PLAGIARISM IN THE DISSERTATIONS AND SCIENTIFIC PUBLICATIONS IN RUSSIA

Andrei Rostovtsev

Abstract: Dissernet is a voluntary organization of Russian scientists and other citizens devoted to detecting and documenting research misconduct, mostly in form of plagiarism in the dissertations. During four years of running of the project, more than 6500 extensively problematic dissertations were identified and made public. Since 2016, the Dissernet has started the Journal Project. The aim of the Journal Project is to investigate misconduct in Russian scientific journals: plagiarism, duplicate publications, gifted and stolen authorship, fake peer-reviews and other violations. By now significant research misconduct in almost 3,000 of 100,000 journal papers was identified. The Dissernet is making an attempt to reconstruct a landscape of the academic fraud over the country. The project has attracted a broad attention of the Russian media and became very popular in the academic society in Russia. Unfortunately, the official governmental bodies consider the project as a dangerous threat to the state.

1   Introduction

The awarding of fake academic degrees to politicians, businessmen, doctors in clinics, professors in universities, and teachers in schools, that is, to all those who wish to use their new academic titles to step onto a faster career route, is widespread in Russia. Fake academic titles are awarded throughout the country. This business is based on the mass scale manufacture of problematic dissertations and scientific publications. Plagiarism is the easiest and fastest way to construct academic texts and this works perfectly well if the scholar attestation system is corrupted. Quite often the work involved in producing a new thesis is reduced to changing a title page of an old one, and submitting a new scientific paper is reduced to changing authorship for an earlier text and rephrasing the title. In early 2013, experts, researchers, and journalists came together to establish the organisation called Dissernet in Russia (Dissernet, 2017). It is a free association aimed at countering fraud and trickery in academia, particularly at the stage of defending dissertations and awarding PhD. By February 2017 Dissernet activists had identified over 6000 falsified dissertations and more than 2000 publications in scientific journals.

It is worth mentioning that the very meaning of plagiarism here is somewhat different from the original meaning of the term 'plagiarism' (from *plagium*—which literally means 'theft' in Latin), an intentional and unlawful incorporation of other people's texts or ideas into one's own text or research paper. Yet in Russia, most of the authors under scrutiny by Dissernet have never really done research; they have, most probably, never written a single page of their theses or publications and might have never read them or even seen them at all. Such 'scientific' works are usually nothing else but a mere compilation of other people's texts. This includes large-scale plagiarism, when tens or even hundreds pages of text are stolen from other authors as a large single block, including figures, tables, references, and typographical errors. Such large text

blocks are easy to identify automatically and this speeds up the checking of problematic texts enormously.

2   Material and Methods

The method used by the Dissernet makes an advantage of index files produced by the public search engines (such as Google or Yandex). First the text under study is split by the machine into many phrases. Then the phrases are submitted automatically one by one to the search engine and the resulting statistics are analyzed by an algorithm such that the most frequent found text sources are identified on each page of the studied text. In addition to the detection of thousands of fraudulent dissertations, the software is the main result of a unique technology developed by Dissernet association. In Russia, along with the dissertation a so-called *avtoreferat* must be made publically available before the Ph.D. defense. The *avtoreferat* consists of a shortened dissertation content (usually 20–30 pages) reporting on the main research results. Importantly, the texts of the *avtoreferats* are indexed by public search engines, while the dissertations itself are not usually indexed. But if the dissertation contains large fragments of plagiarized text, as described above, its *avtoreferat* would contain these text fragments as well. In case these fragments come from another earlier dissertation its corresponding *avtoreferat* provides the evidence for plagiarism. The specific Dissernet software is able to pick up the *avtoreferats* one by one and look for textual coincidences within the whole publicly available corpus of Russian digitized texts, including texts of other avtoreferats. This program runs 24 hours a day and 7 days a week. Thus a few hundred thousands dissertations have been automatically checked. Furthermore, Dissernet takes advantage of the common practice of a chain-like fraudulent dissertation production. As soon as a rampant plagiarism is detected in one dissertation, it is very likely to be detected as well in other dissertations defended by the same dissertation council or with the same supervisor.

Similarly the Dissernet software works with the texts of papers published in the open access Russian scientific journals. Typically more than one hundred of the latest available papers are downloaded for each journal title. By now about thousand journal titles have been checked (corresponding to approximately 100 000 papers) resulting in almost 3000 papers with serious academic misconduct identified.

3   Results and Discussion

3.1   *Plagiarism in dissertations*

The collected statistics over 6500 dissertations with large-scale plagiarism identified allow several important conclusions to be made.

- Most of the problematic dissertations (38%) are produced in economics. Other popular fields are pedagogy (19%) and law (13%), followed by medical sciences, political sciences, engineering, and social sciences. Fake dissertations are rare in the area of natural sciences. Such distribution is symptomatic as it represents the Russia's major everyday problem areas: economy, law, education, etc..

- Since the Dissernet deals with large-scale plagiarism, only the tip of the iceberg is seen. An average rate of fake dissertations identified by the Dissernet varies from three to five per cent in problematic fields and much lower in natural sciences.
- The collected data show that forgery in the academic sphere in Russia is by no means just some fringe phenomenon but an integral part of Russian science. Geographically speaking, problematic dissertations are mass-produced primarily in Moscow—Russia's political and business capital, and in Saint Petersburg rather than somewhere on the outskirts of the country. Other cities and towns fall behind.
- Almost all problematic dissertations in Russia turned out to be produced in the universities. Forgery in academic scientific centres also existing but it is relatively rare phenomenon. Interestingly, the scientific outcome from the academic scientific centres as measured bibliometrically based on the well-known international databases largely overweight one from the universities.

There are few important sociological observations to be made. The academic fraud is abundant where the scientific outcome is miserable. In order to illustrate this axiom one plots a relative contribution to the international corpus of scientific papers indexed by SCOPUS versus the fraction of fake dissertations identified by the Dissernet for different scientific areas. For the natural sciences there is a handful of plagiarized theses only while the relative contribution to the world first class papers rises up to 5%, which is the top record for Russian science. On the other end of the plot are social and humanitarian sciences with thousands fake dissertations and hardly visible scientific contribution less than 0,5% or so. The gap is partially filled by medical and engineering sciences, see Figure 1. As a conclusion, the academic fraud is abundant where an expert scientific community is weekend or missing at all.

As was mentioned above an average probability to find large scale plagiarism in a dissertation is at the rate of few per cent only. Now it is interesting to ask how this rate changes when one selects certain social groups. The Dissernet association has studied several sociological groups: directors of Moscow schools, rectors of Russian universities, heads of regional governments and the deputies of the Parliament. Surprisingly, the rates in these sociological groups is much higher than the average one. In numbers it reads as about 15, 20, 30 and 40% correspondingly. The higher social responsibility or higher social influence of the group studied the easier the representatives of this group go for falsifications or lower their reputational niveau in average. Such trend illustrates a negative sociological selection at work.

3.2   *Plagiarism in scientific papers*

The major problem with many Russian scientific publications is the weakness of the peer review institution. In some journals, the peer review process is only declared but actually does not exist at all. As our experience shows, such journals are an easy target for authors wishing to publish fraudulent papers. The authors' goal is to inflate their publishing activity. Sometimes they send the same text to more than ten journals, just changing a title of the publication for each journal.
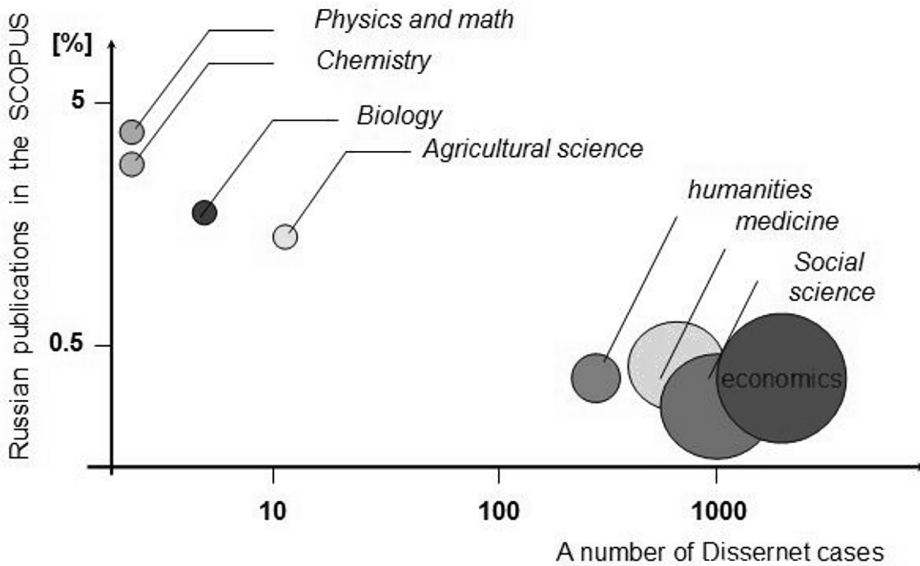
Figure 1. A fraction of Russian papers in SCOPUS as function of a number of problematic dissertations for different scientific fields

The Russian open access journals were tested by selecting at least a hundred of the most recent publications for each of about one thousand journals. In the case that more than half of the text of the paper (article A) is identical to the text of some other earlier work (mainly from one single source and not necessarily from the same journal—article B), then this article A is classified depending on its authorship:

- if the composition of the authors of the paper A does not overlap even partly with the composition of authors of the article B, this case belongs to the category of publications with identified plagiarism in the text. Here we are deliberately talking about the large-scale text coincidences in articles A and B;

- if the composition of the authors of the paper A fully coincides with the composition of the authors of the article B, this case belongs to the category of duplicate publication, where authors suggest the same text often with different title to various editors;

- finally, the intermediate case, when the composition of the authors of the article A and article B partially overlap. This case belongs to the fictitious co-authorship when the place in the author list for one reason or another is offered to third parties who had nothing to do with the original manuscript.

All three of the above-described cases are different kinds of incorrect publication practices and at the same time—different sorts of plagiarism. Statistical analysis of one thousand of cases of incorrect publication practices in Russian scientific journals have shown that large-scale plagiarism and multiple publications occur with approximately

equal probability (38% and 37%, respectively), while the problematic authorship is found in 25% of cases. These results are in a broad agreement with the statistics of publications retraction causes in Western scientific journals.

A detailed analysis of Russian scientific journals which systematically violate ethical standards allows to roughly divide them into three categories.

- The first and perhaps currently the most important category is the so-called "predatory" journals. The very term "predatory" journal have been coined by librarian and professor at the University of Colorado Jeffrey Beall (Beall, 2012). Predatory journals are characterized by an unusually large number of short publications on a wide range of scientific specialties, the rapid growth in the number of publications over time, lack of influential scientists in the editorial Board, short guaranteed period of peer review and unreasonable minimization of editorial and publishing costs. Over the last few years the number of publications in such journals in Russia has increased by almost an order of magnitude. At the same time the dynamics of the traditional publications in university journals shows a steady decline in the number of publications over time. It must be stressed that this decline occurs when the total number of publications in Russian scientific periodicals is growing from year to year. This apparent paradox is explained by the powerful outflow of authors from traditional journals into predatory journals that are increasingly occupying the publication space. Indeed, to publish a manuscript in predatory journal is incomparably easy and fast. In fact, the predatory journals exploit the modern model of Gold Open Access for pure commercial purposes.

- The second category consists of weak, traditional academic journals in which the review procedure exists only nominally. As already noted, in recent years, with the flourishing business of the predatory journals the weak traditional magazines constantly lose authors and exhibit a rapid decrease in the number of publications over time. Some of these magazines cease to exist. Magazines in this category differ from the predatory journals by non-aggressive policies used to attract authors and highly compact geographical location of authors and members of editorial boards.

- Finally, the third category includes magazines, which are a part of the "factories" for mass production of fake dissertations discussed above. In Russia in order to defend theses one is obliged to publish several scientific papers on the subject. Thus the commercial offer of ready to use fake dissertations must include several publications in scientific journals. Partly the members of the dissertation councils at the universities are incorporated to the editorial boards of such journals.

Such a catastrophic situation with Russian scientific periodicals is the result of a serious weakening of the level of expertise in Russian scientific community as a whole, on the one hand, and, on the other hand, high formal bibliometric requirements by the Ministry of education and science applicable to the publication activity of Russian scientists and university staff.

4   Conclusion

This study of misconduct, mostly in form of plagiarism, in Russian dissertations and scientific journal papers is based on rich statistics collected by the voluntary Dissernet association. It helped to reconstruct and visualize a landscape of falsifications in academic world in Russia. The results of this work are important not only for identification of major players on the corrupted market of fake academic degrees, but serve for consolidation of scientific community in the country. The project has attracted a broad attention of the Russian media and became very popular in the academic society in Russia. Unfortunately, the official governmental bodies consider the project as a dangerous threat to the state.

Acknowledgements

Author is very grateful to all colleagues who contributed to the work of the Dissernet association.

Literature

Dissernet [online]. Available at https://www.dissernet.org/

BEALL, J. (2012). Predatory publishers are corrupting open access. *Nature*. 489 (7415): 179.

Copyright statement

Author

Andrei Rostovtsev (dissernet@gmail.com), IITP RAS Moscov, Russia