver et al.[7] errors in reference lists happen due to lack of attention in detail or "careless" transcription of bibliographical data, or the authors' "delegating the responsibility" of verifying reference citations to unqualified assistants: "The ultimate responsibility for accuracy lies with the authors."[8] Unver et al. even believe that only on rare occasions the inaccurate transcription of references by *editorial staff* or *printers* is responsible for bibliographic errors. Only a few publications *mention casually that databases are not error free.*[9]

(3) Interestingly, the literature concerning database biases and/or database errors in *financial analyses* (e.g. financial information on public firms)[10] shows a way more critical attitude. They criticize selection, delisting, omission and survivorship biases as well as misclassification errors and coding policies of the inspected databases and suggest methods of quality control of the competing databases. I think, information scientists could learn from this research area.

(4) Several information / computer science conference papers focus on automatized "*name disambiguation*" methods. There are following name-associated ambiguity issues: (a) One name, different persons (*homonym*). In large international multidisciplinary databases there are many authors with identical surnames and initials, especially Asian names as "Kim, L.". In order to search or to evaluate a *specific* "Kim L." it is necessary to eliminate all doubles of the wanted "Kim L". (b) Different names, but only one person *(synonym)*: Due legal and cultural traditions, life course events like marriages provoke mainly female authors to change their surname (example of an Austrian female social philosopher: from Gröbl to Steinbach-Gröbl to Gröbl-Steinbach to Schuster).

[7] Bayram UNVER *et al.*, "Reference Accuracy in Four Rehabilitation Journals." *Clinical Rehabilitation,* vol. 23, 2009, no. 8, pp. 741–745.
[8] *Ibid.*, p. 744.
[9] *Ibid.*
[10] Kellogg School of Management, *Database Biases and Errors* [online]. 2011. Available at: <http://www.kellogg.northwestern.edu/rc/crsp-cstat-references.htm> [cit. 16. 10. 2013].

The problems are challenging: often proposed as solution for name disambiguation is the combination of the author name with *institutional affiliation*. But modern science policy demands high mobility from the academics. Therefore the search strategy of matching an author name with one or two academic institutions is insufficient yet to retrieve *all* records (or citations) of the targeted person. The second proposed solution to combine a researcher's name with his/her *research field* may be of some success when searching for "narrow" specialists. But multidisciplinary researchers or authors with multidisciplinary visibility and impact (citations) and their publications cannot be isolated only by one specific research field (e.g. by SSCI one journal category).

These unsolved name disambiguation problems can lead to erroneous network study findings as well as to misleading productivity and citation ranking results. Various computer scientists are hopeful in finding reliable software solutions for name disambiguation. Yet Smalheiser – Torvik[11] realistically sum up the name disambiguation literature: "the name disambiguation represents a *major and unsolved problem* for information sciences" (italics by TTF).

3. Commercial Citation Indexing & Their Evaluation Effects

Philosophers of Science have neglected the topic of science evaluation, especially the data employed in research performance rankings. There are two exceptions: Endla Lohkivi et al.[12] analyse the "epistemic injustice" in Estonian research evaluation, in other words: Matthew and Matilda effects (disciplinary differences in publication habits lead to evaluation winners and losers). Philip Mirowski[13] takes a critical stand against "privatizing American science" and the consequences of private scientific data and knowledge ownership (e.g. patents). Mirowski attacks "the lack of openness" of the decision processes of ISI / Thomson Reuters and the transforming of their citation databases from a "helpful tool for researchers" to an evaluation tool for bureaucrats: "What started out as something harmless, rather as a the-

[11] Neil R. SMALHEISER – Vetle I. TORVIK, "Author Name Disambiguation." *Annual Review of Information Science and Technology*, vol. 43, 2009, no. 1, p.1 (1–43).
[12] Endla LÕHKIVI – Katrin VELBAUM – Jaana EIGI, "Epistemic Injustice in Research Evaluation: A Cultural Analysis of the Humanities and Physics in Estonia." *Studia Philosophica Estonica*, vol. 5, 2012, no. 2, pp. 108–132.
[13] Philip MIROWSKI, *Science-Mart. Privatizing American Science.* Cambridge, MA: Harvard University Press 2011.

saurus, has turned in a sharp-edged audit device wielded by bureaucracies uninterested in the shape of actual knowledge and its elusive character."[14] The privatization of bibliometric data leads to the "monetization of university data" (Mirowki refers to a wording by Ellen Hazelkorn).[15] "Ranking individuals, departments, academic institutions, corporations, and the like, according to their 'productivity' as well as their possible relevant to targeted intellectual property (IP), has become *Thomson's stock in trade*."[16] In other words: global public science evaluation is a huge business, based on privately owned data.

There is a widespread opinion that numeric data is objective. But the data relevant for the journal impact factor and for many university rankings are not the product of public science, guided by Robert K. Merton's scientific ethos.[17] They are not compiled according to Merton's institutional imperative "disinterestedness". The bibliographic data used are typically products of commercial activities: as mentioned above, nowadays they are collected, operated and owned by the commercial media corporation Thomson Reuters.

In the current academic evaluation era, *the visibility and impact* of publications, authors, institutions' play a crucial role not only for individual researchers but also for disciplines and organisations. Since the 1970s, the citation indexes SCI, SSCI and AHCI have had the monopolistic market position for decades. Since the millennium, there are two new competitors, which also provide citation data. In 2004 the mighty Dutch publishing company *Elsevier* launched its own subscription based bibliographic database *Scopus* (partly containing abstracts and citations). Only one year later, in 2005, the mighty global search engine *Google* initiated the free access bibliographic database *Google Scholar*. Still, the majority of citation analyses are conducted only based on Thomson Reuters' citation data.

3.1 Thomson Reuters' Influence on University Rankings

Each year the results of international rankings of academic institutions – e.g. *Times Higher Education (THE) World University Rankings* or the *U.S. News*

[14] *Ibid.,* p. 268.
[15] *Ibid.*
[16] *Ibid.,* p. 269; italics added by TTF.
[17] Robert K. MERTON, "The Normative Structure of Science." In: *The Sociology of Science: Theoretical and Empirical Investigations*.Chicago: University of Chicago Press 1973, pp. 267–278.

& World Report Collage Rankings are gaining more and more public attention as well as influence in funding and policy decision-making. Originally higher education institutions rankings were aimed to provide information to students. Currently

> Administrators consider rankings when they define goals, assess progress, evaluate peers, admit students, recruit faculty, distribute scholarships, conduct placement surveys, adopt new programs and create budgets.[18]

In multiple ways, Thomson Reuters is involved in the university ranking business:

Times Higher Education (THE) World University Ranking has been powered since 2009 by Thomson Reuters. Based on the THE World University Ranking Methodology description,[19] I computed Thomson Reuters' data influence on performance indicators (see Table 1). Some introductory remarks: THE has chosen all in all n = 13 performances indicators, grouped into n = 5 areas of evaluation. There are areas with many indicators (for example teaching is evaluated by n = 5 indicators) and there are areas with only one indicator (area citation: research influence by one indicator: TR's citation data). All of the n = 13 performances indicators and in sum all of the five areas have been assigned a "worth" in "% of the overall ranking score"[20] by THE.

One example: In the area of *Teaching: The learning environment* the highest share of the 5 performance indicators has been assigned to the results of the *Academic Reputation Survey*, "a worldwide poll of experienced scholars"[21] carried out by Thomson Reuters. The results of the Academic Reputation Survey with regard to teaching were assigned a worth of "15 percent of the overall rankings score."[22] In summary, *TR's influence on THE University Ranking is more than 70 %.*

[18] Wendy N. ESPELAND – Michael SAUDER, "Rankings and Reactivity: How Public Measures Recreate Social Worlds." *American Journal of Sociology*, vol. 113, 2007, no. 1, p. 11 (1–40).
[19] *The Essential Elements in Our World-leading Formula* [online]. 2013. Available at: <http://www.timeshighereducation.co.uk/world-university-rankings/2012-13/world-ranking/methodology> [cit. 24. 8. 2013].
[20] *Ibid*.
[21] *Ibid*. (invited only participants)
[22] *Ibid*.

Areas of Evaluation	Area's overall weight in (%)	Cumulated share of TR influence on ranking score
Teaching: the learning environment	30	15
Research: volume, income, reputation	30	24
Citations: research influence	30	30
Industry income: innovation	2.5	n.n.
International outlook: staff, students, research	7.5	2.5
TOTAL	100 %	71.5 %

Table 1: *Thomson Reuters' influence on Times Higher Education (THE) World University Ranking*
Source: THE World University Ranking Methodology,[23] own compilation (23. 8. 2013)

THE is not the only one ranking influenced by TR. Globally, there are numerous international and national college and university rankings. A first investigation of the rankings' web-profiles showed: It is more demanding than expected to find out TR's influence, because the information concerning the data bases of the rankings are often not clearly indicated. Till now, I was able to identify at least n = 16 international and national college and university rankings which are employing TR data as indicators.

3.2 General Evaluation Effects: The Gratification of the Chosen

Inter alia, Thomson Reuters' commercial activities have the following consequences:

(1) The successful propaganda of Thomson Reuters has established the common belief – such as amongst the *Taiwanese* Government[24] and the *Austrian Federal Ministry of Science and Research*[25] – that the *coverage* of

[23] *Ibid.*
[24] Chuing Prudence CHOU *et al.*, "The Impact of SSCI and SCI on Taiwan's Academy: An Outcry for Fair Play." *Asia Pacific Education Review,* vol. 14, no. 1, 2013, pp. 23–31.
[25] Universität Innsbruck, Bundesministerium für Wissenschaft und Forschung, *Leistungsvereinbarung 2013–2015* [online]. N.d. Available at: http://www.bmwf.gv.at/uploads/tx_contentbox/Universitaet_Innsbruck_LV_2013-2015.pdf [cit. 18. 9. 2013].

journals by SCI, SSCI, AHCI, meaning the fact that a journal is chosen by Thomson Reuters to be included in its source pool, is *per se* a grant for high quality, due to TR's "exceptionally rigorous selection standards".[26]

(2) Spain found more radical way for rewarding science performance and pays bonuses to individual researcher for research reports in journals with a high impact factor.[27] China, the Philippines and other countries of the so-called Third World pay *financial bonuses* to the authors of JIF-publications / of publications with high impacts (citations). The United Kingdom's *Research Assessment Exercise (RAE)* evaluates the higher education institutions research output and impact. RAE's results determine not only the budgets of institutions, but also the national research priority areas.[28] The upcoming RAE Framework (REF) for 2014 is "to be used from 2015-2016 to selectively to allocate research funding".[29] Richard Naftalin, Emeritus Professor of Physiology, alleges that high article impacts and Journal Impact Factors would be pre-requisitions for institution funding: In elite institutions only papers published in journals with an impact factor of 5 or greater will be submitted for assessment by REF. Papers graded by REF as 'outstanding' will earn their institution £ 100,000 (~ $ 154,000), those rated merely "excellent" will be awarded £ 25,000 (~ $ 38,000), anything less will be given no funding.[30]

It is important to stress that REF's official policy is to measure institutions' research impact by using SCOPUS citation data. Article impacts and Journal Impact Factors are not correlating: a few "hot papers" can raise the JIF manifold. Therefore REF's decision should be valued positively. An anonymous referee pointed out that it is not REF's official policy. Following Naftalin the point is: The universities are obliged to submit the "best" papers of their production for assessment by REF. Many institutions use TR's JIF to select these "best" papers: "To be included as an active research scientist

[26] Thomson Reuters, *Web of Science Coverage Expansion* [online]. 2010. Available at: <http://community.thomsonreuters.com/t5/Citation-Impact-Center/Web-of-Science-Coverage-Expansion/ba-p/10663> [cit. 13. 5. 2013].
[27] Evaristo JIMÉNEZ-CONTRERAS *et al.*, "Impact-factor Rewards Affect Spanish Research." *Nature*, vol. 417, 2002, p. 898.
[28] Keith HOGGART, "Assessing Research, Diluting Outputs, Confusing Institutions and Bedazzling Disciplines." *Progress in Human Geography,* vol. 30, 2006, no. 1, pp. 769–774.
[29] Richard NAFTALIN, "Opinion: Rethinking Scientific Evaluation." *The Scientist* [online], July 16, 2013. Available at: <http://www.the-scientist.com/?articles.view/articleNo/36291/title/Opinion--Rethinking-Scientific-Evaluation/> [cit. 16. 10. 2013].
[30] *Ibid*.

in an elite university's submission to REF requires three recently published papers in journals with high impact factors."[31] The selection of the so-called "best papers" based on JIF is just an example that scientists and scientific organisations are not only "victims". They take numerous shady actions not demanded by the evaluation agencies.

Back to the topic: the financial rewarding of high impact articles is a violation of Robert K. Merton's norm *disinterestedness*, an institutional imperative of Merton's *scientific ethos*:[32] scientists shall strive for knowledge accumulation, not for financial gain. If scientists strive primarily for high impact articles they are in danger to choose topics which are not based on scientific importance, but target strategically sensationalism.

(3) A vast literature criticises the evident *geographical and language bias* in TR's coverage of indexed journals – "the majority of journals are Anglo-American, reflecting and favouring the UK- and US-based ideas, theories and literature published in one language namely English".[33] Nevertheless the science managers remain devoted to private corporation Thomson Reuters' commercial data.

(4) The global dominance of citation indexing and their products (i.e. citation counts and journal impact factors) have devastating consequences mainly for *social sciences and humanities*: still their publication languages are national, but national language publications get fewer citations and are less valued in evaluations; there is a strong pressure to conduct research on international mainstream issues, instead of urgent local-regional context issues; scholarly books, still the dominant publication form in social sciences and humanities, are devalued and downgraded compared with journal articles; single authorships are more frequent in social science and especially in the humanities, therefore downgrading the scientific output in many evaluations: The current use of citation-based metrics to evaluate the research output of individual researchers is highly discriminatory because they are uniformly applied to authors of single-author articles as well as contributors of multi-author papers.[34]

[31] *Ibid.*
[32] MERTON, "Normative Structure."
[33] Manuel B. AALBERS, "Creative Destruction through the Anglo-American Hegemony: A Non-Anglo-American View on Publications, Referees and Language." *Area,* vol. 36, 2004, no. 3, pp. 319–322.
[34] Jozsef KOVACS, "Honorary Authorship Epidemic in Scholarly Publications? How the Current Use of Citation-based Evaluative Metrics Make (Pseudo)Honorary Authors from

The most important point of criticism is the *strong reactivity of public measures*:[35] Output and Impact "measuring" are reactive methods; they exert normative power and they massively influence the decisions of scientists and their institutions. I can only repeat: their guideline is not the scientific ethos (Merton)[36] and growth of scientific knowledge, but only the production of papers in journals indexed by Thomson Reuters, as many as possible, with a journal impact factor as high as possible. Scientific misconduct is spreading, more and more papers have to be retracted. Journals with higher JIF show a higher retraction rate, too.[37]

The results of all these university rankings is not only of academic interest, but the public opinion is highly affected: The national and international sensation-seeking mass media spread the rankings' results and present them as international tournaments of national academic institutions. It is important to emphasize that almost all large university rankings are products of media or media corporations. To say it with Pierre Bourdieu: Mass media exert "intrusion effects"[38] on the scientific field. I claim: they subordinate scientific achievements under their logic of sports competition ("higher, faster, stronger").

Most media reports do not mention nor discuss methodologies and data quality of these rankings. Because of the strong influence of TR data, it is more than necessary to examine their quality.

3.3 Specific Evaluation Effects: Trivial Errors in Thomson Reuters' Data and their Effects

As already mentioned only a few scientometricians or information scientists bear a critical attitude towards Thomson Reuters. To speak of "trivial errors" – trivial in the sense of marginal, insignificant, negligible – can be understood as a euphemistic strategy, as the following two severe problems in TR data computations show. Both authors are "outsiders" and no members of the hard-core of the scientometrics community:

Honest Contributors of Every Multi-Author Article." *Journal of Medical Ethics*, vol. 39, 2013, no. 8, pp. 509–512.

[35] Gerhard FRÖHLICH, "Das Messen des leicht Meßbaren. Output-Indikatoren, Impact-Maße: Artefakte der Szientometrie?" *GMD (Gesellschaft für Mathematik und Datenverarbeitung) Report,* vol. 61, 1999, pp. 27–38; ESPELAND – SAUDER, "Rankings and Reactivity."

[36] MERTON, "Normative Structure."

[37] Ferric C. FANG *et al.*, "Retracted Science and the Retraction Index." *Infection and Immunity*, vol. 79, 2011, no. 10, pp. 3855–3859.

[38] Pierre BOURDIEU, *On Television and Journalism*. London: Pluto 1999.

(1) Anne-Wil Harzing (2013)[39] attacks the *massive false categorization* of articles by the category of "document type" by Thomson Reuters' indexing procedures. According to Harzing, "articles" (i.e. original research reports) were massively falsely classified as "reviews" or as "conferences reports".[40] Thomson Reuters defines every article containing more than 100 references as a "review", as well as every article containing an acknowledgement in the footnote like "the author is thankful for critical discussions with the participants of the workshop XXY" as a "conference report". Why is this categorization discriminating social sciences publications? Social sciences are often text-based. In contrast to natural sciences articles it is common for a social science publication to have a large number of references.

In natural and engineering sciences it is usual to publish conference proceedings prior to their oral presentations. In social sciences symposia presentations are in form of a first draft (mostly only in form of PPT slides). The final version of the eventually published article is highly elaborated and has only marginal similarity to the original presentation.

Thomson Reuters gives no explanation why documents containing more than 100 references are automatically categorised as "reviews" – even if they are original research articles. There is no explanation, why articles – not published in conference proceedings – are classified as "conference reports", too.

Both erroneous document type categorizations have strong evaluation effects: Shanghai University Ranking counts only publications classified, as "articles" in the TR owned citation indexes. Hence all falsely classified articles lead to miscalculation of publication output and impact, meaning heavy losses in terms of number of publications and number of citations for social sciences, universities focusing on social sciences and for individual social scientists.

[39] Anne-Wil Harzing is a critical Australian management scientist who has developed with colleagues the free software *Publish or Perish* which uses the free Google Scholar citation search engine for scientometric studies and rankings. Google Scholar's data quality has been massively attacked in the literature, especially by Péter Jacsó (see footnote no. 49). Over a decade Harzing has published several higly critical studies on Thomson Reuters' data.

[40] Similarly Mike ROSSNER *et al*. ("Show Me the Data." *The Journal of Cell Biology*, vol. 179, 2007, no. 6, pp. 1091–1092) bought and examined the data for several medical and biological journals: "there were numerous incorrect article-type designations. [...] This was true for all the journals we examined."

(2) *Errors in/confusion of journal titles / journal title abbreviations* are a massive problem, because they influence the Journal Impact Factors. The critical study by Lange[41] shows the strong effects of database errors for the two educational science journals *Educational Research* and *Educational Researcher*. The former journal is classified as source journal by Social Science Citation Index and therefore its journal impact factor is calculated. The latter journal is not indexed in the SSCI, therefore its JIF is not calculated. Lange[42] found out that Educational Researcher is suspiciously often cited. The author assumed that the published JIF for Educational Research was based on *erroneous citation counts* in SSCI: due to similar journal title abbreviations ALL citations of the two journals were assigned only to one journal, namely Educational Research. Thomson Reuters were informed already in 1996 about this assumption. This allusion led to the sharp decline of the JIF for Educational Research in 1997 – *from 4.333 to 0.043 (!)*. That means: for almost two decades Educational Research had had a *hundredfold* (!) *exaggerated impact factor*. Thomson Reuters made neither official retraction nor public error correction. To have published articles in a journal characterised by a hundredfold exaggerate JIF is a "godsend" for researchers and their editors – leading to better positions, more citations, higher amounts of grants, media visibility *en masse*. The evaluation losers have been the authors and editors of the second journal Educational Researcher.

3.4 Trivial errors in SSCI: The Case of Pierre Bourdieu

My own first case study focuses on the *author name mutants* of *Pierre Bourdieu* in the Social Sciences Citation Index (SSCI). This famous French philosopher and social scientist was chosen because he is one of the most cited scholars of the 20th century; Bourdieu is an ASCII (American Standard Code for Information Interchange) friendly-name – his surname and given name should be no problem for TR data processing; but Bourdieu is a non-Anglo-American author and editor with world-wide diffusion,[43] many of his papers or papers citing him are in French or German and other non-English languages; the complete works by Pierre Bourdieu inclusive all translations

[41] Lydia L. LANGE, "The Impact Factor as a Phantom: Is There a Self-fulfilling Prophecy Effect of Impact?" *Journal of Documentation*, vol. 58, 2002, no. 2, pp. 175–184.
[42] *Ibid*., p. 177f.
[43] Gerhard FRÖHLICH, "Die globale Diffusion Bourdieus." In: FRÖHLICH, G. - REHBEIN, B. (eds.): *Bourdieu-Handbuch. Leben – Werk – Wirkung*. Stuttgart: Metzler-Verlag, 2009, pp. 376–381.

and reprints are documented in the *HyperBourdieu ©WorldCatalogue*^{HTM}.[44] Why is this of importance? To identify name errors/name mutants in the SSCI is a cumbersome undertaking. It affords systematic knowledge of the authors' complete works, including reprinted and translated versions.

My search strategies and work flow: first I searched for Bourdieu as "cited author"; then I searched for Bourdieu's (most) famous "cited works"; subsequently I compared both lists for Bourdieu's name mutations and cross-checked the data. Till now, I have detected *more than eighty* mutated name variants for *Pierre Bourdieu* in the SSCI only (I have found additional mutants in SCI and AHCI). Due to limited space I will not provide the full list, but only my typology of the found mutations and mutilations (table 2).

Type 1: Surname correct (Bourdieu), given name initial incorrect or missing, e.g. *Bourdieu* (*AD; BP; GPV; JJH; KP; RP; TPR); Bourdieu 248; Bourdieu's*
Type 2: Surname incorrect, e.g. Bordieu, given name initial correct, e.g. *Bourdieux P; Bourdikeu P; Bourrdieu P; Broudieu P*
Type 3: Both surname and given name initial incorrect, e.g. *Bourdieum m*; Boudieu JJH*
Type 4: Fatal mutations / mutilations, e.g. *ourdieu p*; I3ourdieu, (P.); Bour; Pierre B; Pierri B*
Type 5: Author surname Bourdieu hidden or lost, e.g. *anonymous*; *ibid.;* an *empty space* instead of author surname
Type 6*:* Words from different references are lumped together to a new *phantom reference*, e.g. *Atkinson R; *BP*

Table 2: *Name Mutants / Mutilated Names of Pierre Bourdieu in SSCI: Own Typology*
Source: compilation; italics indicate mutations / mutilations in SSCI records.

The typology enlisted in Table 2 needs some exemplifications:

(1) Errors of *Type 1* (surname correct, given name initial incorrect or missing) could be classified as "minor errors". But it is important to stress that in the world-wide community of science there are many, often hundreds of scientists and scholars with the same surname. In Asian countries like Korea

[44] Ingo MÖRTH – Gerhard FRÖHLICH, *HyperBourdieu© WorldCatalogue*^{HTM} [online]. 1999ff. Available at: <http://hyperbourdieu.jku.at/> [cit. 6. 9. 2013].

or China most of people share some few surnames: "The Chinese Academy of Sciences collected 4100 surnames [...] The top 129 surnames are shared by 87 per cent of the Chinese population."[45] The surname *Kim* is the most common Korean family name; over centuries roughly 1/5 of females have born the family name Kim.[46]

Therefore it is of utmost importance for database searches and for quantitative evaluations of individual researchers (such as h-index calculations) to know precisely their complete correct name in order to correctly identify their publications. To distinguish scientists or scholars with the same surnames and first given names we have to know also the *correct middle names*.

(2) Errors of *Type 2* (surname incorrect, given name initial correct) can be grouped into three subtypes: incorrect surname due to *letter commission* (e.g. Bourdieux; Bourdieru; Bourdicu); incorrect surname due to *letter omission* (e.g. Burdieu; Bourdiu; Boudieu; Bourieu); incorrect surname due to *misspelling* or *letter commutation* (e.g. Bourdeui; Borudieu; Bouridue; Broudieu).

(3) Errors of Type 3 (both surname and given name initial incorrect) such as *Boudieu JJH* fall through all cracks (search strategies, individual impact counting). They inevitably result in undervalued h-indices. The same effects are to be expected for the errors type 4, 5, 6.

(4) Errors of Type 4 (e.g. *ourdieu p*; I3ourdieu, (P.); Bour; Pierre B; Pierri B*) I call *fatal mutants*. These severe errors are clearly OCR errors (*I3ourdieu, (P.)*) or parsing errors (*Pierre B, Pierri B).* Such inadmissible errors could easily be detected by any serious quality control, be it automated or by human beings.

(5) Missings (errors of Type 5, author surname Bourdieu is hidden or lost, e.g. *anonymous* or *ibid*. or there is only an *empty space* instead of the author surname) are either human indexing errors or parsing errors. It is usual in juridical, social and cultural sciences to use footnotes and to use common abbreviations indicating repeated references to the same item such as *ibid*.[47] These short citation signals are often misinterpreted by Thomson Reuters'

[45] Liqun DAI, "Chinese Personal Names." *Centrepiece to the Indexer*, vol. 25, 2006, no. 2, pp. C1–C8.
[46] Seung Ki BAEK – Peter MINNHAGEN – Beom Jun KIM, "The Ten Thousand Kims." *New Journal of Physics,* vol. 13, 2011, no. 7, pp. 1–12.
[47] Or "derselbe" / "dieselben", shortform "ders."/"dies" in German language.

automatic indexing procedures, generating *phantom author surnames*.[48] All till now examined anonymous and ibid-type-SSCI-records have shown the same pattern: The original paper contained no errors.

(6) Errors of Type 6 are *phantom references*, e.g. *Atkinson R 1984 Distinction*. They result from *lumping together* fragments from different references (often from the same footnote or reference list, but sometimes also from diverse footnotes). A look into the original citing paper[49] shows: Its bibliography contains three references with the author surname "Atkinson R". Two positions later we find the correct reference for the English edition of Bourdieu's opus magnum *La Distinction*.[50] SSCI indexing has erroneously lumped together author surname and initial of Atkinson's publications and title abbreviation as well as publication year of Bourdieu's book. This severe database error is either an effect of human made indexing error or software, namely parsing error.

I can only repeat: the not detected trivial errors in author surnames and given names and/or initials (or their missing) have adverse effects on database searches and on evaluation. Misspelled or mutilated authors and their publications are not correctly archived in the citation databases. Therefore they are not counted by common citation analyses, resulting in undervalued h-indices.

As mentioned, scientometrics and error researchers blame the authors for making errors in database indexed publications. Contrary to this mainstream opinion, summarising the findings of my doctoral thesis' extensive quantitative case studies on SSCI errors, I claim the opposite generalizations: There are many severe errors in SSCI records**.** The cumbersome comparison between hundreds of detected cited reference records errors in SSCI with the original article's reference list showed almost every time the same result: *the original reference list was error free.* Therefore these detected errors I call *endogenous* database errors. They must be software (OCR, parsing) errors and/or human indexers' errors, indicating a severe deficit in Thomson Reuters' data quality control.

[48] Péter JACSÓ, "Deflated, Inflated and Phantom Citation Counts." *Online Information Review*, vol. 30, 2006, no. 3, pp. 297–309.
[49] Gary BRIDGE, "Perspectives on Cultural Capital and the Neighbourhood." *Urban Studies*, vol. 43, 2006, no. 4, p. 729 (719–730).
[50] Pierre BOURDIEU, *Distinction: A Social Critique of the Judgement of Taste.* London: Routledge 1984.

It was more than hard for me to gain a precise description of TR's work flows and procedures. But Garfield's publications and utterances were more informative. Therefore I decided to take a historical approach.[51]

4. Strategies and Contingencies in the Genesis of Science Citation Indexing

I want to illuminate the genesis of commercial citation indexing for science by interpreting the "*Joshua Lederberg Papers*" (provided by the National Library of Medicine).[52] First, why the utilisation of the papers (letters, notes, materials) of the geneticist *Joshua Lederberg* at all? Eugene Garfield tried to start the citation indexing project already in the mid-1950s – but without success. I assert that Lederberg's social and symbolic capital as 1958 Nobel Prize laureate in Physiology or Medicine as well as his expertise in scientific communication was indispensable for the realisation of citation indexing. Last but not least Lederberg coined the term "*Science Citation Index*". It is important to note that Lederberg's collection of letters and materials seem to be less unselected and more comprehensive than the materials posted on Garfield's homepage.

4.1 The Strategy: Spreading over-optimism, downplaying severe problems

To overcome severe resistance (lack of interest, severe criticism by scientists and by anonymous grant application referees) Eugene Garfield, the "driving force" of the citation indexing project, had to foster overoptimistic attitudes and to downplay the severe problems of global and multidisciplinary citation indexing:

> From this description it will be apparent that, although a great volume of material is to be covered, relatively *unskilled persons* can perform the necessary coding and filing. Professional supervision would still be required, because certain decisions require skilled judgement, for example when ibid. or loc.cit must be carefully interpreted. Footnotes tend to make coding somewhat cumbersome.[53]

[51] I am grateful for Prof. Ingo Mörth for this suggestion.
[52] National Library of Medicine, *Profiles in Science, The Joshua Lederberg Papers* [online]. N.d. Available at: <http://profiles.nlm.nih.gov/BB/> [cit. 16. 10. 2013]. All subsequent cited letters are documented in this archive.
[53] Eugene GARFIELD, "Citation Indexes for Science." *Science,* vol. 122, 1955, no. 3159, p. 111 (108–111).

Garfield and even Lederberg were convinced that *foreign language fluency* is an unnecessary qualification, as Garfield wrote to Lederberg "Russian doesn't really bother me as you can train a girl to transliterate in about one hour."[54]

Numerous letters addressed one topic: *Money*. Several research fund proposals by Garfield were rejected. Eugene Garfield's frustration is best expressed by his wording: "Needless to say, the proposal was turned down."[55]

Lederberg gave Garfield twofold strategic advise. The first one was to downplay the man power cost by pursuing the *automation idea*:

> But for the costs: the job would need mainly money and machines, not professional manpower. It can be conveniently decentralised – even in some places to the point of publication. One way to illustrate its mechanical advantages is to point out that a staff could even index papers in foreign languages *without understanding the text*, just provided they can read the reference lists onto the citation cards. In any case, for a world-wide scheme, a lot of work could be done abroad especially, but not necessarily exclusively for publications in languages other than English. From what I learned of the relative costs of a punch card operator in Italy vs. California, you might well want to farm out a fair part of the work.[56]

In order to get money from the National Institutes of Health (NIH), Joshua Lederberg suggested Garfield twice to propose the citation indexes *as evaluative tool*:

> The NIH administration was interested in making to evaluate the actual impact of NIH support for biological and medical research in this country. The NIH administration was considering a number of rather fancy and insufficient

[54] Garfield to Lederberg May 21, 1959. An anonymous referee qualified this quotation as "rather controversial […] (expressions such as 'train a girl' might be problematic from gender perspective)". The referee might have overseen that the expression she/he found fault with is the original notation. One of Garfield's letters contains pejorative formulations concerning women, which would be qualified nowadays as "problematic from gender perspective". See Garfield to Lederberg, June 23, 1959: "You can't imagine how frustrating it has been in the past five year (or maybe you can) to have had at the helm of scientific documentation activities in NSF *a woman who was neither a scientist or an information specialist, but just a good secretary (a Spanish major)* who *worked her way up by taking good notes at meetings and preparing reports for her bosses*. I would never say this publicly, but that is the absolute truth." (Italics by TTF; NSF is the acronym for National Science Foundation, USA)

[55] *Ibid*.

[56] Lederberg to Garfield, June 18, 1959. (italics added, TTF)

> schemes for doing this. It should take little imagination to see how SCI could accomplish the purpose at a negligible additional cost. In the first place the type of acknowledged support with more or less detail could be one of the keys in the index. Also the impact of NIH supported work could be measured in terms of the frequency of citation to it. Quite seriously with so many agencies anxious to know just what their real effect is, a quantitative measure such as SCI would very readily furnish would be a very valuable tool for them.[57]

There is a widespread myth in the scientometric community, namely that evaluation was not an intended purpose of the fathers of citation indexing. As demonstrated above, that is not the truth.

Eventually Eugen Garfield gave up the idea to get funding for exhaustive citation indexing research: „My conclusion is that nobody wants to do research on this anymore – they just want me to plow into making a citation index."[58] Lederberg arranged as highly reputated geneticist a meeting with the Genetics Study Section of the Institutes of Health (NIH). Finally they got a grant to produce a *Genetics Citation Index*. A year later, Garfield expanded his GCI to the *Science Citation Index***.**

4.2 First Error Reports: "More a Comedy of Errors Than a Real Loss"

Soon after the first citation index specimen sheets were sent out, Garfield was notified of the trivial errors the volumes contained. The following heavy complaint from *J. B. S. Haldane*, to Eugene Garfield / ISI in the year 1963 is found only amongst the Lederberg papers:

> Your specimen sheets are one of the most appalling productions that I have ever seen. I find following surnames: Wilha / hand written correction to "-li" (for Williams), Mit (for Smith), Haldan, Thomps (for Thompson), Spearn (for spear), Falcon (for falconer) Etc. (Commas added by TTF) Many of these errors are repeated. When I get a similar production from an Indian source I do not hesitate to say that it reflects discredit on India and should not be sent abroad. In your case the international distribution of your citation index will be of great value to those who state not without some evidence, that the standard of scientific publication in the US is rapidly deteriorating.[59]

Joshua Lederberg's reply to Haldane was scarce and ambiguous:

[57] Lederberg to Garfield, July 29, 1960.
[58] Garfield to Lederberg, May 21, 1959.
[59] J. B. S. Haldane to Eugene Garfield/ISI, May 18, 1963.

> I am sorry about the misprints that plague the computer outputs. It is a serious problem, not uniquely American. Dr.EG will surely respond directly. If he spent less time in salesmanship, there would be no ISI at all: perhaps that would be preferable by your own reckoning."[60]

There is no answer of Garfield to Haldane documented. But a memo ten years later by Lederberg to Garfield still downplays the fatal errors in author name indexing, using the issue of *Chinese names*: "I just ran into a problem in a way that is *more a comedy of errors than a real loss*, except a few minutes time. SHEN is cited by several authors, but you'd never find it under Shen, he is indexed as CHIUNG."[61]

In other words: The difficulties to handle different formats of references and footnotes, non-Anglo-American names, and of publications in non-English languages *were known to the pioneers* of citation indexing, but they dismissed them.

4.3 Contingencies: The Emergence on US Soil, as Genetics Index, as a Child of the Punch-card Era

Archambault – Larivière consider the geographical contingency of the development of citation indexing and of the journal impact factor:

> The emergence and evolution of this method on US soil [...] likely they had the effect of creating a self-fulfilling prophecy. Indeed, concentrating on the US situation and by positively biasing the sources in favour of US journals, the method placed these journals on centre stage. Had a broader linguistic and national coverage been considered, it might have revealed that these journals were not in fact more cited than others. By creating this centre stage, the measures of JIF made a *selective promotion of US journals*, which could then be picked up, read, and increasingly cited by researchers in the US and also abroad.[62]

Archambault – Larivière conclude:

> Had the Institute for Scientific Information (ISI) emerged as the "Institut für Forschungsinformation", the JCR would undoubtedly have evolved in a sub-

[60] Lederberg to Haldane, May 31, 1963.
[61] Memo Lederberg to Garfield, March 25, 1973; italics added by TTF.
[62] Éric ARCHAMBAULT – Vincent LARIVIÈRE, "History of the Journal Impact Factor: Contingencies and Consequences." *Scientometrics*, vol. 79, 2009, no. 3, p. 4 (1–15); italics added by TTF.

> stantially different form and the aggregate current impact of German journals would likely to be substantially larger.[63]

I agree with Archambault – Larivière , but want to add additional contingencies:

(1) Had the first Citation Index for Science emerged not as *Genetics* Citation Index, but as a "*Sociology* Citation Index" or a "*Philosophy* Citation Index", more effort in detail would have been exercised for indexing surnames and publication titles. In genetics it has been usual to abstain from mentioning the full given names – even in the *author line* of the original paper. In genetics' *reference lists* it has been usual to abstain from listing full given names and even the publication titles. Therefore I think ISI and its successors have not been interested in and have not been not sensitised to guard against the confusion of surnames and given names and to consistently and error-free coverage of the publication titles.

Both shortcomings were connected with the prematurity of the citation indexing enterprise as an automated procedure: It was necessary to be stingy with each of the 80 columns on the punch card. *Citation indexing is a child of the punch card era.*

(2) Concerning the vexed problem of *getting funded*: Had the Armed Forces or NASA believed in citation indexing as at tool for supreme power respectively to advantages in the race to the Moon, they would have paid a plenty of money to Garfield; had Garfield initiated his citation indexing project in the times of *neoliberal "audit cultures"*,[64] foundations and governments would had paid a plenty of money to Garfield.

My thesis goes as follows: The chronicle shortness of money, the severe limitations of hardware and software in the early days of citation indexing as well as the limited disciplinary provenances of the leading actors lead to the strategy to downplay or even ignore the severe error and disambiguation problems of citation indexing.

5. Conclusion: The Inertia of Commercial Citation Indexing and the Demands of DORA

Eugene Garfield was an ardent innovator; he was obsessed with the idea of citation indexing for science. He had to start without forgoing extensive

[63] *Ibid.*

[64] Marilyn STRATHERN (ed.), *Audit Cultures. Anthropological Studies in Accountability, Ethics and the Academy.* London – New York: Routledge 2000.

research; he was forced to find *low-cost ad-lib solutions* (unqualified cheap labour and automated procedures). Garfield had to establish himself as a "scientific-documentary entrepreneur", an unknown role at that time. The banks turned Garfield down, so he had to borrow expensive money from the Household Finance Corporation to survive. *His persistence is admirable.* He had to take enormous financial risks. Therefore the "error-making" version of automatic and cheap-labour citation indexing was maybe the only way to gain momentum in the 1960s.

But nowadays the huge and rich North-American media corporation Thomson Reuters (TR) is the owner of the citation data banks founded by Garfield. Thomson Reuters would have the financial capacities to search and correct the errors and to re-launch their databases. But still there is only patchwork: new data fields, features and services are added, escalating the inconsistencies and errors. TR's strategy is to maintain market dominance and to launch new business areas. No fundamental reforms are in sight. Huge technological systems show a heavy inertness. This insight of technology studies is applicable to the large citation indexes by Thomson Reuters, too. But this inertia is inextricably connected with the profit motive of commercial indexing. As said by Péter Jacsó:

> Many librarians are very vocal in criticizing free Web databases for their deficiencies. They are right to do so, but they should know that respected traditional information providers from ritzy corporate headquarters often deliver far more deficient databases for nifty fees. Compiling databases of accurate information costs a lot of money that few content providers are willing to pay.[65]

To conclude I would like to remind my starting point, referring to Sir Karl Popper: He criticises the "old attitude" of "hiding of our mistakes and to forgetting them".[66] Popper thinks that *to detect, to (publicly) correct and to retract errors* is important for the progress of knowledge accumulation. As common code of practice in serious scientific journals, I would demand *corrigenda / official retractions* from citation database producers, too. My demand was qualified as "awkward" by one anonymous referee.

But since early days the citation database producers Institute for *Scientific* Information (ISI), then called Thomson *Scientific* (!), now called

[65] Péter JACSÓ, *Content Evaluation of Textual CD-ROM and Web Databases*. Englewood: Libraries Unlimited 2001, p. 169.
[66] McINTYRE – POPPER, "The Critical Attitude in Medicine." p. 1920.

Thomson Reuters, they all have raised scientific claims.[67] Apart from that apologies and corrigenda are by all means usual in the database business.[68] The previously mentioned international declaration DORA, *The San Francisco Declaration on Research Assessment,* has been signed by 547 scientific organisations and 12055 journal editors and scientists (reference date: 3. 11. 2014). DORA's essential demand is already formulated in the subtitle of the declaration: „Putting science into the assessment of research".[69]

DORA criticises that the "data used to calculate the Journal Impact Factors are neither transparent nor openly available to the public."[70] „For *organizations that supply metrics*" DORA recommends, among others, „Be open and transparent by providing data and methods used to calculate all metrics. Provide the data under a licence that allows unrestricted reuse, and provide computational access to data, where possible."[71]

Evidently, DORA's demands are guided by the scientific ethos, by Robert K. Merton's institutional imperatives of "communism, universalism, disinterestedness and organized scepticism". [72] Therefore I claim: the minimum quality standard for scientific transparency and verifiability for database publishers would be to provide corrections and retractions.

[67] Erik Jan van Kleef, Thomson Reuters, *Vice President of Sales EMEA* (Europe, the Middle East and Africa), at international conference ODOK 2012, Wels/Austria, Section *Wert des Wissenszugangs - Open Access II*, September 13, 2012, public discussion.
[68] *Correction Notice. Corrections to the Data Tables for the Canadian MIS Database Hospital Financial Performance Indicators, 2008–2009 to 2012–2013* [online]. 30. 8. 2014. Available at: <http://www.cihi.ca/web/resource/en/hfp_correction_notice_2014_en.pdf> [cit. 3. 11. 2014].
[69] *San Francisco Declaration on Research Assessment. Putting Science into the Assessment of Research* [online]. 2013ff. Available at: <http://am.ascb.org/dora/> [cit. 3. 11. 2014].
[70] *Ibid*.
[71] *Ibid*.
[72] MERTON, "Normative Structure of Science."