

DOI: 10.24069/2542-0267-2018-1-2-38-43

МЕЖДУНАРОДНЫЕ СТАНДАРТЫ

Стандартизация и гармонизация издательской политики в отношении исследовательских данных*

Иан Хрынаскевич^а, Александр Бирюков^б, Матиас Астель^с,
Совмя Сваминатан^с, Эми Кенал^а, Варша Ходияр^д

^а Издательство *Springer Nature*, Лондон, Великобритания;
e-mail: iain.hrynaskiewicz@nature.com

^б *Springer-Verlag GmbH*, Гейдельберг, Германия

^с Журнал *Nature Research*, Сан-Франциско, США

^д Журнал *Scientific Data*, Лондон, Великобритания

Резюме: Специалисты Издательского дома *Springer Nature* разработали стандартизованные основы политики в отношении исследовательских данных (*research data*) для реализации во всех его журналах. Основной целью было приблизиться к разрешению проблем, с которыми сталкиваются авторы в ходе публикации результатов научного исследования, а также оценить потенциальные выгоды применения прозрачной политики исследовательских данных (*research data policy*) для всего научного сообщества. Для выявления характерных черт разных видов политики, используемой в журналах *Springer Nature*, и их анализа была сформирована специальная экспертная группа. Эксперты провели интервью с 30 редакторами, представляющими журналы из всех тематических дисциплин издательства. Кроме того, члены экспертной группы провели ряд консультаций с научными редакторами, библиотекарями и представителями финансирующих организаций. Было идентифицировано четыре основных типа политики в отношении исследовательских данных. Обнаружено, что некоторые журналы и научные сообщества проявляют большую по сравнению с другими готовность применять строгую политику. На январь 2017 г. свыше 700 журналов приняло на вооружение разработанную стандартизованную политику, причем их число растет еженедельно. Для того чтобы облегчить процесс унификации и гармонизации политики в отношении исследовательских данных между финансирующими организациями, научно-исследовательскими учреждениями, репозиториями, научными сообществами и другими издателями, разработанные концептуальные основания опубликованы с использованием лицензии *Creative Commons*. Широкое обсуждение данных рамочных положений среди всех заинтересованных сторон процесса научной коммуникации планируется провести в сотрудничестве «Альянсом исследовательских данных» (*Research Data Alliance*, RDA).

Ключевые слова: издательские политики, открытая наука, обмен данными, научные данные, стандарты, научные журналы

Благодарность: Авторы выражают признательность доктору Эндрю Хавтону (Dr. Andrew Hufton) за помощь в составлении перечня репозитория и используемых политик. Редакция журнала выражает благодарность Наталье Геннадьевне Поповой за перевод текста статьи.

Для цитирования: Хрынаскевич И., Бирюков А., Астель М., Сваминатан С., Кенал Э., Ходияр В. Стандартизация и гармонизация издательской политики в отношении исследовательских данных. *Научный редактор и издатель*. 2018;3(1-2):38–43. DOI: 10.24069/2542-0267-2018-1-2-38-43.

* Перевод статьи: Hrynaskiewicz I., Birukou A., Astell M., Swaminathan S., Kenall A., Khodiyar V. Standardising and Harmonising Research Data Policy in Scholarly Publishing. *International Journal of Digital Curation*. 2017;12(1):65–71. DOI: [10.2218/ijdc.v12i1.531](https://doi.org/10.2218/ijdc.v12i1.531).

Standardising and Harmonising Research Data Policy in Scholarly Publishing

Iain Hrynaszkiewicz^a, Aliaksandr Birukou^b, Mathias Astell^c, Sowmya Swaminathan^c, Amye Kenall^a, Varsha Khodiyar^d

^a Springer Nature, London, United Kingdom; e-mail: iain.hrynaszkiewicz@nature.com

^b Springer-Verlag GmbH, Heidelberg, Germany

^c Nature Research, San-Francisco, USA

^d Scientific Data, London, United Kingdom

Abstract: To address the complexities researchers face during publication, and the potential community-wide benefits of wider adoption of clear data policies, the publisher Springer Nature has developed a standardized, common framework for the research data policies of all its journals. An expert working group was convened to audit and identify common features of research data policies of the journals published by Springer Nature, where policies were present. The group then consulted with approximately 30 editors, covering all research disciplines within the organisation. The group also consulted with academic editors, librarians and funders, which informed development of the framework and the creation of supporting resources. Four types of data policy were defined in recognition that some journals and research communities are more ready than others to adopt strong data policies. As of January 2017 more than 700 journals have adopted a standard policy and this number is growing weekly. To potentially enable standardization and harmonization of data policy across funders, institutions, repositories, societies and other publishers, the policy framework was made available under a Creative Commons license. However, the framework requires wider debate with these stakeholders and an Interest Group within the Research Data Alliance (RDA) has been formed to initiate this process.

Keywords: journal policy, research data policy, open science, data sharing, scientific data

Acknowledgements: For contributions to policy and repository list development, Dr Andrew Hufton, Managing Editor, Scientific Data.

For citation: Hrynaszkiewicz I., Birukou A., Astell M., Swaminathan S., Kenall A., Khodiyar V. Standardising and Harmonising Research Data Policy in Scholarly Publishing. *International Journal of Digital Curation*. 2017;12(1):65–71. DOI: 10.2218/ijdc.v12i1.531.

ВВЕДЕНИЕ

Политика в отношении исследовательских данных, практикуемая в научных журналах, в разной степени влияет на готовность исследователей ими делиться [1; 2]. Все большее число организаций, финансирующих науку, а также исследовательских учреждений начинают применять политику, регулирующую процесс обмена исследовательскими данными. Среди них — «Национальные институты здравоохранения США» (*National Institutes of Health, NIH*), «Фонд Билла и Мелинды Гейтс» (*Gates Foundation*), европейская программа «Горизонт — 2020» (*EU Horizon 2020*), фонд «Веллком Траст» (*Wellcome TrustFund*) и семь отделений Совета Великобритании по науке (*Research Councils, UK*) [3]. В учреждениях используется разная политика в от-

ношении данных. Например, некоторые требуют от авторов информацию о том, где их планируется разместить, а другие, такие как «Исследовательский совет по инженерно-техническим и физическим наукам» (*Engineering and Physical Sciences Research Council, EPSRC*), включают в публикуемые научные статьи информацию о том, где данные, полученные в ходе исследования, были архивированы.

Для повышения воспроизводимости публикуемых исследований научные журналы, сообщества и конференции¹ все чаще используют политику регулирования обмена данными, ко-

¹ Например, конференции ECMLPKDD 2016 (<http://www.ecmlpkdd2016.org/submission.html#call-conference>) или ISWC 2016 (<http://iswc2016.semanticweb.org/pages/calls/resource-track.html>).

торая должна отражать как потребности и стандарты, характерные для данного научного сообщества, так и требования финансирующих организаций.

Однако, поскольку и количество издателей, и количество публикуемых ими журналов велико, разброс в подходах к управлению данными достаточно велик [4]. Более того, во многих журналах отсутствует четко сформулированная политика в отношении исследовательских данных, хотя и предполагается, что журналы и издатели гарантируют соблюдение автором требований финансирующей организации, особенно это касается политики открытого доступа (более применимо для статей, а не для данных). Регулирование процесса обмена данными, лежащего в основе опубликованных научных статей, представляет дополнительные выгоды как отдельным исследователям в части увеличения видимости и цитируемости их работ [5], так и издателям в виде возможности распространять более содержательный контент. Кроме того, повышается применимость полученных данных к процессу развития научного знания.

Попытка создать базу всех существующих политик исследовательских данных, опираясь на предыдущие проекты каталогизации политик открытого доступа была предпринята организацией *Jisc*. К сожалению, данная инициатива не была завершена отчасти из-за отсутствия единства в стандартах и гармонизации политик управления научными данными в разных учреждениях и научных сообществах [4]. Часто на веб-сайтах научных журналов представлена неоднозначная информация об обязательных и рекомендательных аспектах их политики. Не всегда также ясно, каким образом она реализуется. Политика в отношении обмена данными в конечном итоге должна способствовать более «открытой» науке, поскольку открытость научной информации есть гарант исследований и инноваций высокого качества, о чем свидетельствуют достижения научных сообществ в области кристаллографии, генетики, археологии и лингвистики [6].

СТАНДАРТИЗАЦИЯ ПОЛИТИКИ ИССЛЕДОВАТЕЛЬСКИХ ДАННЫХ ИЗДАТЕЛЬСТВА SPRINGER

Специалисты Издательского дома *Springer Nature* разработали концептуальные рамки политики исследовательских данных для реализации во всех его 2500 журналах. Основной целью было преодолеть сложности, с которыми сталкиваются

авторы в ходе публикации результатов их работы, а также оценить потенциальные выгоды для научного сообщества от применения прозрачной политики в отношении данных.

Была сформирована экспертная группа для анализа и выявления типичных черт политик, используемых в журналах группы *Springer Nature*. Члены данной группы провели ряд консультаций с 30 научными редакторами, ответственными за журналы Издательства разной тематической направленности. Кроме того, были проведены интервью с научными редакторами, библиотекарями и представителями финансирующих организаций для создания списка часто задаваемых вопросов (FAQ)². Были определены четыре типа политики в отношении управления исследовательскими данными (рис. 1).



Рис. 1. Типы политики исследовательских данных в журналах *Springer Nature*

² Часто задаваемые вопросы, политика *Springer Nature* в отношении исследовательских данных (<http://www.springernature.com/gp/group/data-policy/faq>).

Таблица 1

Типы политик исследовательских данных и их характерные черты

Характерная черта	Описание	Тип 1	Тип 2	Тип 3	Тип 4
Поддержка обмена данных через репозитории	Информация о возможности обмена данными через репозитории представлена в «Информации для авторов»	●	●	●	●
Цитирование данных разрешено	Библиографический стиль, используемый в журнале, разрешает авторам ссылаться в списках литературы на публично-доступные базы данных	●	●	●	●
Служба поддержки по вопросам размещения исследовательских данных	Контакты службы поддержки журнала представлены в «Информации для авторов»	◐	◐	◐	◐
Проверка депонированных данных и их идентификация в отдельных случаях	Факт депонирования данных проверяется на одном из этапов публикационного процесса, если того требует устоявшаяся практика данного научного сообщества	○	◐	●	●
Сообщение о том, где размещены исследовательские данные	Сообщение о том, каким образом получить доступ к данным исследования, размещается в метаданных опубликованной статьи	○	◐	●	●
Требуется и проверяется факт депонирования исследовательских данных и присвоение им идентификатора	Исследовательские данные всех опубликованных работ размещаются в открытом доступе, присвоенные им идентификаторы публикуются вместе со статьей (исключение составляет персональная и другая конфиденциальная информация)	○	○	◐	●
Цитирование данных	Ссылки на цитируемые блоки данных размещаются в списках литературы и их корректность проверяется	○	○	◐	●
Рецензирование данных	В руководстве по проведению процедуры рецензирования размещается информация о деталях экспертизы файлов, содержащих исследовательские данные	○	○	◐	●
Репозитории данных	Электронная система подачи рукописей / процесса рецензирования интегрирована с узкоспециализированными или универсальными репозиториями (напр., Figshare ³)	○	○	◐	●

● — обязательный аспект

◐ — рекомендательный аспект

○ — не требуется

Отдельные журналы и научные сообщества находятся на разных стадиях готовности к поддержке открытости в обмене данными: от тех, кто полностью поддерживает открытый доступ, до тех, кто лишь приступает к обсуждению такой возможности. Выделенные четыре типа обнаруживают эти отличия и позволяют выработать лучшие и прозрачные рамки для продвижения лучших практик обмена данными во всех областях научного знания.

Концептуальные основания политики⁴ исследовательских данных в *Springer Nature* ус-

ловно выделяют девять ее характерных черт (табл. 1). Каждый тип политики имеет определенное число черт, причем политика 1-го типа имеет их наименьшее количество, а 4-го — наибольшее. В рамках политики 1-го типа поддерживается обмен данными и их цитирование, исследователям предоставляется список репозиториях данных. Политика 2-го типа приветствует размещение в статьях информации о том, где депонированы данные, тогда как в рамках политики 3-го типа — это обязательное требование. Политика 4-го типа требует не только размещения данных исследовательской работы в открытом доступе, но и предполагает их экспертное рецензирование.

Важно отметить, что выбор типа политики в области обмена исследовательскими данными не определяется качеством журнала, его престижностью или импакт-фактором. Журналы выбира-

³ Figshare, <https://figshare.com>, репозиторий, где исследователи могут депонировать и обмениваться данными по их исследованиям (таблицами, графиками, видеофайлами и др.) (прим. переводчика).

⁴ Текст политики доступен по лицензии Creative Commons (<http://www.springernature.com/gp/group/data-policy/policy-types>).

ют тот тип политики, который максимально отражает интересы его аудитории. Таким образом продвигаются лучшие практики в соответствующем научном сообществе.

Таблица 2

Примеры журналов, выбирающих тот или иной тип политики исследовательских данных

Тип политики	Журнал	Ссылка на веб-сайт журнала
1	Machine Vision and Applications	http://www.springer.com/computer/image+processing/journal/138
2	Plant and Soil	http://www.springer.com/life+sciences/plant+sciences/journal/11104
3	Palgrave Communications	http://www.palgrave-journals.com/palcomms/about/editorial-policies#Availability
4	Scientific Data	https://www.nature.com/sdata/policies/data-policies

Проведенный нами первоначальный обзор политик, применяемых в журналах *Springer Nature*, позволил обнаружить, что биомедицинские журналы, такие как *Nature* и *BioMed Central*, с большей долей вероятности используют какую-либо политику в отношении исследовательских данных. Однако концептуальные основы политики, разработанные нами без учета дисциплинарной специфики, были приняты на вооружение в журналах разной тематической направленности, включая гуманитарные дисциплины, математику, физику, а также социальные, технические и компьютерные науки.

Такие черты политики, как информация о репозиториях, цитировании данных и наличии службы поддержки характерны для всех четырех типов. Служба поддержки⁵ призвана рекомендовать авторам выполнять требования политики и помочь найти подходящие репозитории. Кроме того, ее функция — помочь редакторам разных журналов выбрать подходящую для их нужд политику исследовательских данных.

Описанные типы политик стали активно применяться во втором квартале 2016 г. в разных журналах *Springer Nature* (собственно издания *Nature*, *BioMed Central*, а также журналы, размещенные на *SpringerLink*) [7]. На январь 2017 г. в списке журналов, применяющих стандартизованную политику, было уже 700 изданий, причем

их число увеличивается еженедельно. На тот период наиболее распространенным типом применяемой политики был тип 3, который выбрали более 340 журналов. Среди них в основном были издания *BioMed Central*, такие как *BMC Genomics*, и журналы *Nature*⁶ [8]. Несмотря на то что есть четкий дисциплинарный тренд в выборе политики типа 2 и 3 биомедицинскими журналами, на этот процесс влияют и другие факторы. Например, в журналах с меньшим количеством сотрудников, где труднее выполнять рутинные проверки на выполнение авторами требований политики, чаще выбирают тип 1 или 2. В журналах, где ранее не практиковалось никакой политики в отношении обмена данными, как правило, также выбирают тип 1 или 2. Политика типа 4, требующая открытого доступа к данным в каждой публикации, пока наименее широко применяется и в основном используется журналами, которые либо привержены идее открытого обмена данными в силу специфики, либо уже функционируют в научном сообществе с установившимися «стандартными» репозиториями. Планируется также внедрять политику исследовательских данных в монографиях и сборниках трудов конференций, в которых представлены результаты оригинальных исследований.

За первые полгода функционирования службы поддержки с вопросами обратилось множество авторов и редакторов. Среди наиболее часто обсуждаемых проблем: поиск репозитория, выбор подходящего типа политики журналами, оформление сообщения о депонировании данных. Авторам предоставляются при этом рекомендации о том, как достигать соответствия требованиям финансирующих организаций, и по многим другим аспектам.

На сайте издательства размещены разнообразные ресурсы для авторов и редакторов, в том числе перечень репозиторий⁷ и руководство по оформлению сообщения о доступности данных (*data availability statement*)⁸. При внедрении политики большинство журналов адаптируют общие положения под свои нужды. Как было обнаружено, авторы, а не издательства,

⁶ Полный список журналов и типов политики, который они используют, доступен по данной ссылке (<http://www.springernature.com/gp/group/data-policy/journal-listing>).

⁷ Репозитории, рекомендуемые издательством Springer (<http://www.springernature.com/gp/group/data-policy/repositories>).

⁸ Руководство по оформлению сообщения о доступности данных (Data Availability Statement) (<http://www.springernature.com/gp/group/data-policy/data-availability-statements>).

⁵ Служба поддержки по вопросам исследовательских данных Springer Nature (<http://www.springernature.com/gp/group/data-policy/helpdesk>).

более склонны выполнять требования журнала о предоставлении доступа к их исследовательским данным [2].

В настоящее время планируется продолжать исследования для выявления издержек и преимуществ внедрения политики исследовательских данных на разных уровнях. Например, важно провести оценку дополнительного времени, затрачиваемого редактором на коммуникацию с авторами по предоставлению ими информации о доступе к исследовательским данным⁹. Такие сведения важны и редакторам научных журналов, и другим участникам процесса науч-

ной коммуникации для принятия взвешенных решений.

Чтобы создать условия для стандартизации и гармонизации политик разных учреждений, финансирующих организаций, репозиториев, научных сообществ и других издательств, мы опубликовали разработанную рамочную концепцию в открытом доступе по лицензии *Creative Commons* [9]. Однако требуется более широкое обсуждение данной концептуальной основы среди всех заинтересованных лиц. Мы планируем осуществить такое обсуждение в сотрудничестве с «Альянсом исследовательских данных» (*Research Data Alliance*, RDA)¹⁰.

Перевод Н. Г. Поповой

⁹ Предварительные данные (http://www.stmassoc.org/2016_12_06_Digital_Publishing_Hrynaskiewicz-Research_data.pdf).

¹⁰ См. <https://www.rd-alliance.org/groups/data-policy-standardisation-and-implementation>.

СПИСОК ЛИТЕРАТУРЫ

1. Meadows A. To share or not to share? That is the (research data) question... *The Scholarly Kitchen*. Nov. 11, 2014. Available at: <https://scholarlykitchen.sspnet.org/2014/11/11/to-share-or-not-to-share-that-is-the-research-data-question/>
2. Schmidt B., Gemeinholzer B., Treloar A. Open data in global environmental research: The Belmont Forum's open data survey. *PloS One*. 2016;11(1):e0146695. DOI: [10.1371/journal.pone.0146695](https://doi.org/10.1371/journal.pone.0146695).
3. Hahnel M. Global funders who require data archiving as a condition of grants. *Figshare*. 2015. DOI: [10.6084/m9.figshare.1281141.v1](https://doi.org/10.6084/m9.figshare.1281141.v1).
4. Naughton L., Kernohan D. Making sense of journal research data policies. *Insights the UKSG Journal*. 2016;29(1):84–89. DOI: [10.1629/uksg.284](https://doi.org/10.1629/uksg.284).
5. Piwowar H. A., Vision T. J. Data reuse and the open data citation advantage. *PeerJ*. 2013;1:e175. DOI: [10.7717/peerj.175](https://doi.org/10.7717/peerj.175).
6. HEFCE, Research Councils UK, Universities UK, Wellcome Trust. *Concordat on open research data*. 2016. Available at: <http://www.rcuk.ac.uk/documents/documents/concordatonopenresearchdata-pdf>.
7. Hrynaskiewicz I. *Promoting research data sharing at Springer Nature. OfSchemes and Memes* [Blog]. 2016. Available at: <http://blogs.nature.com/ofschemesandmemes/2016/07/05/promoting-research-datasharing-at-springer-nature>.
8. Announcement: Where are the data? *Nature*. 2016;537(7619):138–138. DOI: [10.1038/537138a](https://doi.org/10.1038/537138a).
9. Over 600 Springer Nature journals commit to new data sharing policies. Available at: <http://www.springernature.com/gb/group/media/pressreleases/over-600-springer-nature-journals-commit-to-new-data-sharingpolicies/11111248>.