

ОТКРЫТАЯ НАУКА

С.И. Паринов

Центральный экономико-математический институт РАН

Аннотация. В статье обсуждаются требования Открытой Науки, а также конфликт этих требований с коммуникационными возможностями традиционной системы научных публикаций. Рассматриваются возможности научных информационных систем для обеспечения научного сообщества средствами коммуникации в духе Открытой Науки. Представлены возможности системы Соционет, в которой реализованы все основные требования Открытой Науки.

Введение

Данная статья продолжает серию моих публикаций, посвященных Открытой Науке. В первой статье из этой серии [6], опубликованной в 2009 г. было высказано мнение «... формируется комплекс условий для появления новой более эффективной организационной формы научно-исследовательской деятельности, которую мы назвали Открытой Наукой». В статье от 2010 года [7] мы представили концепцию виртуальной научной среды, предполагающую существенно более высокую степень открытости и прозрачности в деятельности ученых и научных организаций. С этого времени началось создание виртуальной среды "Открытая Наука", на базе научной информационной системы Соционет (<https://socionet.ru/>). Данная статья, третья по счету в этой серии, представляет краткую историю развития идей Открытой Науки в международном научном сообществе (раздел 1), текущее переосмысление требований к Открытой Науке (раздел 2), обсуждение препятствий для развития Открытой Науки, создаваемых традиционной системой научных журналов (раздел 3). В разделе 4 приведены данные из обзора действующих научных информационных систем. Раздел 5 посвящен системе Соционет и ее текущему вкладу в поддержку средств и методов Открытой Науки. В заключении подведены краткие итоги.

1. Краткая история Открытой Науки

Термин Открытая Наука (Open Science) начал активно использоваться в 2008-2009 годах, как представляется, в результате осмысления научным сообществом важных изменений, происходящих в практике научной работы и, в частности, в научных коммуникациях, под влиянием новых информационных технологий.

Первоначально, это понималось как расширение идей и подходов, продвигаемых движением за программное обеспечение (ПО) с открытым кодом (Open Source Software), на другие аспекты научной деятельности [4]. Основной

контекст Открытой Науки в этом случае представлялся так, как это изображено на Рис. 1.

При этом цели Открытой Науки определялись следующим образом [1]:

- Прозрачность в методологии экспериментов, наблюдений и сбора данных;
- Публичная доступность и повторное использование научных данных;
- Публичная доступность и прозрачность научных коммуникаций;
- Использование онлайн-средств для поддержки научного сотрудничества.

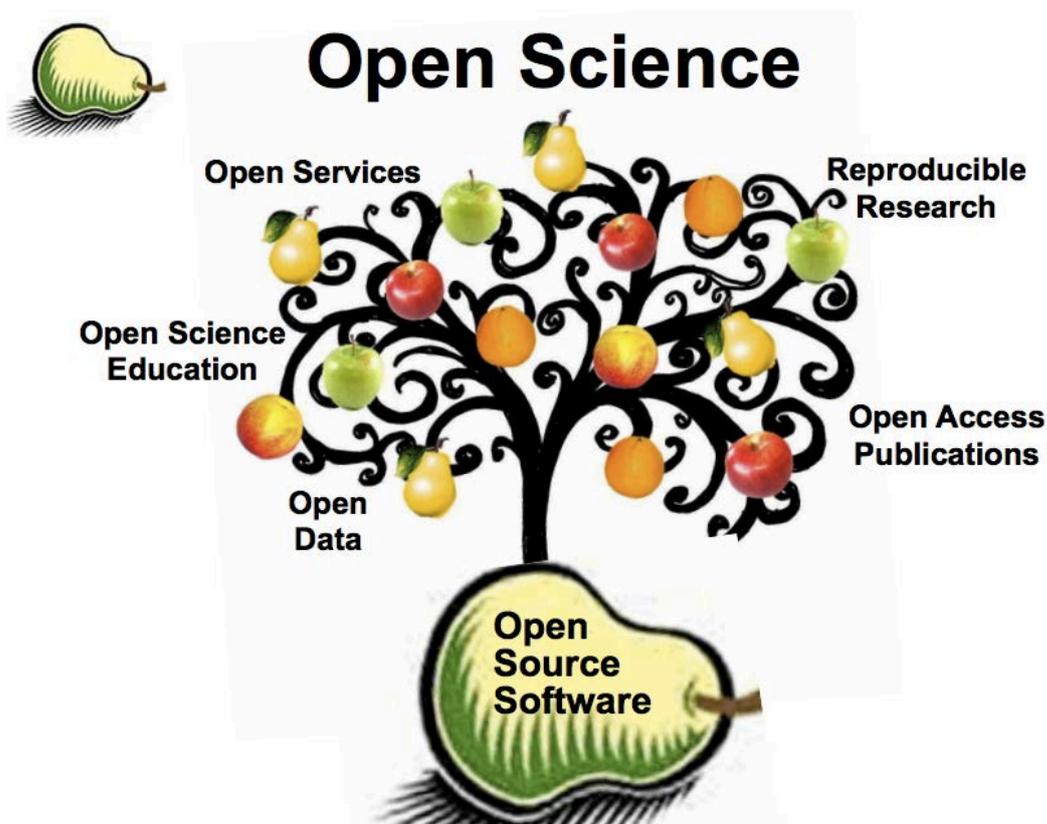


Рис. 1. Открытая Наука как развитие идей движения за ПО с открытым кодом.
Источник - [2]

В последние годы контекст Открытой Науки существенно расширился и детализировался.

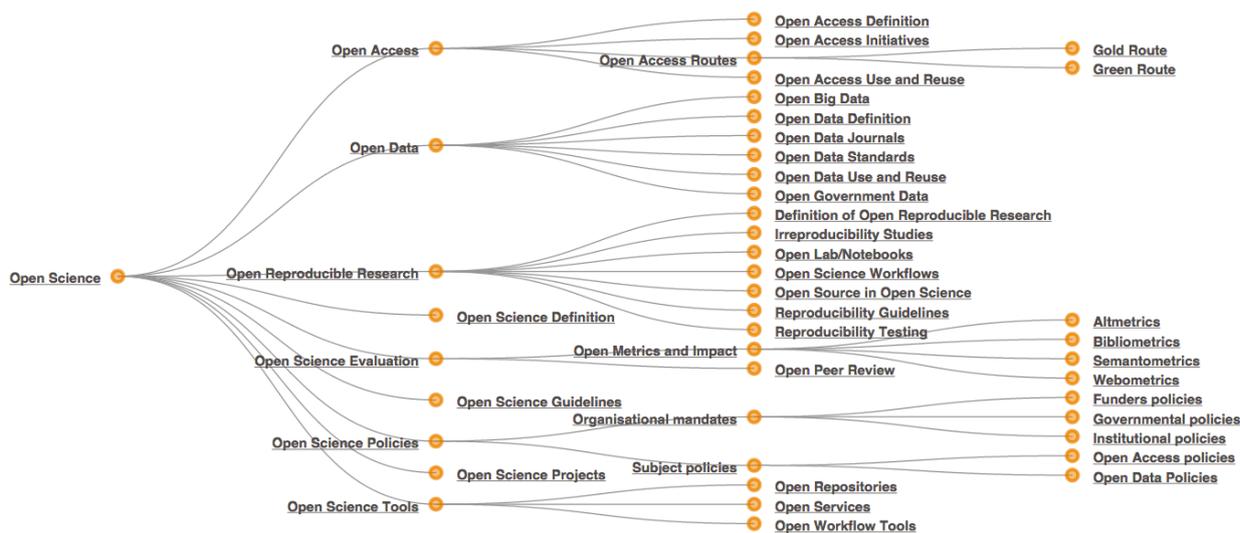


Рис. 2. Дерево понятий Открытой Науки. Источник – проект FOSTER.

Например, проект FOSTER¹, целью которого является создание устойчивых механизмов для европейского научного сообщества по ускорению внедрения открытой науки в повседневную деятельность ученых, Открытую Науку представляет в виде довольно большого дерева понятий (и связанного с ним интерактивного глоссария), изображенного на рис. 2.

Участники проекта FOSTER определяют Открытую Науку следующим образом: проведение научных исследований таким образом, чтобы все желающие могли способствовать или участвовать в нем, при этом создаваемые исследовательские данные, лабораторные заметки и т.п. должны быть свободно доступны для использования, распространения и воспроизведения данного исследования².

Это определение Открытой Науки, опубликованное недавно, видимо в 2014 г., по сути почти не отличается от приведенного выше определения 2009 года. Единственный новый момент – требование открытости процесса исследования для участия в нем посторонних исследователей.

Кроме естественного расширения содержания Открытой Науки, которое иллюстрируют рис. 1-2, усиливается критика традиционной научной системы с позиций Открытой Науки. Примером этого является доклад "Mapping the Future of Scholarly Publishing" [3], опубликованный в феврале 2015 г. международной рабочей группой «Инициатива Открытая Наука»³. Доклад содержит анализ с позиций Открытой Науки недостатков действующей системы научных публикаций и коммуникаций и предлагает некоторую программу действий. Среди принципов и направлений совершенствования

¹ <https://www.fosteropenscience.eu/>

² <https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>

³ <http://nationalscience.org/portfolio/open-scholarship-initiative/>

системы научных публикаций и коммуникаций, в данном документе отмечаются следующие [3, стр. 5-6]:

- Научные работы должны быть свободно доступны, как только автор чувствует, что они готовы;
- Оценка работы должна начинаться с момента, когда она становится доступной, и продолжается пока это имеет смысл;
- Процесс оценки должен быть полностью открытым и каждый должен иметь возможность в этом участвовать;
- Оценка работы должна быть комплексной и включать такие факторы, как обоснованность, важность, направленность, этика и многие другие;
- Финансовая поддержка системы научных коммуникаций должна осуществляться организациями с гарантированной устойчивостью;
- Форма публикации и инфраструктура должны быть оптимизированы для поддержки разнообразного использования опубликованных работ

Эти формулировки полностью соответствуют идеям прозрачности и открытости научных исследований из ранее приведенных определений Открытой Науки. Однако в данном случае формулировки носят вид рекомендаций для ученых и научного сообщества по изменению традиционной формы научной деятельности.

2. Требования Открытой Науки

Для уточнения деталей различных видов научной открытости, упоминаемых в трех выше приведенных источниках, рассмотрим их в контексте следующих укрупненных компонент научного процесса: (1) доступ ученого к существующему научному знанию; (2) оценка научной значимости и анализ применимости доступных ученому результатов исследований по отношению к своей работе; (3) использование ученым отобранных результатов исследований для производства нового научного знания; (4) коммуникации ученого с «соседями» по цепочкам глобальной научной кооперации; (5) оценка ученым своего текущего места в научной системе и возможных направлений для его улучшения.

Применительно к каждой из этих 5 компонент научного процесса требования Открытой Науки могут быть сформулированы в виде следующих требований:

1. Открытый доступ к результатам исследований. Это базовое и принципиальное требование, без выполнения которого не возможно все остальное. Содержание его достаточно хорошо известно, т.к. движение за Открытый доступ в науке имеет уже достаточно длинную историю (более 15 лет). Подборка публикаций на эту тему приведена в списке литературы [3, стр. 44].

2. Открытая научная оценка результатов исследований. Также достаточно известное требование, которое означает, что встретив публикацию заслуживающую внимания, ученый должен дать ей профессиональную оценку. Это требование реализовано в некоторых научных информационных системах в виде публичного (открытого) рецензирования⁴. Оценка, как правило, делается на основе заданного классификатора. Авторы оцененной публикации должны получать имейл-уведомление и система должна давать им возможности отреагировать на содержание оценки. Сделанные оценки, а также реакции на них авторов оцененных публикаций должны быть общедоступны на странице описания соответствующей публикации. Оценки также могут отображаться в данных о научной репутации как ученого, который их сделал, так и ученого, чьи работы были оценены.

3. Открытое использование результатов исследований. Новое требование. Технические возможности и примеры его реализации появились сравнительно недавно, в том числе, с развитием техники семантических связей в научных информационных системах [8]. Это требование означает, что ученый должен указать как или для чего им были использованы процитированные источники из списков литературы его публикаций. Выбор вариантов использования - на основе заданного классификатора. В комментариях ученый может дать пояснения. Система должна уведомлять авторов о всех случаях использования их результатов. В ответ авторы использованных результатов могут дать ценные советы и/или доработать свои результаты, чтобы пользователи этих результатов получили больший эффект. Подобные коммуникации между учеными должны содействовать возникновению долгосрочной научной кооперации между ними.

4. Открытые научные коммуникации. Требование не новое, но в нем появились новые моменты, вытекающие из предыдущих требований. Если ученый использует результат другого ученого, то он должен быть готов к ответной реакции в вариациях от протеста до предложения совместного развития идей. Содержание подобных коммуникаций - публично и доступно читателям публикаций, а также в профиле соответствующих авторов. Публичность содержания таких коммуникаций создает условия для конкуренции между учеными за право предоставить лучший результат для использования. Кроме этого система накапливает данные о подобных коммуникациях и обновляет статистические распределения показателей о том, как ученый реагирует на использование его результатов исследований, как другие ученые реагируют на использование их результатов исследований данным ученым, как ученый реагирует на реакции других ученых и как другие ученые реагируют на его реакции.

5. Открытые данные о научной результативности и научной репутации ученого. Предыдущие требования создают некоторые новые

⁴ <https://peerj.com/>, <http://cyberleninka.ru/about>, <https://socionet.ru/> и др.

моменты для реализации этого вида открытости. Для каждого автора публикаций в научной информационной системе автоматически накапливается статистика, суммирующая различные характеристики его публикаций. Эта статистика включает: общее текущее количество публикаций, суммарное количество распознанных в системе цитирований публикаций автора, статистические распределения характера использования публикаций автора и т.п. Эти данные регулярно обновляются и характеризуют текущую результативность ученого не только с точки зрения выполненного им объема работ, но и демонстрируют, как научное сообщество использовало его результаты. В системе также автоматически собираются данные о научных оценках публикаций автора, о его реакции на эти оценки, а также о его оценках чужих публикаций и реакциях других ученых на оценки данного автора. Статистические распределения, построенные на основе этих данных, дают представление о том, какие категории оценок доминируют в настоящее время у определенного ученого по отношению к чужим работам и какова структура оценок его работ другими учеными. Авторы и все пользователи системы видят эти ежедневно обновляемые данные о своей научной репутации и могут на основе этого корректировать свои действия. Также они могут учитывать данные о научной репутации других ученых, например, если собираются с ними сотрудничать, или реагировать на их оценки по поводу своих публикаций.

В случае реализации этих 5 принципов открытости их совместное действие должно существенно изменить научный процесс. Эти изменения обещают разнообразные улучшения в работе научной системы. Однако на пути реализации этих требований стоят достаточно серьезные препятствия.

3. Традиционная система научных публикаций против Открытой Науки

Уже упоминавшийся выше доклад "Mapping the Future of Scholarly Publishing" [3] также содержит достаточно подробный анализ недостатков традиционной системы научных публикаций. Один из выводов, который можно сделать из этого анализа - неспособность традиционной системы научных публикаций и коммуникаций, основанной, в первую очередь, на глобальной системе научных журналов, соответствовать требованиям Открытой Науки.

Проанализируем как выглядит процесс глобальных коммуникаций между авторами и пользователями опубликованных результатов исследований в том виде, как его поддерживает традиционная система научных журналов (рис. 3):

ученые получают доступ к журналам → читают статьи → выделяют в этих статьях важные для них фрагменты (научные артефакты) → формируют новое научное знание через создание как собственных новых артефактов, так и научных отношений между всеми связанными артефактами → превращают набор артефактов и отношений между ними в научную статью с цитатами, списком литературы и др. → добиваются публикации статей в журналах и полагаются на издателей в том, что потенциальные пользователи их

результатов исследований будут иметь доступ к этим журналам. Далее цикл повторяется.



Рис. 3. Научные коммуникации на основе традиционной издательской инфраструктуры

Про недостатки этого механизма коммуникации уже много написано (см. например, список литературы в [3, стр. 44]). Отметим еще некоторые новые важные моменты:

- потенциальные пользователи результатов исследований обычно сегментированы по доступу к системам агрегации журналов и автор, как правило, не может обеспечить всем им равный и немедленный доступ к его опубликованным результатам;
- при создании нового научного знания ученые мысленно оперируют научными артефактами и отношениями между ними, но для передачи этого знания другим ученым им приходится тратить силы и время на придание научному выводу формы статьи;
- при использовании (цитировании) результатов исследований действующая система не гарантирует что автор использованного результата узнает о всех случаях как, с какой целью и кем этот результат был использован, и тем более она не обеспечивает коммуникации между автором и пользователем, которые, если бы они были, могли бы улучшить качество результатов исследований обеих сторон (автора и пользователя), а также превратиться в продуктивную научную кооперацию данных ученых;
- реальное научное значение и влияние опубликованных результатов выясняется по итогам процесса «проб и ошибок», который возникает, когда ученые пытаются использовать эти результаты в своих исследованиях; качественные итоги этого процесса «проб и ошибок» не могут быть выражены просто наличием цитаты на соответствующий результат (они отсутствуют в индексе цитирования) и никак не связаны с импакт-фактором журнала, где этот результат опубликован.

С точки зрения современного уровня развития информационных технологий традиционная система научных журналов как способ глобальной коммуникации между авторами научных результатов и их пользователями выглядит устаревшей и неэффективной. Она и не может быть модернизирована в соответствии с требованиями Открытой Науки, т.к. эти недостатки заложены в ее конструкции.

Более эффективная система научных коммуникаций, создающая условия для применения принципов Открытой Науки, должна быть создана за рамками традиционной системы научных издательств.

В качестве одного из способов решений этой проблемы в упомянутом докладе [3] обосновывается необходимость создания единого научного репозитория (All-Scholarship Repository) для размещения исследователями своих научных выходов [3, стр. 35].

На наш взгляд, этого не достаточно. Технология открытых репозиториях достаточно хорошо реализует только первый из требований Открытой Науки: она обеспечивает открытый доступ к результатам исследований. Для реализации остальных принципов требуется более сложные информационные технологии. Здесь требуется научная информационная система или взаимосвязанный комплекс таких систем, которые предоставят ученым возможности создавать научные произведения, оперируя научными артефактами и отношениями между ними, а также будут включать сервисы, реализующие открытые коммуникации, аккумулирующие необходимую статистику и т.п.

4. Научные информационные системы

Текущая ситуация характеризуется большим количеством и разнообразием уже существующих научных информационных систем, в которых реализованы те или иные требования Открытой Науки. Хорошее представление об этом дает обзор «101 инновация в научных коммуникациях»⁵.

Данный обзор, к сожалению, не охватывает российские научные информационные системы⁶. Все перечисленные в нем научные информационные системы разбиты на следующие категории (рис. 4):

- Процесс научного открытия (Discovery)
- Научный анализ (Analysis)
- Написание научного произведения (Writing)
- Публикация научного произведения (Publication)
- Пропаганда и популяризация научных результатов (Outreach)
- Оценка научных результатов (Assessment)

⁵ 101 Innovations in Scholarly Communication, <https://innoscholcomm.silk.co/>

⁶ Авторы анонсировали новую версию данного обзора, в котором уже более 400 систем - <http://bit.ly/innoscholcomm-list>

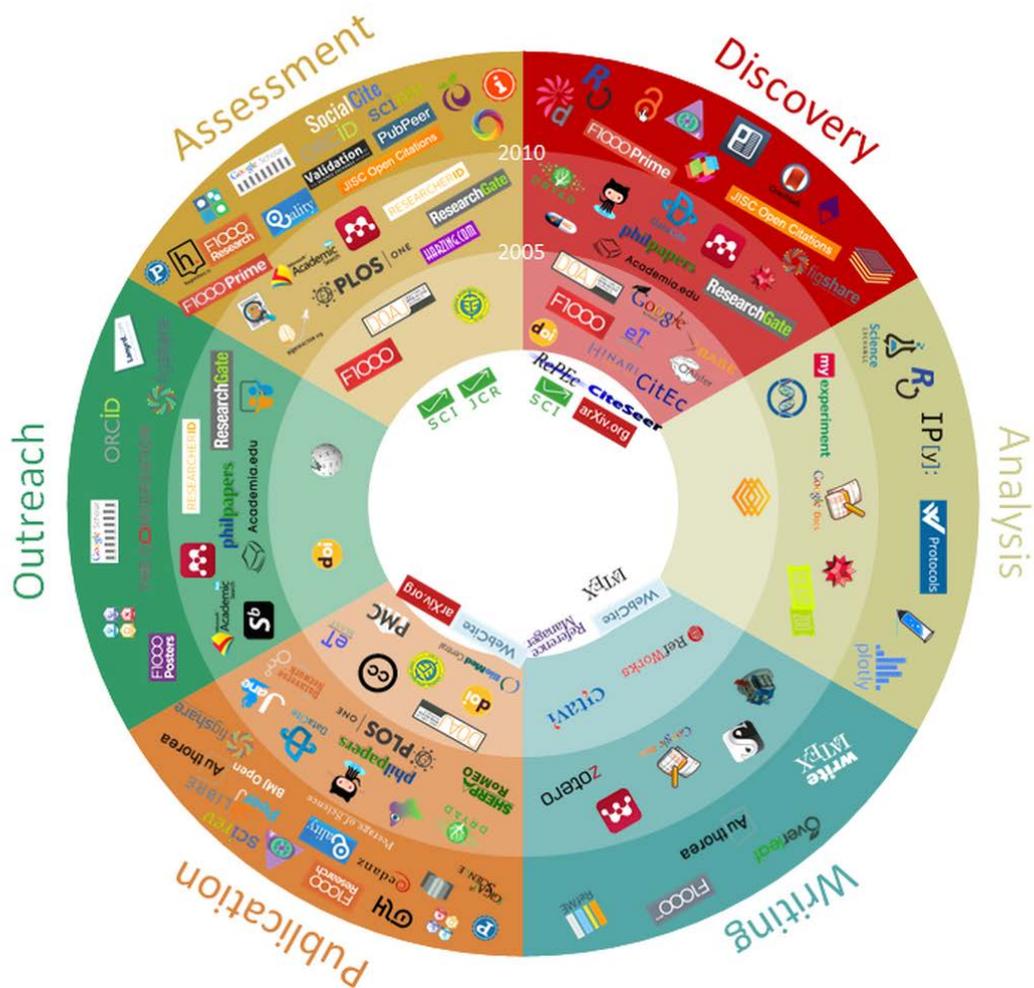


Рис. 4. 101 инновация в научных коммуникациях. Источник - <https://innoscholcomm.silk.co/>

Отнесение информационной системы к определенной категории сделано с определенной долей условности: авторы обзора сами определяли основную функциональность каждой системы.

Использованная здесь структура категорий отражает стадии научного процесса. Однако ее нельзя прямо соотнести с компонентами научного процесса, для которых мы выше рассматривали требования открытости.

По нашему мнению, ни одна из перечисленных в этом обзоре научных информационных систем не реализует все 5 требований Открытой Науки. Однако среди них можно выделить несколько наиболее популярных общедоступных систем: Academia.edu (<http://www.academia.edu/>), F1000 Research (<http://f1000.com/research>), Mendeley (<http://www.mendeley.com/>) и др.

К сожалению, в данный обзор не попала наша собственная система Соционет, которая реализует полный набор перечисленных требований открытости.

5. Научная информационная система Соционет

По состоянию на май 2015 г. общедоступная научная информационная система Соционет (<https://socionet.ru/>) - единственная, предлагающая пользователям функциональность, основанную на реализации всех 5 требований Открытой Науки:

1. авторы могут без ограничений самоархивировать свои научные выходы и выполнять таким образом требование открытого доступа к результатам исследований;
2. ученые могут выражать свое профессиональное мнение о доступных в Соционет материалах, заслуживающих внимание;
3. авторы могут для всех своих публикаций указать, в каком качестве ими были использованы (процитированы) источники из списка литературы;
4. авторы могут получать имейл-уведомления о всех случаях создания отношений (профессиональное мнение или использование) с их публикациями и публично реагировать на эти действия, при этом система Соционет информирует читателей о содержании этих открытых коммуникаций;
5. система Соционет собирает разнообразную статистику о действиях ученых и генерирует статистический портрет для каждого из них, которые включают данные о текущей научной результативности и научной репутации ученого.

Принципы организации и функционирования системы Соционет достаточно подробно описаны в публикации [5].

Одна из стратегических задач развития в Соционет технологий Открытой Науки – создание прототипа глобальной научной информационной системы для решения проблемы, создаваемой существующей научной издательской системой, которая как посредник между авторами результатов исследований и их пользователями плохо приспособлена для эффективных научных коммуникаций и формирования цепочек глобальной научной кооперации между исследователями, использующими результаты друг друга.

В Соционет созданы новые средства работы ученого, способствующие возникновению между авторами и пользователями результатов исследований прямых коммуникаций, сотрудничества и их более высокой ответственности за свои действия.

Практическое применение созданных в Соционет прямых средств сведения друг с другом «авторов» и «пользователей» результатов исследований в условиях конкуренции участников процесса научной кооперации создает ситуацию, при которой чем быстрее и с меньшими затратами найдут друг друга участники глобальной научной кооперации, тем эффективнее работает научная система.

Этот механизм функционирует следующим образом:

- Автор получает сигналы, когда кто-то использовал его результаты исследований;
- Автор может оказать помощь в правильном использовании его результатов исследований (кооперация);
- Автор может доработать свои результаты, чтобы усилить эффект от их использования (кооперация);
- Коммуникации между автором и пользователем формализованы и публичны (конкуренция);
- Все данные накапливаются и формируют публичный портрет ученого, включая его результативность и научную репутацию;

В этих условиях ответственность ученого за свои действия существенно повышается.

Заключение

Одна из наиболее обсуждаемых проблем Открытого доступа и, как следствие, Открытой Науки – ослабление «входного» контроля качества публикаций, который выполняет рецензирование рукописей редколлегией журналов. Распространение открытого доступа путем самоархивирования учеными результатов своих исследований часто не предполагает рецензирование публикаций, а, значит, отсутствует их проверка на соответствие научным стандартам качества. Известно, что многие ученые не хотят тратить время и силы на чтение статей, если они заранее не знают уровень их качества. Разработанный в Соционет комплекс средств в определенной степени решает и эту проблему. Реализованные требования открытости порождают более жесткий самоконтроль и более высокую ответственность ученого за качество публикуемых материалов, а также за использование и оценку чужих результатов исследований и за свое поведение в профессиональных коммуникациях. С другой стороны, на страницах публикаций читателю доступны данные о том, кто и как их использовал, а также профессиональные мнения о них.

ЛИТЕРАТУРА

1. Gezelter, D. What, exactly, is Open Science? Posted on July 28, 2009 by Dan Gezelter - <http://www.openscience.org/blog/?p=269>
2. Harris, N. Welcome to BOSCO 2013. 14th Annual Bioinformatics Open Source Conference. July 19-20, 2013 Berlin, Germany - http://www.open-bio.org/bosc2013/day1/BOSC2013_Intro_Nomi_Harris.pdf
3. Open Science Initiative Working Group. Mapping the Future of Scholarly Publishing, 1st edition // Seattle: National Science Communication Institute, January 2015 - <http://nationalscience.org/wp-content/uploads/2015/02/OSI-report-Feb-2015.pdf>

4. Woelfle, M., Olliaro, P., Todd, M.H. Open science is a research accelerator // Nature Chemistry 3, 745–748 (2011).
<http://www.nature.com/nchem/journal/v3/n10/full/nchem.1149.html>
5. Parinov S., Kogalovsky M., Lyapunov V. A Challenge of Research Outputs in GL Circuit: From Open Access to Open Use // The Gray Journal (TGJ). An International Journal on Grey Literature. v. 10, i. 2, pp. 87-94, 2014 -
<https://socionet.ru/publication.xml?h=RePEc:rus:isyigw:article-448>
6. Паринов С.И. Развитие электронных библиотек – путь к Открытой Науке // Труды 11й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2009, Петрозаводск, Россия, 2009. -
<https://socionet.ru/publication.xml?h=RePEc:rus:mqijxk:21>
7. Паринов С.И. Концепция виртуальной научной среды "Открытая Наука" // Научный сервис в сети Интернет, суперкомпьютерные центры и задачи: труды Международной суперкомпьютерной конференции (г. Новороссийск, 20-25 сентября 2010 г.). — М.: Изд-во МГУ, 2010. — С. 473-480. — URL:
<https://socionet.ru/publication.xml?h=RePEc:rus:mqijxk:24>
8. Паринов С.И., Когаловский М.Р. Технология семантического структурирования контента научных электронных библиотек // Труды 13й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2011, Воронеж, Россия, 2011 -
<https://socionet.ru/publication.xml?h=RePEc:rus:mqijxk:28>