

Local, Networked or External?

Inclusion of Regional Journals in WoS and its Effect

Valeria Aman

iFQ – Institute for Research Information and Quality Assurance
Schützenstraße 6a, Berlin, Germany
aman@forschungsinfo.de

Abstract

Between 2007 and 2009 Thomson Reuters started to accept journals beyond the international high-impact literature, and included 1,600 regional journals in Web of Science (WoS). The study sets out to analyze the effect on countries that in terms of absolute article numbers profited most from the inclusion process. The comparison of a country's newly included journals with those in existence in WoS before 2007, and the distinction of articles according to their origin (local, networked or external) allows to detect various effects. The reception of these two sets of journal articles is likewise distinguished by the origin of the recipients. Results show that some countries suffered a loss in international visibility and are now represented in WoS more locally oriented than before the altered journal selection process. Other countries experienced a growth in external citations as well as publications from international authors, being everything but regional.

Keywords: Regional journals, Coverage, Productivity, Web of Science

In: F. Pehar/C. Schlögl/C. Wolff (Eds.). *Re:inventing Information Science in the Networked Society. Proceedings of the 14th International Symposium on Information Science (ISI 2015)*, Zadar, Croatia, 19th–21st May 2015. Glückstadt: Verlag Werner Hülsbusch, pp. 358–369.

1 Introduction

What Gibbs (1995) laconically calls “Lost science in the Third World” refers to the locally published journal articles that are virtually invisible for the international scientific community. The low visibility derives from the fact that the majority of their publications are disseminated through domestic scholarly journals or other publication means with a geographically and linguistically limited range of dissemination (Arunachalam, 2004). Since the inception of the *Science Citation Index* (SCI) by Eugene Garfield more than half a century ago, the goal of the journal selection process was to identify journals that build the core literature of the sciences. These were mostly English-language journals meeting the need of a broad research community. In 1973, journals from the Social Sciences were added to SCI, followed by Arts & Humanities in 1978. Journals included in Web of Science were regarded as the ones with the highest citation impact of all journals published. Over a period of more than 40 years, the set of international highly-cited journals was relatively constant. Between 2007 and 2009 Thomson Reuters changed its journal policy and included around 1,600 so called “regional journals”. These journals present research from a regional perspective, while their international impact is secondary. The aim of this paper is to analyze the effect on those countries that profited most from the inclusion process that started in 2007. The distribution of locally publishing, networked and external authors and how it changed with the inclusion of regional journals will be addressed. Likewise, the distribution of the origin of incoming citations, just as the citation counts, indicate the effect of the journal inclusion on the orientation of the countries, thus whether they are interconnected, locally or externally oriented.

2 Background

The community that uses Web of Science (WoS) in their daily work has expanded over the years. Whereas in the past universities and research institutions in the US and Western Europe formed the core, nowadays major universities and research facilities all over the world show broad interest in Web of Science. With the expanse of the circle of users, Thomson Reuters adapted

quickly a strategy to serve the need of the newcomers. They dissociated from the original journal selection process and moved into “the realm of the regional journal literature” (Testa, 2009). According to Testa (2009) regional journals are those published outside the US or UK. They focus on topics of regional interest or present them from a regional perspective. The authors are from the region rather than from international research groups. The importance of regional journals is measured by the content and not simply by their citation impact. Requirements for the inclusion of regional journals in WoS are the publication on time, a formal peer-review and bibliographic information in English (with the exception of Arts & Humanities). Testa calls it a “monumental task”, which was initiated by the Editorial Development Department of Thomson Reuters in 2006. Subject editors compiled a list of more than 10,000 journals from all areas of scholarship. Within 12 months they selected 700 journals for coverage in WoS. Another 900 were added by the end of 2009 (Testa, 2011). The number of regional journals covered by WoS will significantly influence any analysis of the country that is affected by the inclusion of regional journals. Basu (2010) examined the effects on a country’s scientific productivity, if the SCI-indexed journals covered keep changing with time. He concluded that an apparent increase of a country’s productivity may simply result from the inclusion of new journals in the database. There are several previous studies focusing on regional countries, their visibility and the influence of publication language on impact. Van Leeuwen et al. (2001) discussed the consequences of the relative language coverage of journals in the SCI, when comparing national research performance on a global scale. They concluded that the citation rate of a paper depends primarily on the language it was published, with an outstanding advantage of English-language papers. Tijssen, Mouton, van Leeuwen & Boshoff (2006) analyzed South African journals, including international journals indexed in bibliographic databases and local journals. The analysis of more than 200 South African journals in regard to output and citation impact showed that the majority of local journals are invisible for the global scientific community. Collazo-Reyes et al. (2008) studied the publication and citation patterns and growth dynamics of Latin American and Caribbean journals covered in WoS from 1995–2003. They found little inter-citation among local papers, while the highest cited papers by extra-regional authors were those published in English. Lermarchand (2010) determined the journal coverage of 12 Iberoamerican and Caribbean countries and analyzed the evolution of the cooperation networks among them in the 1973–2006 period. What

most of the previous studies have in common is that they take the presence of regional journals in WoS for granted, showing that regional journals are disadvantaged in terms of visibility and impact due to their publication language or outreach. Kosanović & Šipka (2013), and Collazo-Reyes (2014) studied critically the effect of Thomson Reuters change in journal policy. Their analyses of South East European journals and Latin American and Caribbean journals respectively, show that the recently covered journals in WoS are low in citation impact. They conclude that overrepresentation in WoS can neither be the long-term interest of the country affected, nor of Thomson Reuters.

3 Data and methods

The study was conducted by means of an in-house-database version of Thomson Reuters' Web of Science. The journal analyses include the *Science Citation Index* (SCI), *Social Sciences Citation Index* (SSCI) and the *Arts & Humanities Citation Index* (A&HCI). To prevent any bias among countries, the analyses are restricted to journals and articles only. In a first step, journals had to be assigned to the country they are published, and the year of their inclusion into WoS had to be identified. Based on these information, for each of the countries of interest a distinction between "new" journals, those that were by definition included after 2007 and "old" journals, those that have been included prior to 2007 in WoS, was possible. For both of these sets of journals (new and old) articles published between 2007 and 2013 were classified into one of the following three groups, according to the geopolitical location of the author's institution and the journal:

- (L) *Local articles*, where all of the authors affiliated to an institution are situated in the country where the publishing journal is edited.
- (N) *Networked articles*, where at least one of the authors is from the local country, the target journal is edited in and one author whose affiliation belongs to an institution situated in a country other than that where the publishing journal is edited.
- (E) *External articles*, where all of the publishing authors work at institutions located outside the country where the respective journal is edited.

These three sets are disjoint and allow comparisons with respect to the following indicators: Total numbers and percentages of articles published be-

tween 2007 and 2013 in a country, distinguished by the three groups of articles (L, N, E) and the set of journals (old vs. new). The same distinction has been applied for the citation analyses. Local citations are those from the country, where the article is published, networked citations originate from articles where a local author co-operated with an author from abroad. External citations are accumulated by publications where all of the authors are located in a country other than that where the cited article was published.

4 Results and discussion

The countries of interest have been chosen on the basis of the highest number of incoming journal articles, resulting from Thomson Reuters' inclusion of regional journals. Table 1 lists the 20 countries in a descendant order, according to their absolute number of articles published between 2007 and 2013 in newly included journals. On the right part of the table we can see the number of articles that were published in journals that have been already covered in WoS before 2007. A column indicating the number of "new" and "old" journals is provided next to each of the sets of articles.

We can infer from the table that the lion's share of articles in newly included journals are of US-origin, followed by Great Britain. Brazil is on rank 3, followed by Poland and Germany. According to Testa (2011) UK, US, Germany and the Netherlands faced the greatest increase in journal coverage between 2005 and 2010. We can see that Turkey's number of journals represented in WoS has grown by a factor of ten. The ratio of articles and journals results in the indicator "Journal Packing Density" (JPD), which is defined as the average number of papers in journals for a given country (see Basu, 2010). The example of Pakistan shows that the JPD is extremely high for the two journals included prior to 2007. On average, the ratio results in 280 articles per journal per year. In contrast, Pakistan's JPD for the recently included journals is much lower. To illustrate the countries that profited most from the journal inclusion, the following figure provides a bar chart. It visualizes the percentages of articles published between 2007 and 2013 in journals accepted after 2007, in relation to articles in journals that have been covered in WoS before 2007.

Table 1. Overview of the effect of the journal inclusion process on article numbers and journal numbers published between 2007 and 2013 in SCI-E, SSCI, and A&HCI for countries with the highest growth in absolute articles numbers.

Country	No. of articles in journals incl. after 2007	No. of journals incl. after 2007	No. of articles in journals incl. prior to 2007	No. of journals incl. prior to 2007
USA	256,814	766	3,287,377	4,251
GB	146,719	622	1,422,221	2,988
Brazil	44,290	98	30,414	35
Poland	31,349	84	35,581	68
Germany	31,264	141	275,294	510
China	31,037	67	94,650	68
Netherlands	30,968	141	730,690	701
South Korea	27,365	63	41,357	39
India	23,195	63	43,056	52
Turkey	22,244	63	2,620	6
Romania	21,046	44	8,442	11
Switzerland	18,776	73	171,143	184
Italy	17,310	74	36,972	126
Japan	14,995	49	101,822	145
Spain	13,022	79	21,106	65
Pakistan	12,278	12	3,916	2
Iran	11,139	44	1,371	8
France	9,545	52	95,897	208
Australia	8,437	33	13,762	58
Croatia	8,317	40	4,871	17

Depending on the number of journals that were already in presence in WoS before 2007, the percentages vary immensely among countries. Evidently, Turkey, Iran and Pakistan are the countries who owe their today's presence in the database to the inclusion of regional journals. More than 50% of all articles published in Romania, Croatia and Brazil between 2007 and 2013 arise from the recently included journals. Although the number of journals from USA and GB has grown in absolute terms, effects on the proportion of all articles published in these countries are rather weak.

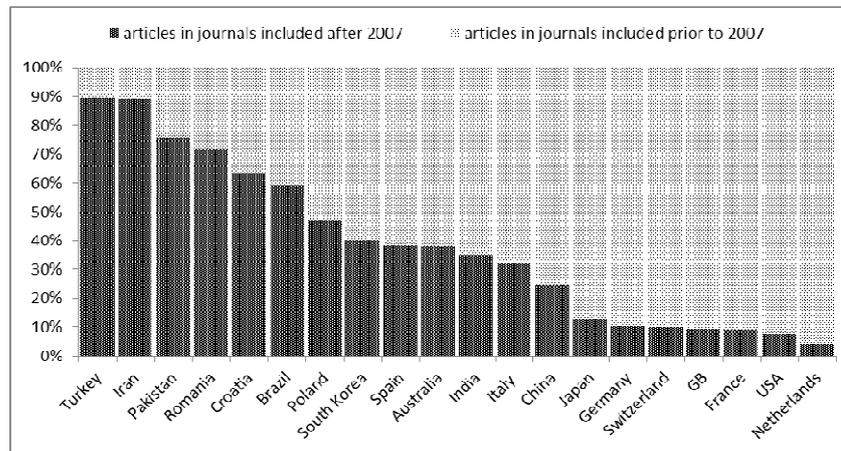


Figure 1. Countries that profited most from the journal inclusion process and their relation of article numbers published between 2007 and 2013, distinguished by articles in journals included after 2007, and those included prior to 2007.

To show the effects of the implemented journal selection process, it is now of interest to distinguish between “old” and “new” journals and the origin of the published article. Therefore, the following figure presents a bar chart for the 20 countries of interest. The left bar represents the “old” journals (included before 2007), whereas the right bar represents the set of newly included journals. Each of the bar consists of three groups of articles, according to their origin (L, N, E). The countries in figure 2 are not in an arbitrary order, instead they are arranged in accordance to the highest increase of the share of locally published articles. Starting with Pakistan, we can see that in the 2 journals already included in WoS before 2007, 70% of the articles published between 2007 and 2013, are exclusively by local authors. The right bar for Pakistan with the newly absorbed (“regional”) journals shows that these journals are not as regional as expected. Only 36% of articles are local, whereas 60% of articles are published exclusively by authors from abroad. Thus, the regional journals included are evidently of relevance for the international community, since many external authors place their articles in these journals.

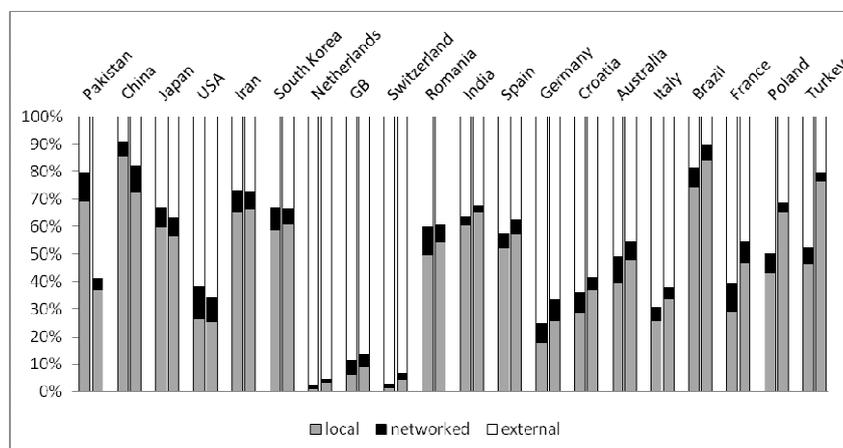


Figure 2. Distribution of the origin (L, N, E) of the articles published between 2007 and 2013. The left bar for each country represents articles in journals included prior to 2007, whereas the right bar represents the articles in journals included after 2007.

The same effect becomes visible for China. The journals included prior to 2007 show the highest rate of locally published articles (85%) in figure 2. The 67 Chinese journals included after 2007, show a rate of 72% of local articles. The recently included journals from Japan and USA also show a lower rate of locally published articles. For Iran and South Korea a small increase of locally published articles is visible, with around two-thirds of their articles being local. The Netherlands, GB and Switzerland show different characteristics. Together with UK and Germany they are centers of international scholarly publishing and experienced the greatest increase in journal coverage in WoS on a routine yearly basis (Testa, 2011). The journals published in these countries have a high influx of external publications. According to table 1 Romania, India and Spain faced an increase in regional journals. The bars in figure 2 show that the journals included after 2007 consist of more than 50% of local articles. A high number of regional journals was also included in Germany and Croatia. Just as China, Brazil is outstanding as a country with a high share of local articles – in absolute article numbers, and in relation to all articles published in the country. Different from China, the bars show that the newly included journals from Brazil are to a higher share local than those covered in WoS before 2007. The highest increase in the share of local articles becomes evident for Turkey. Whereas articles in “old” journals are to a rate of 50% external, the inclusion of regional journals let

this share sink to 20%. The number of networked papers has diminished as well. In terms of article counts, we can observe different effects on countries that were affected by the inclusion of 1,600 regional journals. It is now of interest to study the reception of articles published in these countries. A citation window of three years is applied. Thus, only those citations are counted, which were received in the year of article publication and the two following years. Figure 3 can be read as figure 2, with the difference that each bar represents the share of the origin (L, N, E) of the incoming citations. Whereas the two journals from Pakistan, included prior to 2007, received 35% of their citations from external authors, the newly included journals owe 68% of their citations to this group. Pakistan marks the highest increase in external citations from the countries displayed in figure 3. The results show that the recently absorbed journals are not as regional as stated in Testa's documentation (2011). Iran, Romania, India and China too, receive to a higher degree external citations in comparison to the journals that have been already integrated in WoS before 2007.

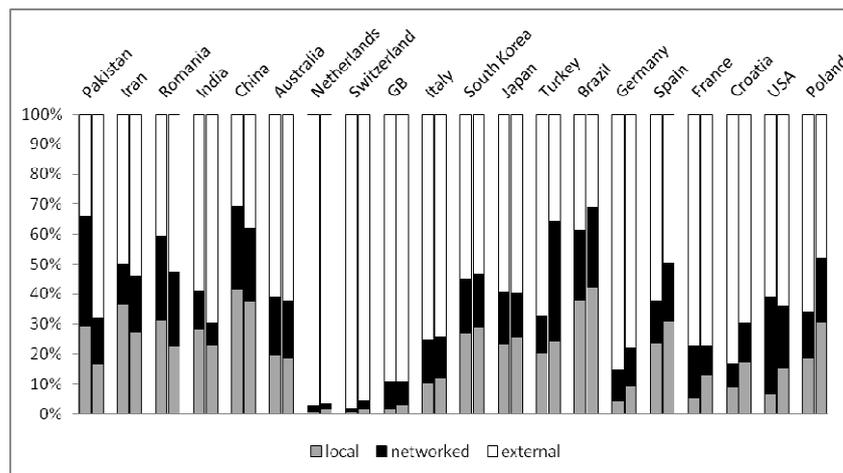


Figure 3. Distribution of the origin of the citations to articles published between 2007 and 2013 in the country as indicated. The left bar for each country represents articles in journals included prior to 2007, whereas the right bar represents articles in journals included after 2007.

The central countries of publishing, Netherlands, Switzerland and GB show high external citation rates. South Korea and Japan show with ca. 30% a stability in their share of national self-citations. Turkish journals that are

recently covered in WoS, are to a high degree cited by networked publications. One-third of all incoming citations results from Turkish authors, who have collaborated with authors from other countries than Turkey. The countries on the right of figure 3 show an increase in national self-citations in journals covered in WoS from 2007 on. These countries are accompanied by lower external citation rates. The increase of local citations in US should be interpreted with caution, because a size effect is at stake. Since the US-output is relatively large (table 1), US articles constitute a large citation “target” (see Moed 2005: 293). Having presented the origin of citations, we can now turn to the citation impact. The following figure indicates the citation rates to the group of articles distinguished (L, N, E). Pakistan on the left shows that more than 50% of all citations received in the 2007–2013 period point to the external publications in Pakistan’s journals. We can conclude that the recently included journals from Pakistan are not regional, since the majority of cited articles are those being external.

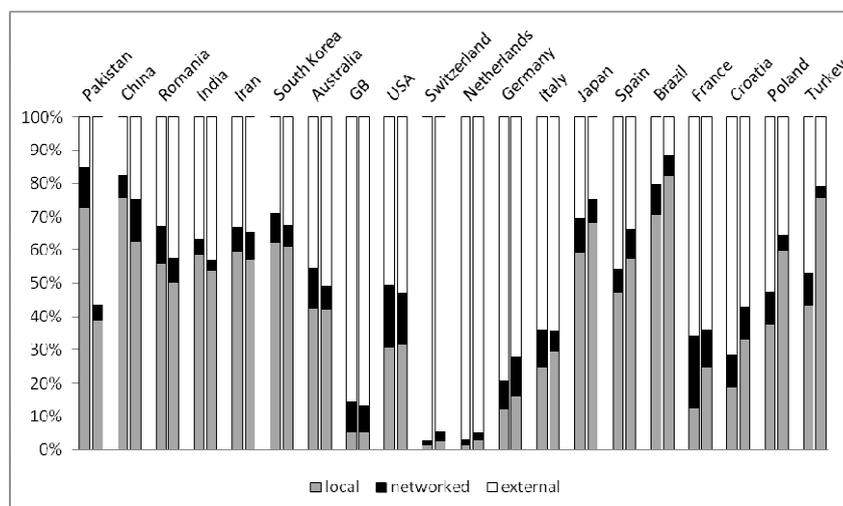


Figure 4. Distribution of the origin of a country’s articles that receive most of the citations between 2007 and 2013. The left bar for each country represents articles in journals included prior to 2007, whereas the right bar represents articles in journals included after 2007.

China, Romania, India, Iran and South Korea show that with the newly included journals a higher share of the external articles accounts for citations, than in the journals included prior to 2007. At the same time a lower share of

local articles accounts for incoming citations. This means, that these countries became less regional despite the inclusion of regional journals. On the opposite, we can see on the right the countries who faced an increase of citations to locally published articles. Local articles account for more than 80% of all citations to articles from Brazil. Articles of French origin and published by local authors account for higher citation rates at the expense of networked articles. Three quarter of all citations to Turkish articles in newly included journals, target to local articles. This is in accordance with the high share of local articles in Turkey (fig. 2).

5 Conclusion

Thomson Reuters extended the journal coverage by focusing on regional journals all over the world. The goal was to enrich the collection of international journals with those whose focus is on specific regional themes. Countries such as Turkey, Iran, Pakistan and Romania grew from very narrow representation to significant coverage. By the end of 2010, 87 countries were represented in Web of Science on the journal level, 14 of which for the first time (Testa, 2011). As Michels and Schmoch (2012) state, this growth in the articles number in WoS should not be interpreted as an increased scientific activity. It is primarily an artefact of the inclusion of regional journals since 2007, and the tolerance to accept journals regardless of their impact. The distinction by local, networked and external articles shows that some countries are now represented as more locally than before the journal expansion. The results of the citation impact indicate nevertheless that regional studies are referenced from external publications and do not go unnoticed in the international community of researchers. To conclude, the inclusion of regional journals since 2007 had different effects on the visibility and the reception of these journal articles for each of the countries presented.

References

- Arunachalam, S. (2004). Science on the periphery: bridging the information divide. In: H. Moed et al. (Eds.). *Handbook of Quantitative Science and Technology Research*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Basu, A. (2010). Does a country's scientific 'productivity' depend critically on the number of country journals indexed? *Scientometrics*, 82 (3), 507–516.
- Collazo-Reyes, F. (2014). Growth of the number of indexed journals of Latin America and the Caribbean: the effect on the impact of each country. *Scientometrics*, 98, 197–209.
- Collazo-Reyes, F. et al. (2008). Publication and citation patterns of Latin American & Caribbean journals in the SCI and SSCI from 1995 to 2004. *Scientometrics*, 75 (1), 145–161.
- Gibbs, W. W. (1995). Lost science in the Third World. *Scientific American*, 273, 76–83.
- Glänzel, W., Leta, J. & Thijs, B. (2006). Science in Brazil. Part 1: A macro-level comparative study. *Scientometrics*, 67 (1), 67–86.
- Kosanović, B., & Šipka, P. (2013). Output in WoS vs. Representation in JCR of SEE Nations: Does Mother Thomson Cherish All Her Children Equally? *Centre for Evaluation in Education and Science*, 125–137.
- Lemarchand, G. A. (2010). The long-term dynamics of co-authorship scientific networks, Iberoamerican Countries (1973–2006). <http://arxiv.org/abs/1001.2837>.
- Michels, C. & Schmoch, U. (2012). The growth of science and data bases coverage. *Scientometrics*, 93, 831–846.
- Moed, H. F. (2005). *Citation analysis in research evaluation*. Springer, Berlin.
- Testa, J. (2008). A regional perspective on science. *KnowledgeLink newsletter*, July 2008.
- Testa, J. (2009). Regional content expansion in Web of Science®: opening borders to exploration. GlobalHigherEd.
- Testa, J. (2011). The globalization of Web of Science 2005–2010. <http://wokinfo.com/media/pdf/globalwos-essay.pdf>.
- Tijssen, R. J. W., Mouton, J., van Leeuwen, T. N. & Boshoff, N. (2006). How relevant are local scholarly journals in global science? A case study of South Africa. *Research Evaluation*, 15 (3), 163–174.
- Van Leeuwen, T. N., Moed, H., Tijssen, R. J. W., Visser, M.S., Van Raan, A. F. J. (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics*, 51, 335–346.

Bibliometric Analysis of the Field of Folksonomy Research

Tomislav Ivanjko, Sonja Špiranec

University of Zagreb, Faculty of Humanities and Social Sciences, Department
of Information and Communication Sciences
Ivana Lučića 3, 10000 Zagreb, Croatia
{tivanjko, sspiran}@ffzg.hr

Abstract

The area of researching folksonomies is still in development, so theoretical perspective and research methods are still being defined. This study conducts a webometric and bibliometric analysis of the folksonomy research in the Library and Information science (LIS) field by collecting data from Web of Science (WOS), SCOPUS and Google Scholar in July 2014. It utilizes a total of 346 papers with 2660 citations from WOS and 1581 papers with 8848 citations from SCOPUS. In addition, Google Scholar database search was also included for providing a wider coverage of works published in conference proceedings, books and to include a wider journal base. Based on these results, research identifies most influential papers and authors across all three databases.

Keywords: Folksonomies, Social tagging, Collaborative tagging, Bibliometric analysis, Webometrics

1 Introduction

With the rise of Web 2.0, a new wave of user participation in creating and describing online resources instigated a new approach in knowledge repre-

In: F. Pehar/C. Schlägl/C. Wolff (Eds.). *Re:inventing Information Science in the Networked Society*. Proceedings of the 14th International Symposium on Information Science (ISI 2015), Zadar, Croatia, 19th–21st May 2015. Glückstadt: Verlag Werner Hülsbusch, pp. 370–379.