

X Всероссийская научно-практическая конференция

«Фонды библиотек в цифровую эпоху: традиционные и электронные ресурсы, комплектование, использование»

О результатах поиска дублированных научных публикаций в базе eLibrary

Юрий Чехович, к.ф.-м.н.,
Исполнительный директор компании Антиплагиат

Самоцитирование

Самоцитирование (text recycling) – повторное использование автором текстов или фрагментов текстов из своих ранее написанных произведений

Термин self-citation не имеет отношения к обсуждаемому вопросу

Самоцитирование и самоплагиат

По лезвию бритвы: как самоцитирование не превратить в самоплагиат¹

А. В. Кулешова^{1,2}, Ю. В. Чехович³, О. С. Беленькая³

¹ *Всероссийский центр изучения общественного мнения (ВЦИОМ), Москва, Россия*

² *Ассоциация научных редакторов и издателей (АНРИ), Москва, Россия*

³ *Компания «Антиплагиат», Москва, Россия*

Кулешова А.В., Чехович Ю.В., Беленькая О.С. По лезвию бритвы: как самоцитирование не превратить в самоплагиат. *Научный редактор и издатель*. 2019;4(1-2):45-51. <https://doi.org/10.24069/2542-0267-2019-1-2-45-51>

Самоцитирование в зависимости от жанра научной работы

Из:\nВ:	Тезисы	Доклад	Научная статья	Диссертация	Монография	Научный отчет
Тезисы	1 ✓	1 ✓	✓	✓	✓	4 !
Доклад	1 ✓	1 ✓	✓	✓	✓	4 !
Научная статья	1 ✓	1 ✓	2, 3 !	✓	✓	4 !
Диссертация	1 ✓	1 ✓	7 !	8, 9 ! -	✓	4 !
Монография	1 ✓	1 ✓	-	✓	-	4 !
Научный отчет	5 !	5 !	5 !	5 !	5 !	6 !

- 1 – зависит от требований программного комитета конференции
- 2 – объем самоцитирования определяется требованиями к общему допустимому объему неоригинального текста, установленного редакцией
- 3 – см. разделы «Дублирующая публикация» и «Перепубликация на другом языке»
- 4 – ограничения по самоцитированию в научном отчете устанавливаются заказчиком
- 5 – возможность дальнейшего использования материалов научного отчета

- регулируется заказчиком
- 6 – допустимо из промежуточного отчета в итоговый
- 7 – редакция может устанавливать ограничения, например, не более 1 года после защиты
- 8 – из кандидатской диссертации в докторскую с ограничениями по объему относительно докторской диссертации
- 9 – если диссертации относятся к разным областям знаний

Нарушение публикационной этики и не только



Text recycling

Text recycling guidelines

These guidelines are intended to guide editors when dealing with cases of text recycling.

Text recycling, also known as self-plagiarism, occurs when sections of the same text appear (usually un-attributed) in more than one of an author's own publications. The term 'text recycling' has been chosen to differentiate from 'true' plagiarism (i.e. when another author's words or ideas have been used, usually without attribution).


A separate issue, not to be confused with text recycling, is redundant (duplicate) publication. Redundant (duplicate) publication generally denotes a larger problem of repeated publication of data or ideas, often with at least one author in common. This is outside the scope of these guidelines and is covered elsewhere^{1,2}.

Journals should also ensure that they have a clear policy on duplicate publication, detailing what is considered a previous publication and informing authors of the need to declare any potentially overlapping publications and cite them.

How can editors deal with text recycling?

Editors should consider each case of text recycling on an individual basis as the 'significance' of the overlap, and therefore the most appropriate course of action, will depend on a number of factors. These factors will be discussed in more detail below and include:

- How much text is recycled
- Where in the article the text recycling occurs
- Whether the source of the recycled text has been acknowledged
- Whether the article is a research or non-research article
- Whether there is a breach of copyright
- In some circumstances, cultural norms at the time and place of publication

 **BioMed Central**
The Open Access Publisher

biomedcentral.com Part of **SPRINGER NATURE**

Дублирование публикаций (duplicated publication)

ГРАЖДАНСКИЙ КОДЕКС РОССИЙСКОЙ ФЕДЕРАЦИИ

ЧАСТЬ ЧЕТВЕРТАЯ

N 230-ФЗ

Принят
Государственной Думой
24 ноября 2006 года

Одобен
Советом Федерации
8 декабря 2006 года

С последующими изменениями и дополнениями

18 декабря 2006 года

Количественный анализ. Цель исследования

Выяснить насколько распространено такое явление как дублирование научных публикаций в русскоязычных статьях

Исключение: не исследовалось самоцитирование в условиях смены жанра (например, из статьи в диссертацию, из диссертации в монографию, из тезисов в статью)

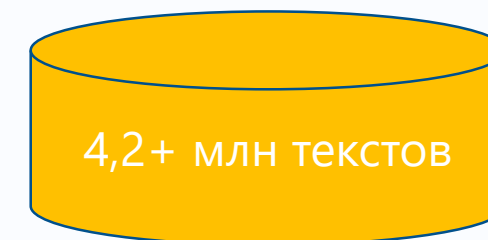
Данные для исследования

Коллекция полнотекстовых документов Elibrary



1. Тип публикации: научные публикации, с типом «Научная статья» и категорией «Статья в журнале» (Article), «Краткое сообщение», «Обзорная статья», «Не определен»
2. Язык публикации: основным языком – русский
3. Размер публикации: длина текста не менее 5000 знаков
4. Полнота метаданных: в метаданных публикации содержится информация об авторах и годе публикации.

Данные для исследования



Алгоритм выявления дубликатов

- Для каждой отобранного документа выполняем проверку на заимствования по описанной выборке с помощью системы Антиплагиат
- Все документы, с которым находится пересечение более, чем на 80 % от размера текста исходного документа, мы объединяем в кластер, включая в него и исходный документ
- Удаляем из кластера сборники и журналы, размещенные одним файлом

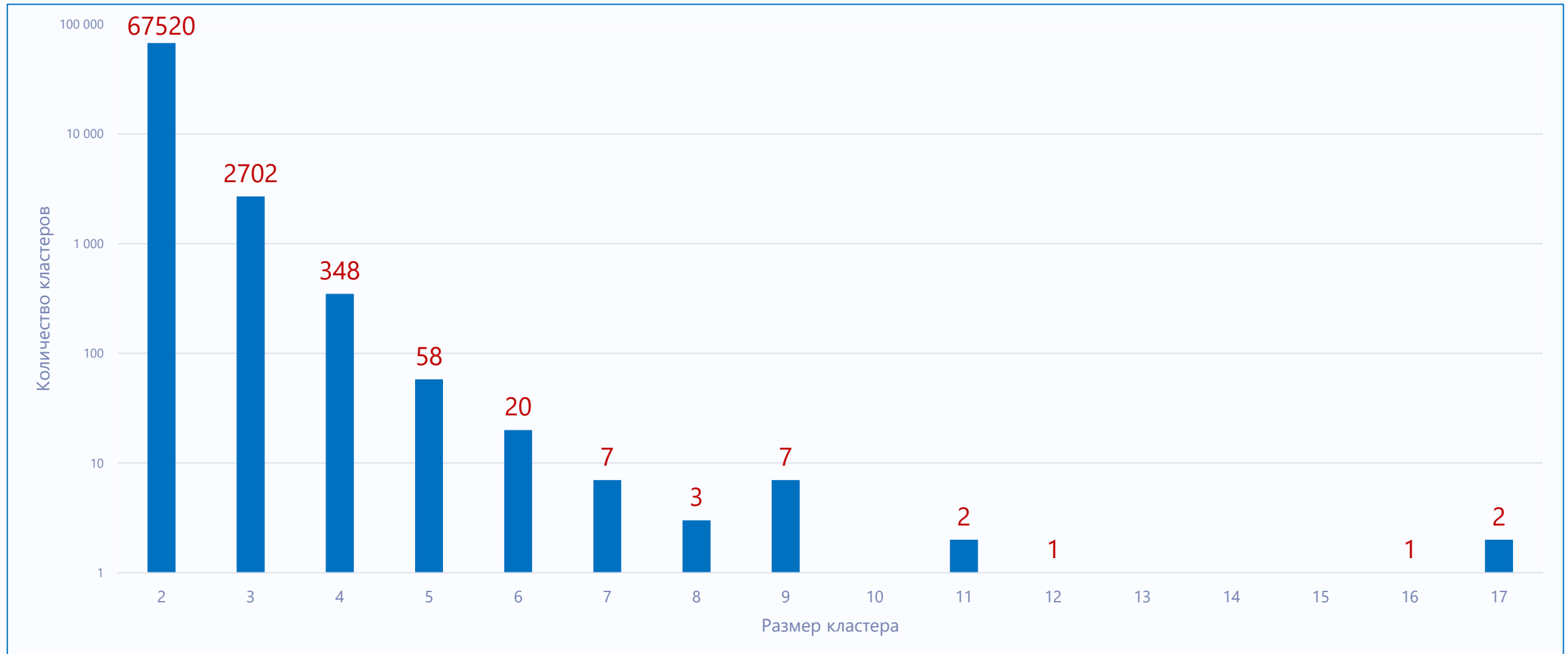
Результаты поиска

Каждый кластер включает в себя все статьи, основанные на одном тексте

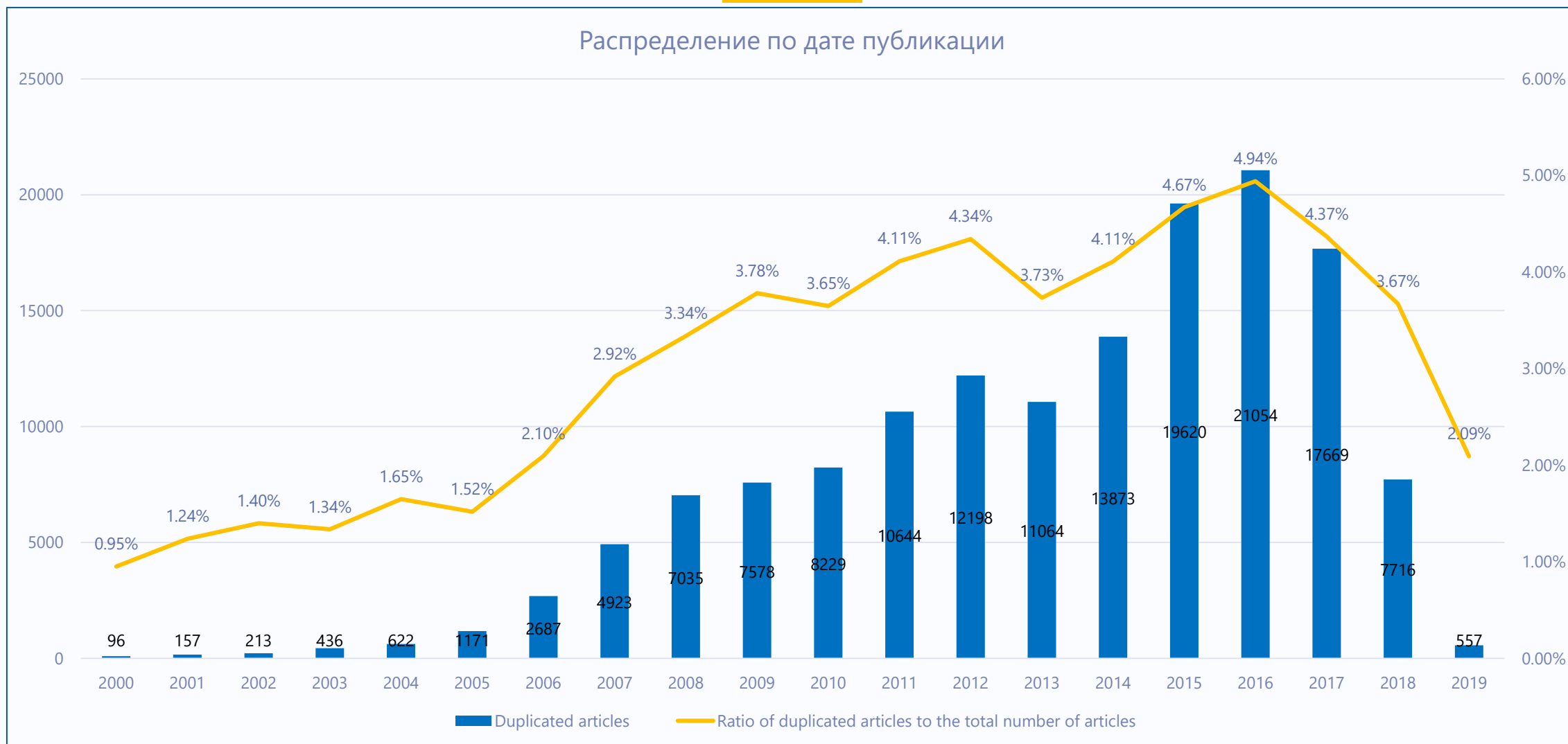
**Всего было обнаружено 70 671 случаев дублирования,
включающих 145 168 статей**

Результаты поиска

Сколько раз можно опубликовать один и тот же текст?




Результаты поиска



Почему возникают дублированные публикации?

- Наличие у многих авторов желания нарастить свою публикационную активность в отсутствие реальных результатов для публикации



Эффект Тараса Бульбы

Я тебя породил,
я тебя и...

повторно опубликую!

Почему возникают дублированные публикации?

- Наличие у многих авторов желания нарастить свою публикационную активность в отсутствие реальных результатов для публикации
- Авторы используют «веерные рассылки» рукописей, которые сильно снижают шансы журналов выявить дублирование
- Слабость (или полное отсутствие) процесса рецензирования у многих журналов, в том числе, отсутствие контроля рукописей на наличие заимствований

Некоторые выводы

- Дублирование публикаций действительно распространенное явление
- Значительная часть дублированных статей в настоящее не отозваны
- Действующий издательский процесс лишает журналы возможности своевременно обнаружить попытку самоплагиата
- Нужен сервис для отслеживания самоплагиата после публикации

Первая публикация

7-я международная конференция
Электронные научные и образовательные ресурсы:
создание, продвижение и использование

Количественный анализ случаев самоплагиата в российских научных публикациях

Юрий Чехович, Исполнительный директор компании Антиплагиат, к.ф.-м.н.
Андрей Хазов, Руководитель группы обработки данных компании Антиплагиат

Презентация: <https://doi.org/10.13140/RG.2.2.35600.76801>

Видеозапись: https://conf.neicon.ru/index.php/science/overseas2019/pages/view/overseas2019-video?video=/27/27_Chekhovich

Последние новости

- Здесь представлены результаты предварительного исследования.
- В настоящее время по выверенным результатам подготовлена и подана в научный журнал статья (в один(!) журнал 😊)
- Данные о найденных дубликатах будут размещены в открытом доступе

И еще

Похоже, что значительную часть из 140 тысяч статей придется отзывать...

Спасибо за внимание!

Ваши вопросы

Юрий Чехович,

Исполнительный директор компании Антиплагиат, к.ф.-м.н.

chehovich@antiplagiat.ru

+7 495 223 23 84

8 800 777 81 28