

***Hitching or Nitching?  
Scientific re-orientation in Economics after the German  
reunification***

***EARLY DRAFT – please do not cite or quote***

**Matthias Geissler\***  
**Andreas Rehs<sup>\$</sup>**

\*Technical University of Dresden  
Junior Research Group “Knowledge and Technology Transfer”  
Muenchner Platz 2-3  
01187 Dresden (Germany)  
E-Mail: [matthias.geissler1@tu-dresden.de](mailto:matthias.geissler1@tu-dresden.de)

and

International Centre for Higher Education Research (INCHER-Kassel)  
Moenchebergstrasse 17  
34109 Kassel (Germany)

<sup>\$</sup>University of Kassel  
Moenchebergstrasse 19  
34109 Kassel (Germany)

Print date: Mai 26, 2015

**Abstract:**

Keywords: PhD theses, Bibliometrics, German reunification  
JEL codes:

## 1 Introduction

Alternative research streams and approaches often face strong opposition in their own scientific fields. It is often hard to legitimize relatively new thoughts especially if they threaten what has been established as scientific “facts” in a research community. Opposition to the new might be fierce and opponents are likely to have better reputation and equipment. “A new scientific truth”, Max Planck already remarked, “does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it” (Planck, 1948, p. 22).

The German reunification in the 1990s led to the dismantling of a great number of chairs, institutes and whole research organizations and a broad institutional restructuring in academia in the eastern part of the country. Reasons included political motives but in several instances also a lack of fit between what had been researched under the “old” system and what was considered interesting in the “new” one. This was especially the case for Economics (including Business Administration). Whole faculties were completely re-build and often newly founded at several East-German universities. Therefore, the 1990s saw a comparatively large number of vacancies in this discipline. The new generation of professors quite often came from Western Germany and was well acquainted with the “scientific truths”. The question is, whether these researchers used the opportunity to explore new fields and ideas that had had difficulties to stand their ground against established (and perhaps more reputed) researchers in the West or whether they chose to mimic research on well-established and less ambiguous topics. In the former case we would expect a long-lasting divide between research in Economics in the eastern and western part of Germany, whereas in the latter a growing coherence could be expected.

We analyze PhD-Theses in Economics and Business Administration handed in at universities in Western and Eastern Germany before and after 1990 to shed light on this issue. We extract text corpora out of theses’ titles, which constitute a very condensed form of their research content. The 20 most frequent words of these corpora are compared in a rank correlation test. Results are compared to dissertations in Chemistry, a “hard” scientific realm, much less affected by socialist ideology before the reunification and much less a subject to restructuring thereafter.

Not surprisingly, results indicate two completely separated research streams in Economics in the two Germanies before 1990. In Chemistry this divide also exists but is less pronounced. After 1990 research topics converge in both research fields, but Economics shows much more dynamism in this process. This hints at an imitation of established and reputed research topics rather than an exploration of niches by newly established Economics departments in Eastern Germany.

The paper proceeds as follows: Section 2 elaborates on the scientific systems in the two Germanies before the reunification. We consider differences in organization of scientific work, tertiary education, granting of PhDs and involvement of the government as most relevant for our study. Furthermore we discuss the impact of the reunification, compare our two research fields, Economics and Chemistry and derive hypotheses for the development of research topics. Section 3 introduces our data and methods. Results are presented in section 4 and discussed in section 5, which concludes the paper.

## **2 Research in Germany before and after reunification**

Since the birth of both Germanies in 1949 the inner-German relationship was characterized by a competition of political (and economical) systems. Walter Ulbricht, prominent veteran Socialist of the German Democratic Republic (GDR) was renowned for his saying of “overtaking without catching-up”. The early Socialists strived to demonstrate the superiority of Socialism over Capitalism, with a central role of scientific and technological achievements. Even the constitution of the GDR claimed that the foremost aim of a Socialist society was to increase the effectiveness of scientific-technological development and labor productivity (GDR, 1974, §2, 1).

This orientation of scientific advancement on productivity considerations dated back at least to Lenin (1976) and had consequences for the academic landscape of the GDR. Industrial application of findings was heavily emphasized. Basic research was almost exclusively carried out by universities, but freedom of choice of the research objects was increasingly difficult and almost impossible since the 1960s (Gruhn and Lauterbach, 1979). PhD candidates had minimal freedom in choosing their research objects. The case of the Humboldt University in Berlin shows that roughly two thirds of the topics of dissertations followed the five-year research plans of the government (Wollgast, 2001). Furthermore, international contacts were more or less limited to members of the Council for Mutual Economic Assistance (COMECON) and access to Western academics and their publications were difficult (Mann et al., 1979). Moreover, limited financial resources made internationally competitive research impossible in the majority of scientific fields.

However, the conditions of career advancement in academia closely resembled those in the West. The average student in the GDR had to complete a basic part and an advanced (or specialized) part of his study to earn a degree. Afterwards, a dissertation (Promotion A) had to be written to obtain the degree “Dr.” in a scientific field. Contrary to the Republic of Germany, the GDR had universal requirements for the award of a PhD title that included a fair amount of ideology (GDR, 1968, §5, 1). PhD degrees could be earned through research studies (2-3 years, similar to a graduate school), employment at a university chair (usually four years contract) or through distinction in industrial and societal engagement (similar to an external PhD candidate) (Belitz-Demiriz and Voigt, 1990a; Guenther, 1989).

The transition and political change in Germany in 1990 had a deep impact on academic institutions, most notably in scientific fields that were heavily affected by socialist ideology. The prime example is Economics that was almost completely dismantled and rebuild from the ground, often involving new personnel, structures and research agendas. This fundamental change is of noteworthy interest, because it opens up the opportunity to analyze how a relatively large amount of suddenly open positions was filled and what kind of research newly appointed chairs preferred. Three avenues of development seem to be quite likely a priori: First, the newly emerging “Eastern” Economics could have formed a microcosm of its own inheriting the research focus of Economists from the GDR before 1990. For reasons mentioned above (institutional and personnel discontinuity) this seems rather unlikely, but could be discernable e. g. in research subjects (for instance with a pre-dominant focus on comparative studies between the GDR and FRG). Second, the newly instituted chairs could have fully blended in with their Western colleagues and focus on “mainstream” topics in their research field. This seems rather likely especially because many appointed chairs came from Western Germany. However, it is noteworthy that this would cause direct competition between established and new research groups and would put the newly instituted Economic faculties under severe

pressure. A third possibility is therefore that the new faculty used the historical opportunity to establish an own research agenda that was competitive but quite dissimilar from the whole field of Economics in the West. This would also resonate well with the quote by Max Planck because the newly appointed professors in the East were most likely acquainted with the newest “scientific truths” and had the opportunity to pursue research in new directions free of the shackles of evolved (and perhaps encrusted) institutions.

### 3 Data and Method

#### ***Dataset***

We probe into the issue described above by relying on PhD theses handed in at universities in the East and in the West of Germany both before and after the reunification as a formalized representation of scientific work. We examine the evolution of PhD theses’ titles in Economics and comparing them to those in Chemistry, a scientific field arguably less affected by the reunification. Our work rests on a number of presumptions: First, in Germany the advisor (often dubbed the *Doktorvater*) has a strong influence on the advisee and his choice of research topic. Moreover, the advisor is usually required to have a chair at a university as only these are entitled to grant PhDs. The second important assumption is that the title of a thesis represents its content in a very condensed form. Together, both assumptions lead to the conjecture that the research focus of an individual (the advisor) is reflected in the titles of theses handed in at an entity (a university) that he is presumably affiliated with.

We utilize the online catalogue of the German National Library (Deutsche Nationalbibliothek, DNB) as basis for our analysis. The catalogue lists the vast majority of PhD theses handed in at German universities including the former GDR.<sup>1</sup> Entries sum up to roughly one million entities for PhD theses which are classified according to subject. We use this classification to distinguish between Economic and Chemistry and employ information on university location (cities, name of university or a combination of both) to separate East from West.<sup>2</sup> Our final sample includes 11,420 dissertations.

Attention has also been paid to issues of timing. Since we are interested in the effect of the reunification, the year 1990 marks a natural cut-off point. However, we want to exclude dissertation projects that had been already underway but not completed before 1990 as those might tell a completely different story (which is a possible avenue for further research). Historical accounts moreover show slightly different averages for completing a PhD in Economics/Chemistry<sup>3</sup> especially before the reunification. Before 1990 a PhD in Economics and Chemistry took six years in the East and four years in the West (Belitz-Demiriz and Voigt, 1990; Wienke, 1989). After 1990 the average was about five years in Economics (Bomman and Enders, 2001) and could not be reliably determined in Chemistry (which is therefore assumed to also last five years on average). We take these two issues into account by choosing base years accordingly: 1978 till 1982 before the reunification

---

<sup>1</sup> The DNB has the official mission of collecting a volume of all literature (and other materials) issued in Germany including PhD theses at least since 1969 (DNB, 1969).

<sup>2</sup> The institution “University of Berlin” could not be reliably classified and entries have been removed from the data.

<sup>3</sup> Before 1990 data is only available on the level of social sciences and natural sciences, which are used for Economics and Chemistry respectively.

(corresponding to completion years 1982-1986 in the West and 1984-1988 in the East because of the different durations) and 1991 till 1995 thereafter (corresponding to completion years 1996-2000). In order to offset variation resulting from idiosyncrasies in specific years we always analyze a moving average of the focal year and its predecessor (e. g. base years 1977/78, 1978/79, etc.).

Lastly, the differentiation into East and West (see above) revealed that East Germany (both before and after the reunification) seems to produce only a fraction of PhD theses of the West. This does not seem to be a data artifact as our numbers of theses for the East is well in line with prior findings (Guenter, 1989). The practice of keeping theses secret in the GDR should also be irrelevant since all those dissertations have been made publicly available after the reunification (Bleek and Mertens, 1994). Therefore, the lower number is most likely due to a smaller number of universities in the East (and in connection with this to a lower number of inhabitants). Table 1 gives an overview.

**Table 1:** Number of dissertations per two-year-window

	Economics				Chemistry			
	East		West		East		West	
Base Year	Corr. Compl. Year	# of theses	Corr. Compl. Year	# of theses	Corr. Compl. Year	# of theses	Corr. Compl. Year	# of theses
1978	83/84	673	81/82	1,824	83/84	402	81/82	3,561
1979	84/85	678	82/83	1,853	84/85	431	82/83	2,627
1980	85/86	700	83/84	1,811	85/86	460	83/84	2,562
1981	86/87	697	84/85	1,864	86/87	447	84/85	2,743
1982	87/88	710	85/86	1,981	87/88	502	85/86	3,063
1991	95/96	276	95/96	4,060	95/96	554	95/96	6,288
1992	96/97	479	96/97	4,102	96/97	628	96/97	6,304
1993	97/98	439	97/98	4,100	97/98	693	97/98	6,199
1994	98/99	555	98/99	4,203	98/99	756	98/99	5,875
1995	99/00	586	99/00	4,276	99/00	754	99/00	5,469

To mitigate problems that may result from different sample and subsequently text corpora size, we draw samples from West German theses that match the size of theses in the East.

### ***Text Mining Preparations***

The analysis of our data follows a text mining approach labeled KDD (knowledge discovery in databases) that consists of seven phases (Maimon and Rokach, 2010). In phase three and four data is to be cleaned and transformed, which is the main concern of this section. In the cleaning phase two steps, “stopwords” and “stemming”, are employed that deserve some explanation. Our data so far consists of a collection of dissertation titles. The stopwords algorithm deletes all words that resemble common vocabulary, which is nearly identically distributed in any document of a specific language. These words mostly consist of articles, denominators, prepositions and the like (an algorithm is available for the statistical software “R”). We supplemented the typical German stopwords with a list of common scientific vocabulary that does not seem to be of any value in distinguishing either the time, nor the location of the dissertation (e. g. words like “newly”, “different”, “example”). The stemming algorithm subsequently reduces all words which survived the cleaning to its basic lexicographical form (called “lemma”). This ensures further standardization and reduces double-counting of similar words (Kilgarriff and Rose, 1998). Finally, in phase four of the KDD process the

now cleaned titles are split into text corpora according to base year. We therefore end up with a total of 40 such corpora containing all theses' titles of two subsequent years, five for each period (before and after reunification), location (East and West) and scientific field (Economics and Chemistry).

### ***Statistical Testing***

We rely on quantitative linguistics to analyze our collection of corpora. This presumes that corpora share the same language. This assumption seems to be justified as the vast majority of our titles are in German and stem from a very standardized kind of document that should also share a common scientific vocabulary.

Our approach is simple and straightforward. We sort individual words according to their frequency and compare ranks using Spearman's correlation coefficient. To do this, links of words to titles have to be dissolved which is justified in our case since we are comparing whole scientific fields, not individual titles. Moreover, Zipf's law constitutes an inverse proportionality between the rank and the frequency of a given word (Zipf, 1968 and 1972). Therefore, the top-most words in our listings should be representative of the majority of the corpus. Specifically, according to Zipf's law the first 20 ranks already contain any randomly drawn word of the corpus with a probability of almost 50%. Thus, we focus our analyses on the first 20 ranks of our listings. Since we always compare an East German to a West German corpus, we could use either top-twenty list as the basis for comparison, which may influence the result. Since our corpora should contain roughly the same amount of titles from both locations (see the sampling in West German titles above), the solution is to combine both lists into a single one (cf. Kilgariff, 2001). Finally, words that are not included in one of the lists (but in the other) cannot be compared. To get rid of this problem we assign the lowest rank ("20" in our case) to any word that is affected by this issue.

The analyses of ranks and the Spearman coefficient alone are not quite sufficient to allow for conclusions on the question of how fast titles converged (or diverged) after the reunification. We therefore conduct a test of structural breaks according to Chow (1960). We regress the Spearman coefficient on the year and test whether two separate regressions for the years 1978-82 and 1991-95 have a higher explanatory power than a single regression for the whole time span (in each scientific field).

## **4 Results**

### ***Economics***

A look at the top-twenty ranks of the combined list in Economics already reveals some interesting results (see table 2). Before the reunification, there are a total of 34 words that exist only in the East (yellow) or West (green), whereas after 1990, there is only one word. This is already suggestive of an increasing overlap in theses' topics after the reunification.

**Table 2:** Ranks of individual words in Economics (East/West combined) according to base year (green = words only in the West German corpus, yellow = only in the East German corpus; \* = rank correlation calculated taking tied ranks into account (cf. Bortz, 2005))

Rang	1977	1978	1979	1980	1981	1991	1992	1993	1994	1995
1	Entwicklung	Entwicklung	Entwicklung	Entwicklung	Entwicklung	Entwicklung	Entwicklung	Entwicklung	Entwicklung	Entwicklung
2	Planung	Planung	Planung	Gestaltung	DDR	Unternehmen	Unternehmen	Unternehmen	Unternehmen	Empirische
3	sozialistischen	DDR	Gestaltung	Planung	Planung	Deutschen	deutschen	Empirische	deutschen	Unternehmen
4	DDR	sozialistischen	DDR	DDR	Gestaltung	Neuen	Europäischen	deutschen	Empirische	Deutschland
5	Anwendung	Deutschland	sozialistischen	Industrie	sozialistischen	Europäischen	Deutsches	Europäischen	Deutschland	deutschen
6	Probleme	Gestaltung	Probleme	sozialistischen	Anwendung	Deutschland	neuen	Deutschland	Europäischen	Europäischen
7	Deutschland	Probleme	Deutschland	Deutschland	Leitung	Theoretische	Empirische	Bewertung	neuen	Theoretische
8	Bundesrepublik	Bundesrepublik	Industrie	Bundesrepublik	Nutzung	Controlling	Recht	Recht	Bewertung	Management
9	Gestaltung	Leitung	Bundesrepublik	Nutzung	Bewertung	Probleme	Gestaltung	Theorie	Ansatz	Bewertung
10	ökonomischen	Nutzung	Leitung	Probleme	Probleme	Empirische	Bewertung	ökonomische	Recht	neuen
11	Leitung	Anwendung	Nutzung	Anwendung	Produktion	Management	Konzeption	Transformation	Integration	Recht
12	Industrie	Vervollkommen	Anwendung	Arbeit	Deutschland	Rechtsvergleichende	Theoretische	Gestaltung	ökonomische	ökonomische
13	Empir	ökonomischen	VEB	ökonomischen	Arbeit	Recht	Internationalen	neuen	Internationalen	Internationalen
14	Vervollkommen	Bewertung	deutschen	Leitung	Bundesrepublik	Internationale	Theorie	Ansatz	Grenzen	Rahmen
15	Arbeit	Industrie	Vervollkommen	Vervollkommen	ökonomischen	Grenzen	Integration	Grenzen	Basis	Gestaltung
16	Organisation	Empir	Produktion	Produktion	Empir	Integration	Planung	Konzeption	Theoretische	Integration
17	Untersucht	Organisation	Empir	Entwicklungsländern	Vervollkommen	ökonomische	Bundesrepublik	Internationalen	Haftung	Basis
18	Nutzung	Arbeit	ökonomischen	Empir	Untersucht	Planung	Probleme	Rahmen	Bundesrepublik	Grenzen
19	Bewertung	Effektivität	Arbeit	ökonomische	ökonomische	GmbH	Transformation	Steuerung	Gestaltung	Deutscher
20	Effektivität	Republik	Bewertung	Untersucht	Industrie	Konzeption	Grenzen	Union	Planung	Theorie
$\rho_s$	-0,1659*	-0,1096*	-0,6274*	-0,3659*	0,0523	-	0,1804*	-0,1173*	0,6246*	0,8388*

Furthermore, the rank correlation ( $\rho_s$ ) is generally increasing over the years. This can also be illustrated in a heat map (table 3). Before the reunification a substantial amount of the top-twenty words has a high rank difference between the East and West German text corpora (indicated by yellow/red color). With the exception of the base year 1991, rank differences are less pronounced after 1990. It was not possible to calculate a rank correlation coefficient for the base year 1991, because the number of tied/missing ranks is too high. This also supports our initial conjecture that PhD projects directly affected by the reunification have to be treated separately.

**Table 3:** Heat map of rank differences in Economics between East and West (green = small difference, red = large difference)

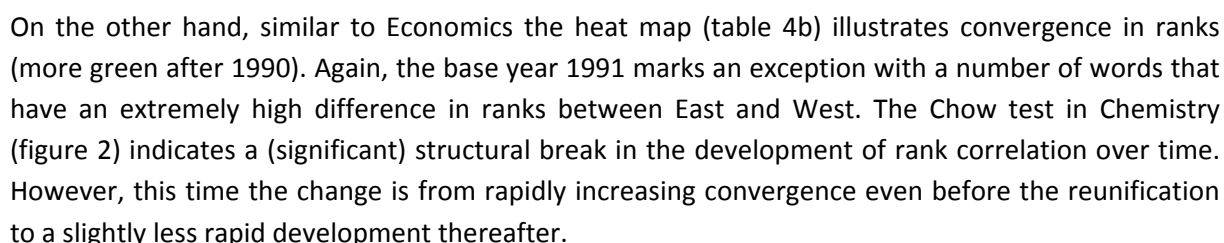
1977	1978	1979	1980	1981	1991	1992	1993	1994	1995	
Entwicklung	Arbeit	Probleme	Anwendung	Entwicklung	Unternehmen	Entwicklung	europäischen	europäischen	Entwicklung	
Planung	Entwicklung	Industrie	Gestaltung	Planung	Entwicklung	Deutschen	Unternehmen	Unternehmen	Deutschland	
Anwendung	Republik	Anwendung	Probleme	Industrie	Neuen	europäischen	Empirische	Empirische	Unternehmen	
Industrie	Planung	ökonomischen	Entwicklung	Untersucht	Europäischen	Bewertung	ökonomische	Deutschland	Deutschen	
Nutzung	ökonomischen	Planung	Produktion	Anwendung	Deutschen	Deutschland	Entwicklung	neuen	Rahmen	
Effektivität	Anwendung	Entwicklung	ökonomischen	ökonomischen	Controlling	theoretische	Steuerung	Grenzen	Deutscher	
untersucht	Effektivität	Produktion	Planung	Vervollkommenung	Probleme	Grenzen	Deutschen	Integration	Bewertung	
Probleme	Bewertung	Gestaltung	ökonomische	ökonomische	Planung	Planung	Deutschland	Entwicklung	neuen	
Arbeit	Industrie	Arbeit	Untersucht	Gestaltung	Management	Probleme	Bewertung	Deutschen	ökonomische	
Gestaltung	Probleme	Bewertung	Vervollkommenung	Nutzung	Empirische	Gestaltung	Grenzen	Bewertung	Empirische	
Bewertung	Organisation	Vervollkommenung	Nutzung	Produktion	Deutschland	Integration	Union	Recht	europäischen	
ökonomischen	Gestaltung	Verb	Arbeit	Bewertung	Theoretische	Unternehmen	Konzeption	internationalen	internationalen	
Vervollkommenung	Nutzung	DDR	Leitung	Arbeit	Recht	Konzeption	Ansatz	Bundesrepublik	Grenzen	
Leitung	Vervollkommenung	Deutschen	DDR	Leitung	GmbH	Empirische	internationalen	Gestaltung	Theorie	
Organisation	Leitung	Nutzung	Industrie	Sozialistischen	Konzeption	Recht	Rahmen	Planung	Basis	
DDR	DDR	Leitung	Entwicklungsländern	Probleme	Grenzen	neuen	Recht	theoretische	Gestaltung	
Bundesrepublik	Sozialistischen	Empir	Sozialistischen	DDR	Integration	Bundesrepublik	Transformation	Haftung	theoretische	
sozialistischen	Empir	Sozialistischen	Empir	Empir	Ökonomische	internationalen	Gestaltung	Basis	Integration	
Empir	Deutschland	Bundesrepublik	Bundesrepublik	Bundesrepublik	Internationale	Transformation	Neuen	Ansatz	Management	
Deutschland	Bundesrepublik	Deutschland	Deutschland	Deutschland	Rechtsvergleichende	Theorie	Theorie	ökonomische	Recht	
$\rho_s$	-0,1659	-0,1096	-0,6274	-0,3659	0,0523	-	0,1804	-0,1173	0,6246	0,8388

Finally, the Chow test (Figure 1) illustrates a structural break in the development of the rank correlation coefficient that is significant at a 5-percent level. Whereas rank differences before 1990 indicate neither convergence, nor divergence of topics, thereafter rank correlation rapidly converges

**Figure 1:** Chow test for structural break in the development of rank correlation (Economics)

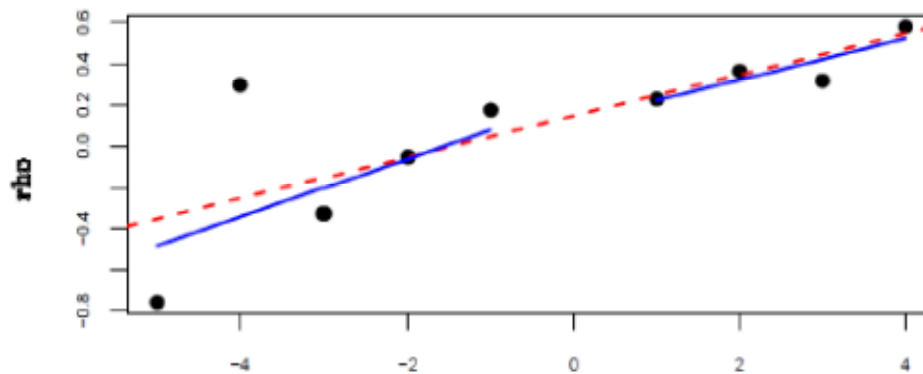


**Table 4a and 4b:** Ranks of individual words in Chemistry and heat map (for color coding see table 2 and 3 above)





**Figure 2:** Chow test for structural break in the development of rank correlation (Chemistry)



## 5 Discussion and Conclusion

We analyzed PhD-theses in Economics (and Chemistry) handed in at universities in Western and Eastern Germany before and after 1990 to shed light on this issue whether suddenly open positions at universities had been used to establish new research streams (“nitching”) or to follow more established routes of research (“hitching”). Our results indicate that the latter seems to be the case. Topics of PhD theses in the field of Economics written in the East and West of Germany, rather divided before the reunification, rapidly converge after 1990. We take this as evidence that new chair holders at universities in the East often coming from universities in the West did refrain from the opportunity of pursuing radically new ideas, but rather tended to rely on already established research topics.

Our research is not without some limitations. The usage of Spearman’s rank correlation coefficient has some drawbacks as it weights all distances between different ranks equally. Moreover, it is susceptible to tied and missing ranks. Cluster and network analysis may provide some superior instruments, but have tighter requirements on data and computational capacity. Furthermore, the text corpora that we end up with are comparatively small, because theses’ titles contain few words comparatively to what is usually used in quantitative linguistics (e. g. abstracts and full-text documents). Another challenge in this regard may come from the sampling of the West German dissertations that had to be employed to make the size of corpora comparable. Finally, the stopwords and stemming algorithm are sensitive to variations. We have already put some effort into better validation of the stopwords-list and will continue to improve on this in the future.

## References

- Belitz-Demiriz, H., Voigt, D. (1990): *Die Sozialstruktur der promovierten Intelligenz in der DDR und in der Bundesrepublik Deutschland 1950 - 1982 der Einfluss der politischen Systeme auf die unterschiedliche Entwicklung in den beiden deutschen Staaten*, Bochum: Brockmeyer.
- Bleek, W., Mertens, L. (1994): *DDR-Dissertationen - Promotionspraxis und Geheimhaltung von Doktorarbeiten im SED-Staat*, Opladen: Westdeutscher Verlag.
- Bortz, J. (2005): *Statistik für Human- und Sozialwissenschaftler* (6th ed.), Heidelberg: Springer.
- Chow, G. C. (1960): *Tests of equality between sets of coefficients in two linear regressions*, *Econometrica - Journal of the Econometric Society*, 591-605.
- GDR (1968): Regulation on Academic Degrees, 6<sup>th</sup> November 1968, <http://www.documentArchiv.de/ddr/verfddr.html>, retrieved 2014-11-27.
- GDR (1974): Constitution of the German Democratic Republic, 4th June 1968, Version of 7th Octobre 1974, <http://www.documentArchiv.de/ddr/verfddr.html>, retrieved 2014-11-27.
- Guenther, K.-H. (1989): *Das Bildungswesen der Deutschen Demokratischen Republik: Gemeinschaftsarbeit der Akademie der Paedagogischen Wissenschaften* (3<sup>rd</sup> ed.), Berlin: Volk und Wissen.
- Gruhn, W., Lauterbach, G. (1979): *Die Organisation der Forschung in der DDR*, in: I. f. G. u. W.-. IGW (Ed.), *Das Wissenschaftssystem in der DDR* (pp. 127-225), Frankfurt a.M./New York: Campus.
- Kilgarriff, A. (2001): *Comparing corpora*, *International journal of corpus linguistics*, 6(1), 97-133.
- Kilgarriff, A., Rose, T. (1998): *Measures for Corpus Similarity and Homogeneity*, Paper presented at the EMNLP.
- Lenin, W. I. (1976): *Works* (March-August 1919), Vol. 29, pp. 408-417, Berlin: Dietz.
- Mann, R., Schwarzenbach, R., Mehringer, W. (1979): *Internationale Wissenschaftsbeziehungen*, in: I. f. G. u. W.-. IGW (Ed.), *Das Wissenschaftssystem in der DDR*, pp. 233-255, Frankfurt a.M./New York: Campus.
- Maimon, O., Rokach, L. (2010): *Data Mining and Knowledge Discovery Handbook* (2<sup>nd</sup> ed.), Boston, MA: Springer US.
- Planck, M. (1948): *Wissenschaftliche Selbstbiographie - Mit einem Bildnis und der von Max von Laue gehaltenen Traueransprache*, Johann Ambrosius Barth Verlag (Leipzig), p. 22, as translated in: *Scientific Autobiography and Other Papers*, trans. F. Gaynor (New York, 1949), pp. 33-34-
- Wienke, P. (1989): *Die promovierte naturwissenschaftlich-technische Intelligenz in der DDR: eine empirische Untersuchung ihrer Karrieremuster im Vergleich mit denen von promovierten Naturwissenschaftlern und Ingenieuren aus der Bundesrepublik Deutschland*, Bochum: Brockmeyer.
- Wollgast, S. (2001): *Zur Geschichte des Promotionswesens in Deutschland*, Dr. Frank Graetz Verlag.

Zipf, G. K. (1968): *The psycho-biology of language : an introduction to dynamic philology* (2. printing ed.), Cambridge, Mass.: The M.I.T. Press.

Zipf, G. K. (1972): *Human behavior and the principle of least effort an introduction to human ecology*, New York: Hafner.