

**СОДЕРЖАНИЕ**

<b>Бадд Д. М.</b> Неразбериха в научной коммуникации: может ли помочь экономический анализ?	3
<b>Телволл М.</b> Оценка влиятельности журналов: вебометрическая перспектива	12
<b>Йёрг Б., Хёлригл Т., Сисилия М.-А.</b> Сущности и идентификаторы в исследовательских информационных системах	21
<b>Сенгупта Ш.</b> Хранилища открытого доступа: сценарий для стран Азии с особым акцентом на библиотековедение и информатику	27

## РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Академик РАН **Ю.М. Арский** (Российская Федерация) — *главный редактор*, ВИНТИ РАН, 125190, Москва, ул. Усиевича, 20. Телекс 411249

Проф. д-р. **Р.С. Гиляревский** (Российская Федерация) — *заместитель главного редактора*, ВИНТИ РАН, 125190, Москва, ул. Усиевича, 20. Телекс 411249

**С. Дж. Паркер** (Канада) — *заместитель главного редактора*, IDCR, P.O. Box 8500, Ottawa, Ontario K1G 3H9, Canada

**А. Джилкрайст** (Великобритания) — CURA Consortium and GAVEL g.e.i.e, 38 Ship Street, Brighton BN1 1AB, UK

**М. Дрейк** (США) — Технологический институт шт. Джорджия, Библиотечный и информационный центр, 704 Cherry Street, Atlanta, Georgia 30332-0900, USA

**Н. Дюсулье** (Франция) — 192 Chemin du Cannet, 06220 Vallauris, France

**А. де Кемп** (Германия) — Издательство “Springer-Verlag”, Postfach 10 52 80, D-69042 Heidelberg, Germany

Д-р **Т. Кеннон** (Великобритания) — Отдел исследований и разработок Британской библиотеки, 2 Sheraton Street, London W1V 4BH, UK

**М. Мидадтон** (Австралия) — Школа информационных систем, QUT Gardens Point Campus, 2 George Street, Brisbane, 4000 QLD., Australia

**Т. Молвиг** (Норвегия) — Национальное управление по научной информации, вузовским и специальным библиотекам, P.O. Box 2439 Solli, N-0201, Oslo, Norway

**Х. Ринкон Феррейра** (Бразилия) — Бразильский институт информации по науке и технике (IBICT), SAS— Quadra 5, Lote 06, Bloco H, 700-70-000 Brasilia D.F., Brazil

**С. Феррейро** (Чили) — Чилийский университет, Системы информационных и библиотечных служб, Casilla de Correo 10D, Santiago, Chile

Проф. **Ю. Фудзивара** (Япония) — Университет Цукуба, Институт электроники и информатики, Tsukuba-shu, Ibaraki, 305 Japan

Д-р **М. Хименес** (Испания) — Испанское общество по научной документации и информации, Fuencarral, 123-6° dcha., 28010, Madrid, Spain

Д-р **А.И. Чёрный** (Российская Федерация) — ВИНТИ РАН, 125190, Москва, ул. Усиевича, 20. Телекс 411249

# Неразбериха в научной коммуникации: МОЖЕТ ЛИ ПОМОЧЬ ЭКОНОМИЧЕСКИЙ АНАЛИЗ?\*

Джон М. БАДД  
(John M. BUDD)

Школа информатики и образовательных технологий, Университет шт. Миссури, г. Колумбия, США

*Вводится попытка осуществления подхода, основанного на теории игр и принятия решений, который стремится рассмотреть элементы сложности научной коммуникации как экономического предприятия. Как индивидуальные, так и институциональные виды игр анализируются с целью определить, какие факторы будут влиять на реальное экономическое использование теорий игр и принятия решений. Существуют взаимосвязи между двумя видами, которые добавляют сложность к любому возможному применению. Более того, этот анализ включает идеальные и практические факторы, которые оказывают влияние на реальное экономическое применение. Как показывается здесь, имеются серьезные проблемы применения этих теорий, но также есть важные показатели для продвижения индивидуальных и институциональных интересов посредством обсуждения.*

## ВВЕДЕНИЕ

Новаторская работа лауреата Нобелевской премии Джона Нэша была сделана в новой (к тому времени) области теории игр. Обычно считается, что фон Нейман и Моргенштерн [1] «изобрели» теорию игр за несколько лет до появления удостоенной премии работы Нэша. Нэш вывел понятие баланса, который может встречаться, когда игроки представлены в затруднительном положении, в котором они могут общаться и сотрудничать, и когда они рациональны и знакомы с игровой ситуацией. Каждый игрок может выбрать сотрудничество или отказ от сотрудничества в процессе игры. Иногда не сотрудничающие игроки склонны достичь безвыходного положения, где нет никаких выгод.

Вид игры, продемонстрированный *Дилеммой заключенного*, предполагает *рациональность* со стороны игроков и является показателем классической теории игр. Эти игроки будут делать свой выбор на основе наиболее вероятных выгод. Желательным результатом в классической теории игр является, как определил Нэш [2], баланс. Объяснение Нэша является математически сложным; оно потребовало от экономистов совсем немного времени, чтобы включить его в свои теории. Герберт Гинтис [3] четко формулирует баланс Нэша: «при игре с двумя игроками имеется пара стратегий, каждая из которых является *лучшим ответом* на другую; т.е. каждая дает использующему ее игроку наивысший возможный результат, учитывая

стратегии другого игрока» [3, с. 34]. Однако в сложной игре может быть много возможных способов достичь баланса, отчасти потому, что каждый дополнительный игрок приносит несколько возможных результатов. Когда экономисты стремятся применить теорию игр, они заботятся о расположении решений к играм, и поэтому выгоды очерчиваются заметно. Игроки ищут получения оптимальной выгоды от процесса игры. Утилитаризм является проблематичной философской позицией и, как предполагает Роберт Сагден [4], мера выгодности вводит концептуальную и практическую трудность в теорию, так как она может оцениваться в ряде способов даже в рамках одной игры. Появляется другая трудность, также определенная Сагденом; классическая теория игр была разработана скорее как концептуальный, а не эмпирический подход. Преимущество этой теории заключается в ее фундаментальной применимости.

Анализ, представленный в этой статье, будет состоять в рассмотрении научной коммуникации как типа игры, включающей особенности согласования, заключения сделки и сотрудничества. Полезность в этом анализе не игнорируется, однако больше внимания уделяется справедливости и эффективности, чем результативности. В этом исследовании используется потенциальный аналитический инструмент для изучения сложной экономической и свойственной человеку системы научной коммуникации. Родственный инструмент теории решений (о котором писали Бинмор и другие теоретики игр) также будет применен. Поскольку игровой аспект этого вида коммуникации включает смешение индивидуальных и институциональных игроков, оба типа игроков будут включены, рассмотрены и, в частности, будет изучаться

\* Перевод Budd J.M. Scholarly communication's mess: Can economic analysis help? — 2012. — <http://libres.curtin.edu.au/libres22n1/index.cfm>

вопрос о том, что характеризует их отношения в наше время. В частности, реальные и практические элементы научной коммуникации будут помещены в центр этого исследования, с тем чтобы гипотетические применения могли быть превосходящими в попытке поиска решений экономической дилеммы коммуникации. По необходимости система научной коммуникации должна быть определена в качестве экономической, чтобы можно было провести подобный анализ. Типы и характеры игр индивидуумов и учреждения, также должны быть объяснены, чтобы применяемые теории имели смысл. «Неразбериха», которая обсуждается в этой статье, происходит из-за сложности оценки материальных и нематериальных затрат и выгод и некоторых ошибок, которые могут быть допущены, если не делать различий между этими двумя типами игр. В конечном счете, в этой статье выясняется возможная эффективность теории игр и принятия решений как научного метода, примененного к проблеме экономики научной коммуникации. Теория игр безусловно включает некоторую сложную математику, но имеется аспект этой теории, который является даже более важным для применения в некоторых установках – справедливость (или объективность). Другими словами, теория игр сегодня может значить больше, чем узкий экономический инструмент, который применяется только к чисто рациональным и утилитарным ситуациям (и здесь появляются некоторые утилитарные проблемы). В частности, может быть задан следующий вопрос: что подразумевается под наиболее эффективно гарантированным сотрудничеством, чтобы избежать конфликта между игроками и дать их результаты? Одним способом достижения желаемого результата (в кооперативных играх) является разработка способа, позволяющего всем игрокам «победить». Здесь необходимо подчеркнуть, что «победа» не означает проигрыш оппонента, а работу с другими участниками, чтобы можно было реализовать оптимальную взаимную выгоду. Короче говоря, «дилемма» можно избежать, если игра характеризуется сотрудничеством. Здесь нет достаточного пространства, чтобы тщательно исследовать все моральные причины применения теории игр, но статья Кена Бинмора [5], хотя и проблематичная в некоторых моментах (обсуждаемых ниже), будет основой для последующего исследования. Для Бинмора Дэвид Хьюм является героем теории игр, так как Хьюм ясно выразил необходимое взаимодействие, которое должно быть в равновесии в центре любого общества. Хьюм возможно привлек внимание Бинмора, поскольку моральная философия Хьюма является попыткой признать баланс в человеческой природе между благотворительностью и эгоистичной жадностью. Бинмор [6] выступает за рациональное изучение возможностей, которые возникают в ситуациях, где взаимодействуют игроки (несмотря на то, имеют ли игроки полную или неполную информацию), но он отклонил метафизическую рациональность Канта в пользу Хьюма. Позиция Бинмора должна быть тщательно изучена, и позже будет отражена в этой статье. Имеется немного применений теории игр в библиотечном деле и информатике и всего несколько применений в теории принятия решений. Те применения, которые существуют, в первую очередь касаются разработки систем [7-8] или интеллектуального управления [9]. В литературе по высшему образованию упоминается также ряд применений, которые существуют и имеют тенденцию быть присутствующими области экономического образования [10].

## ЭКОНОМИКА ПУБЛИКАЦИИ

Вероятно первое, что необходимо отметить относительно научной коммуникации как экономического механизма, – это ее «специфика» [11]. Библиотекари стремятся быть комплектаторами информационных материалов

в традиционном смысле подписки на печатные издания и лицензионных ресурсов, это сообщество включает университетских преподавателей и других ученых как производителей и потребителей, и их организация оплачивает расходы. Добавим к этому вторичных производителей информационных продуктов, издателей и распространителей (особенно в той мере, в какой распространители добавляют стоимость через индексирование и т.п.). В настоящее время и почти на протяжении всей истории научной коммуникации вторичные производители влияют на негибкость цен. Предполагается, что эти производители добавляют стоимость к материалам в форме письменных сообщений, индексирования и реферирования и другим действиям. Это предположение сегодня ставится под сомнение, особенно создателями ресурсов открытого доступа и цифровых институциональных хранилищ. Роль вторичных производителей сопровождается примерами негибкости цен. Отсутствие гибкости является следствием особого монополистического характера информационных сообщений. В предметной области может быть больше одного журнала, скажем, в области физики плазмы, но каждая статья в литературе является в значительной степени уникальной. Если университет, его профессорско-преподавательский состав и студенты, а также его библиотека хотят иметь доступ к основным работам в области физики плазмы, они должны платить за большое число документов. Они делают это путем подписки на многочисленные журналы и/или на лицензионный доступ к одной или более базам данных, которые будут предоставлять полный текст статей.

Если любая игра должна управляться решениями в наилучших интересах игроков, тогда эти решения должны быть рациональными. Имеется значительная литература, которая это исследует, хотя многие моменты повторяются во всем массиве такой литературы. Например, в сложных играх, как наблюдают Харт и Мансур [12], «предполагается, что каждый игрок вначале знает только *свою собственную выигрышную функцию*, а не других игроков» [12, с.108]. Никто из игроков не является всеведущим, поэтому общая рациональность (в смысле априорного знания возможной выгоды, удовлетворяющей всех игроков) является труднодостижимой. Цели игроков в системе научной коммуникации разнообразны, настолько разнообразны, что меры полезности и выгоды, трудно установить. В частности, трудно формулировать полезность с точки зрения предпочтений для всех игроков. Эта проблема обостряется, так как основная природа принятия решений страдает неуверенностью. Например, с точки зрения потребления, то, что содержание журнала *A* было полезным для исследователя *X* в 2010 г., не означает, что содержание журнала *A* будет полезно исследователю *X* в 2011 г. Положение дел может описываться так: «Пелена неуверенности представляет реальное отсутствие знания о событиях, которые еще должны произойти» [13, с.84]. Другой способ понять различия в оценках зависит от осознания того, что исследователь *X* может оценить полезность в контексте, а контексты 2010 и 2011 гг. могут быть совсем разными (такими как сосредоточение на различных проектах исследования).

Существуют многочисленные условия, свойственные системе научной коммуникации, которые чрезвычайно затрудняют рациональные решения. Хотя существуют многочисленные жалобы, адресованные вторичным производителям из-за назначенных ими цен, вторичные производители работают внутри негибкой экономической структуры и в их собственных интересах. Эти вторичные производители имеют предысторию создания документов способами, которые включают эффектив-

ность, нацеленную на удовлетворение потребностей рынка (включая внутренний процесс рецензирования и комплектование информационных элементов данных в выпуски). У журналов есть редакторы и редакционные коллеги, и они используют экспертов для рецензирования каждой рукописи, с тем чтобы реализовать функцию отбора. Эта функция нацелена на обеспечение уровня качества, подразумевающего гарантии относительно соответствия методологий, важности научных вопросов и проблем и обоснованности полученных результатов. Существуют затраты, связанные с деятельностью по отбору материалов. Престижный научный журнал может получать сотни, возможно даже тысячи, статей ежегодно. Если деятельность журнала требует услуг редактора, работающего на полную ставку, и возможно даже штата редакции, то эти затраты могут быть значительными. Даже если редакторам не компенсируется их труд напрямую, они могут получать непрямые компенсации от своих организаций (в форме работы ассистентом, освобождения от некоторых преподавательских обязанностей или других вещей, которые имеют цену). Если до сих пор существует печатная версия журнала (и имеется много таких журналов к настоящему времени), то налицо производственные расходы и расходы по распространению. Если журнал доступен в электронном формате, существуют затраты, связанные с соответствующей технологией. Ясно, что научный журнал среднего размера может иметь ежегодные затраты в 500 тыс. долл. США или больше.

### ИГРЫ ИНДИВИДУУМОВ

Производственные затраты журнала, как показано выше, зависят среди прочих факторов от числа статей, публикуемых в журнале. Другая группа игроков вступает в игру на этой стадии. Профессорско-преподавательский состав в университетах и колледжах стремится общаться с другими людьми в какой-то области или подобласти. Отчасти они публикуются, чтобы выполнить требования финансирующих агентств (чтобы показать, что финансирование достигает некоторого осязаемого результата). Отчасти они публикуются по их собственной мотивации, которая может включать их желание внести вклад в общественное благо растущего знания. Так как частью мотивации авторов является их оценка третьими сторонами (такими как комитеты по пребыванию в должности и продвижению), то будет использоваться некоторая метрика для измерения относительного (абсолютного) вклада авторов. Либо относительная, либо абсолютная стоимость публикаций, хотя и неполностью определенная, может быть оценена. Определенное число публикаций в идеальном массиве журналов может вносить вклад в пользу благоприятного решения относительно пребывания и продвижения в должности, что имеет общую материальную выгоду для преподавателя. Большое число публикаций в самом маленьком массиве ведущих журналов может привести как к материальной, так и нематериальной (репутационной) выгоде для преподавателя. Стоимость публикаций, тем не менее, не является постоянной в рамках дисциплин, поэтому принятие решений на уровне кампуса является непростым. Другой проблемой решений на основе теории игр и теоретического изучения системы научной коммуникации является установление переменных, связанных с авторством и функцией полезности, на которые авторы могут опереться. Если больше значит лучше (что само является обсуждаемым условием), то может ли быть здесь мера, которую можно использовать в матрице игры или решения? Более сложный вопрос, но служащий показателем затруднения, с которым сталкиваются игроки, состоит в том, может ли

что-то, подобное дереву принятия решений, описывать предполагаемые выборы. Дерево принятия решений может применяться любым игроком для изучения относительных выгод различных альтернатив.

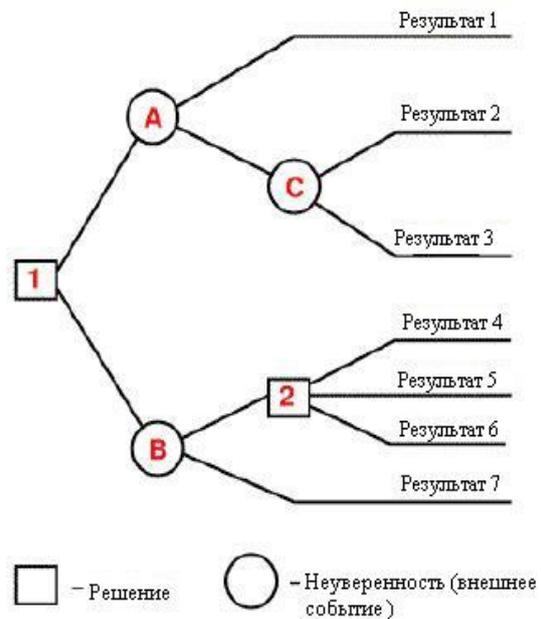


Рис. Дерево решений

Решение может действительно привносить определенного рода неуверенность. Например, решение о лицензировании базы данных может быть введено в «А» в модели; это, как результат, может представлять неуверенность относительно будущего использования источника преподавателями и студентами. «С» может быть неуверенностью, связанной с будущим увеличением цен. Последующим результатом может быть сохранение или отмена лицензии. Некоторое разрешение неуверенности (скажем, в случае «В») может привести к последующему решению, «2». Эта модель, при условии применения как средства предвосхищения возможных переменных, особенно переменных, для которых имеется неуверенность, может помочь принятию решений в библиотеках.

Одной из сложностей, связанной с поведением университетских преподавателей является то, что оно не является естественно или обязательно основанным на принципах сотрудничества и поддержки либо по отношению к издателям или администраторам в своих кампусах, либо друг к другу. Последний из соревновательных факторов становится очевидным, так как признается, что престижный журнал может опубликовать только  $X$  статей за год. Если преподаватель в кампусе публикуется в таком журнале, другой преподаватель в этой области имеет  $X-1$  шанс в нем опубликоваться. Поэтому, до тех пор, пока большее число публикаций улучшает индивидуальный рост и представляет интерес для индивидуума, индивидуум может стремиться публиковаться больше. Остается открытым вопрос, является ли разумной практика максимизации числа публикаций (в более широком философском смысле врожденного логичного действия, основанного на коллективной пользе или благе). Гипотетически не существует высшего предела для публикации индивидуума, хотя практически имеется так много времени, которое любой может посвятить исследованию или написанию авторской работы. Многие индивидуумы стремятся увеличить число статей, опубликованных (при условии, что это является чьей-либо изначальной мотивацией) через соавторство, что определенно явля-

ется своего рода сотрудничеством, но таким, которое может быть основано на максимизации публикаций всех преподавателей, а не на оптимизации возможностей роста знания среди читателей литературы. При помощи соавторства работа может распространяться и может быть осознан больший результат при условии, что больше рук было приложено к работе. Соавторство может также быть тактикой, используемой для максимизации ограниченных (если не недостаточных) ресурсов, включая денежные средства, рабочее пространство, аспирантов и время. Имеются неоспоримые преимущества в улучшении научных основ дисциплин, и существует сложный экономический императив правительств, отдающих предпочтение финансированию совместных работ. Однако соавторство несомненно проблематично по многим причинам, не самой малозначимой из которых является законный интерес со стороны авторов продолжать практику включения многих людей в список авторов, помимо выделения некоторых людей в качестве «почетного» автора. Отход от проблем настоящего исследования (без сомнения важного) является вопросом ответственности. Если ряд имен приведен в списке в качестве авторов статьи, то кто реально ответственен за содержание [14]? Более относящимся к проблемам изучения вопросов теории игр и принятия решений является видимое расширение авторства, которое может создать нарушение в анализе.

Если бы мы должны были создать гипотетический сценарий, то мы могли бы посмотреть на факты, доступные сегодня из индивидуальных практик. Доцент астрономии в главном исследовательском университете имеет 63 публикации ко времени вступления в должность и пересмотра на предмет последующего продвижения. Такое высокое число публикаций, возможно, не является характерным для доцента. Другой доцент (социологии) имеет 19 публикаций, что, возможно, является характерным для продуктивного доцента в этой дисциплине. Согласно базе данных Scopus, профессорско-преподавательский состав Стэнфордского университета (без сомнения главного научного университета) создал 37200 публикаций за период 2005-2009 гг. Доля Стэнфорда в создании публикаций является характерной для научных университетов. Этот университет (согласно его сетевому сайту) имеет 1903 преподавателя, что означает, что средний сотрудник опубликовал 19,5 статей за этот период времени (если сократить подсчеты до упрощенных параметров). Эти цифры дают некоторое представление (хотя неполное и имеющее уклон в сторону большего числа публикаций) об активности преподавателей в научных университетах. Совместные действия ученых со временем приведут к ресурсу, подобному ScienceDirect, имеющему около 9 млн. статей в своей базе данных.

Кроме того, данные указывают на некоторые сложности системы научной коммуникации, исследуемой в качестве игры. Каждый преподаватель с каждым представлением статьи вовлекает журнал в элементы игры, в которой преподаватель «выигрывает», если статья принимается для публикации. Журнал меньше инвестирует в решение относительно типичной представленной статьи (по крайней мере, в плане того, что касается отдельного автора и его статьи) и делает свой выбор на основе будущих рассмотрений. Издатель журнала – и в значительной степени автор, но с потенциально различными мотивациями – хочет, чтобы содержание журнала прочитывалось. Подтверждением того, что статья прочитана, является ссылка на нее в будущей литературе, поэтому критерии, подобные фактору влияния, пристально прослеживаются издателями. С учетом этих критериев сложность становится очевидной; люди не ссылаются

на содержание журнала как таковое, а цитируют статьи, которые оказываются напечатанными в журнале. Таким образом, издатели хотят принимать решения, которые приведут к принятию к публикации статей, которые будут иметь влияние на читателей. Становится ясно, что игра, выраженная таким образом (с потенциальными победами как издателя, так и журнала), не согласуется с обычной теорией игры с участием двух игроков [15].

Например, проект *Science Watch* отслеживает наиболее цитируемые статьи из 20 стран за период с 1998 по 31 августа 2008 гг. (<http://sciencewatch.com/dr/soc/2009/09janALLPAPRS/>). Что касается США, то самой цитируемой статьей была статья Р. Росса «Механизмы болезни» в журнале *New England Journal of Medicine* в 1999 г. Росс выиграл, став автором самой цитируемой статьи за почти десятилетний период (а также получив значение цитирований – более, чем 7000). Журнал также значительно выигрывает от публикации самой цитируемой статьи. Даже если нет необходимости в любом априорном сотрудничестве, Росс и *New England Journal of Medicine* не имели взаимоотношений, основанных на теории игр, относительно его статьи до того, как он ее представил, и даже если бы имел место возможный антагонизм как часть процесса рецензирования, оба игрока в итоге выиграли. Ситуация, которую необходимо проанализировать, если нужно получить выигрыш от сложностей научной коммуникации, все-таки включает рецензентов статьи, которых выбрал издатель. В случае Росса можно предположить, что положительная рецензия на представленную статью была получена издателем. Предположим, что один или более рецензентов были способны определить Росса в качестве автора, даже при помощи двойного анонимного рецензирования. Предположим далее, что один или более рецензентов были соперниками Росса и направили отрицательные рецензии издателю. Журнал *New England Journal of Medicine* мог не опубликовать эту статью из-за влияния трех сторон и мог не получить выгоды, ожидаемой от публикации статьи. Может ли сложность этой версии игры быть представлена посредством символов?

Нет простого ответа на упомянутый вопрос, но некоторые условия для возможного представления могут быть определены. Что касается первого, только что описанная игра (независимо от автора или журнала) может быть описана как игра переносимой полезности, то есть можно утверждать, что существует коалиция между автором и издателем, по крайней мере, на определенный момент времени. Эта коалиция очевидна, когда статья принята для публикации и редактор работает с автором над пересмотром, нацеленным на улучшение статьи. Калаи [16] предполагает, что эта игра переносимой полезности имеет реально значимую функцию (нечто осязаемое, из чего игроки могут получить выгоду и могут получить ее максимально за счет работы в коалиции). Эта функция, в частности, значение взаимодействия ( $v$ ), может быть выражена следующим образом:

$$v=(v(S)), S \in j \quad [16, \text{с.128}].$$

Это выражение предполагает коалицию игроков ( $S$ ) с множеством явно определенных игроков ( $S \in j$ ). Трудность с этим представлением состоит в том, что оно не включает предпубликационные действия, включая рецензирование. Весьма вероятно, что любой журнал имеет в виду, чтобы предпубликационная стадия была также действием коалиции. Редактор может стремиться отбросить беспристрастные рецензии, предназначенные для того, чтобы предложить наиболее информативную поддержку решения издателя. Если эти условия выполнены,

последующее символическое представление может использоваться для отношения на уровне издатель – рецензент. Однако, если один или более рецензентов не работают в структуре коалиции, то баланс Нэша не может быть достигнут. Поскольку отсутствует основной элемент сотрудничества, то игра действительно терпит неудачу. Иными словами, журнал и автор несут убытки, хотя они играли в игру честно. Экономика примера пристрастных рецензентов затрудняет использование теории игр в этой ситуации. Не потому, что имеется неполная информация, доступная игрокам, здесь могут быть конкурирующие интересы, которые препятствуют сотрудничеству. Авторы, как и игроки, могут не обратить внимания на знание, которое утаили рецензенты. Обычные возможности принятия статьи к публикации не достигнуты.

Пока обсуждалась только часть игры, в которую играет преподаватель. Почти все преподаватели также вовлечены в некоторые отношения с своими учреждениями. Учреждения, такие как университеты, подобны журналам в том, что они также имеют некоторый уровень индифферентности, когда дело касается отдельных индивидуумов. Администраторы университетов связаны в первую очередь со своими организациями; они хотят престижа, признания и уважения для своих учреждений. Такие индивидуумы, заботящиеся об этих позитивных факторах, вероятно, должны поддерживаться с помощью утверждения на постоянную должность, продвижения по службе и роста заработной платы. Как в случае с преподавателями, пытающимися опубликовать статьи в журналах, постоянная должность, продвижение и решения о зарплате включают третьи стороны. Существуют комитеты, кафедры и деканаты, которые рассматривают личные дела преподавателей и выносят рекомендации относительно решения этих вопросов. Бескоалиционные среды также могут существовать в этой игре. Как показал Дэвид Хьюм [17] в своей книге *Treatise of Human Nature*, существование описательных утверждений не может иметь отношения к предписывающим или нормативным утверждениям. Существование комитетов, ответственных за беспристрастное и объективное рассмотрение этих достижений, не означает, что последуют беспристрастные и объективные рецензии. Другими словами, нет гарантии равноправного отношения ко всем преподавателям в процессе рассмотрения (как статей, представленных в журнал, так и личных дел, поданных в комитеты). Как будет показано в разделе «Заключение», неудача этого метаэтического императива следует вместе с неудачей утилитаризма, примененного к теории игр.

## УЧРЕЖДЕНЧЕСКИЕ ИГРЫ

В целом экономика университетов и ресурсы, на которые полагаются преподаватели и студенты, является даже более сложной, и, безусловно, более дорогой, чем иллюстрируется отдельными примерами. Например, согласно данным за 2007-2008 гг., представленным Ассоциацией научных библиотек (США), 111 университетских библиотек потратили всего 627 707 869 долл. на сериальные выпуски. Эта ассоциация также сообщает, что учреждения имели в штате 186 759 преподавателей; стоимость сериальных выпусков на преподавателя составляет 3361,06 долл. Эти университеты сообщили, что потратили всего 554 637 844 долл. на действующие электронные ресурсы [18]. Без учета трех университетов, не сообщивших о затратах, стоимость электронных ресурсов на преподавателя составляет 2 969,81 долл. Принимая во внимание расходы на преподавателя в 2007-2008 гг., цифра для библиотечных ресурсов составляет 6 405,14 долл. Безусловно, преподаватели не являются единственными

пользователями (получателями) этих ресурсов, студенты, вероятно, также интенсивно пользуются электронными ресурсами. Этот подсчет необязательно равняется долларовой цене ресурсов, используемых каждым преподавателем, а является очень грубым показателем размеров финансовых ресурсов, которые должны использоваться для обеспечения университетского сообщества чем-то, приближающимся к полному доступу научной информации. «Особенность» этой экономики подразумевается, когда эти данные рассматриваются как часть игры.

Если это специфическое положение дел составляет игру, тогда кто является игроками? Как было показано в играх индивидуумов, двумя игроками являются преподаватели – со своими понятиями полезности и блага – и отдельные университеты, которые имеют только отчасти родственные, но, по существу, различные понятия полезности и блага. Что касается преподавателя (для настоящего рассмотрения мы должны добавить всех преподавателей, а не только тех, кто нацелен на постоянную должность и продвижение), вопросы полезности имеют личностный характер (или более личный, чем учрежденческий). Количество публикаций может привести к дальнейшему карьерному продвижению и соответствующему росту зарплаты. Если только личная выгода является целью, то большее число публикаций может дать преподавателю определенные рычаги в обсуждении позиции в должности в другом университете с более высокой заработной платой. Хаусмен и МакФерсон [19] пишут, «не только большая часть теории игр определяет достоинство и выбор агентов с точки зрения удовлетворения предпочтений, но и анализ социальных взаимодействий теории игр безоговорочно предполагает, что единственными (и поэтому правильными) перспективами для индивидуумов, чтобы адаптироваться в своих социальных взаимодействиях, является индивидуальная максимизация» [19, с. 189]. **Забота о благе может вытекать из желания вносить вклад в совокупность знаний, таким образом продвигая исследование и обучение той дисциплине, которую ведет преподаватель, но теория игр стремится «исключить коллективную перспективу – перспективу, которая рассматривает, что нам следует сделать, и какими будут последствия для нас»**[19, с.196]. **Позиция университета будет обсуждаться ниже.**

Имеются другие игроки в этой версии игры. Хотя преподаватели, студенты и другие люди создают научные знания и результаты исследований, которые публикуются, производители делают эти научные знания и исследования доступными. Эти игроки известны библиотекарям и преподавателям. Они устанавливают цены и определяют, как много статей будет опубликовано в определенном году. Если все эти игроки принимаются за участников игры, быстро становится очевидным, что игра включает некоалиционных игроков, каждый из которых нацелен на индивидуальную максимизацию выгоды. Преподаватели имеют собственные требования и понятия выгоды, которые иногда не учитывают других игроков. Соавторство может основываться на настоящем сотрудничестве, поэтому разделенные мотивации должны учитываться там, где существует такая деятельность. Производители стремятся максимизировать цены путем установления их цен в долях, которые будут нести рынки (библиотеки). Это не является новым, но при просмотре контекста теории игр источники конфликта становятся ясными. Отсутствие сотрудничества очевидно в действующих маржах (прибылях) некоторых ведущих производителей. Например, журналы издательства Elsevier сообщили, что маржа операционной прибыли за 2010 г. составила 2,5% (<http://reports.reedelsevier.com/>

ar10/financial\_review/chief\_financial\_officers\_report.htm). Издательство Springer сообщило о действующей марже за 2010 г., оценив ее в 33,9% до подсчета «амортизации и ослабления нематериальных активов» (<http://www.springer.com/about+springer/company+information/annual+report?SGWID=0-175705-0-0-0>). В 2008 г. издательство Wiley сообщило о прямом вкладе в прибыль своего подразделения по науке, технологии, медицине и образованию, составившем 39,4% ([http://www.wiley.com/legacy/annual\\_reports/ar2009//10kWiley2009.pdf](http://www.wiley.com/legacy/annual_reports/ar2009//10kWiley2009.pdf)). Если эта игра будет совместной, то ресурсы более умеренной стоимости станут доступными для научных библиотек и, в основном, для студентов и ученых, их использующих. Если это действительно так для библиотек, играющих в эту игру, то они должны что-то потерять.

Стоимость некоторых видов ресурсов высока, и цены повышаются ежегодно. Данные Ассоциации научных библиотек (США) показывают, что затраты на серийные издания ежегодно увеличиваются на 7,3% с 1986 по 2008 гг. Ежегодные затраты на монографии выросли за этот период только на 2,9%. Кроме того, число статей, предоставляемых по межбиблиотечному абонементу, выросло в среднем на 6,0% за 1986-2008 гг. [18] (показывая, что был задействован еще один – коалиционный – вид игры среди библиотек, стремящихся обмениваться статьями, которыми владеет каждое учреждение). Производители журналов выигрывают от ряда решений, которые принимают библиотекари (независимо от того, предпочитают ли библиотекари принимать или не принимать такие решения). Вопрос, который может и должен быть задан относительно сокращений, состоит в том, является ли это действие рациональным. Если плата за сокращение информационных ресурсов влияет на некоторых преподавателей, но не на других лиц, тогда может иметь место селективность, которая полностью не руководствуется причиной. Иначе говоря, реакция на то, что выбирает делать один ряд игроков, не является несогласованной; выгода распространяется на одного игрока или ряд игроков, хотя польза и выгода других страдает.

Как утверждают законодательные органы, полное применение теории игр также должно включать других игроков, внешних по отношению к колледжам и университетам. Хотя эти органы могут иметь несколько ограниченное влияние на частные организации, законодательство периодически увеличивает возможность роста преподавательской нагрузки в школах штатов. Если такое движение будет успешным в любом штате, тогда динамика усилий будет нарушена. Решения законодательных органов относительно ассигнований также оказывают сильное влияние на финансирование библиотек, которое, естественно, влияет на доступ к ресурсам и услугам для научных сообществ. Сокращения ассигнований имеют негативное влияние на все аспекты жизни колледжей и университетов. Сложность данных, которые являются доступными, иллюстрирует эмпирические ограничения теории игр в реальных экономических ситуациях. Точно так же, как индивидуумы, публикующиеся в журналах и оцениваемые коллегами, необязательно разделяют цели издателей, поступают и библиотеки. Библиотекари в кампусах колледжа и университета хотят обеспечить доступ к оптимальному числу и самому выгодному контенту информационных статей для сообщества кампуса. Вторичные производители хотят оптимизировать прибыль. Последняя группа стремится осуществить свою цель путем убеждения как можно большего числа библиотек подписаться или приобрести лицензию на их продукты и убедить библиотеки заплатить установленную цену. До сих пор вторичные производители были победителями в этой игре. Эта тенденция совсем необязательно

продолжится в будущем, и теория принятия решений может показать способы, чтобы эта тенденция могла прерваться. Теория игр и принятия решений полагает, что процессы сделок и согласований могли привести к «минимаксному» результату для всех игроков. Согласно Уинклеру [20], критерий минимакса предполагает, что «для каждого действия найдется возможная потеря, и тогда следует выбрать действие, для которого эта крупнейшая потеря будет самой маленькой» [20, с. 210]. Это кажется слишком крученым (замысловатым), но это, по существу, сводит возможную потерю к минимуму. Даже вычисление минимакса является неэффективным и неполным в настоящее время из-за неопределенного общего экономического климата. Вероятность Бейеса будет означать, что вероятность события  $A$ , обусловленного событием  $B$ , основана на априорной (предшествующей) вероятности  $A$ . Поэтому мы должны быть способны подсчитать  $P(A | B)$ , но мы не можем сделать это, так как априорная вероятность  $A$  не сообщает будущие вероятности.

Снова имеются сходства между институциональной и более индивидуальной играми. Существуют другие стороны, которые влияют на результаты. Администраторы колледжей и университетов принимают решения о бюджетах и диктуют, могут ли библиотеки больше или меньше год от года тратить. Чтобы усложнить ситуацию, индивидуальная игра влияет на институциональную. В частности, поскольку большие числа имеют важное значение в этой оценке отдельных преподавателей, то вторичные производители выигрывают как от представления все большего и большего числа статей в журнал, так и от давления на библиотеки в целях обеспечения доступа к журналам, в которых публикуются преподаватели. Возникает вопрос: может ли институциональная игра быть решена без решения индивидуальной игры? Это уклончивый вопрос, но можно утверждать, что согласование библиотекарями сокращения стоимости согласованных статей может оказывать влияние на положение отдельных преподавателей и редакторов журналов. Сокращение числа статей, представленных в журнал, и числа опубликованных журналов и/или выпусков может привести к общему сокращению затрат. Согласование этого типа основано на особом предположении, которое направит библиотекарей к принятию решения. Это предположение состоит в том, что современная система коммуникации производит больше информационных статей, чем требуется для создания общей интеллектуальной базы. Если, например, ученый (или группа ученых) вовлечен в создание наименьших печатных единиц (LPUs – Least Publishable Units) с целью повысить число своих публикационных записей, то дополнительные статьи могут даже скрыть доступ ко всему количеству научных результатов. [LPU относится к делению продуктов на части, приводящему в результате исследования к самым маленьким элементам, которые могут быть опубликованы.] Это поведение может быть принято кем-то неохотно: «То, чему я научился как администратор, состоит в том, что хотя я могу предпринять усилие, чтобы распознать качество записи в послужном списке преподавателя, но чем длиннее перечень его должностей и профессиональных программ, тем труднее для администраторов или членов соответствующего комитета сделать это» [21]. Те же самые люди, которые допускают наличие трудностей в оценке, тем не менее, могут указывать и на некоторые преимущества такого образа действий преподавателей: «Наименьшая печатная единица будет поддерживать перспективность профессора. Когда мы накапливаем данные для большой статьи, имеется тенденция позволить им накапливаться, пока инерция непроанализированных, не скопированных, непродуманных данных сделает перспективу ее написания

почти непреодолимой» [21]. Оба эти фактора могут вступить в игру, но они могут также омрачить *рационализациями* действительно рациональное принятие решений. Устранение многословных или ненужных публикаций может привести к системе, которая является более экономической, включая решения относительно назначения на постоянную должность и дальнейшего продвижения.

В этот момент возникает другой вопрос: может ли коммерческая производственная система внести вклад в необходимые реформы системы научной коммуникации? Хотя ответом на этот вопрос могло быть «да», вторичные коммерческие производители должны побеспокоиться о доходах и прибылях. Эти производители будут играть в игру в соответствии с интересами и предпочтениями, которые отражают их коммерческий характер. Хотя сотрудничество возможно, эти интересы и предпочтения очень отличаются от интересов библиотечных и преподавателей. Библиотекари могут ожидать проведения переговоров со вторичными производителями, но игроки подходят к игре с двух разных точек зрения. На языке теории игр может существовать доминирующий стратегический баланс стратегической игры. Если можно сказать, что библиотекари/ученые учреждают одного игрока, а традиционные производители/вторичные производители учреждают другого игрока, то последний игрок находится в диктаторской позиции. Существует вероятность, что открытое и общее согласование может быть заблокировано различными интересами, поэтому применение вероятности Бейеса в ситуации теории игр, вероятно, не должно встретиться, и еще меньше должно быть успешным.

Институциональная игра может быть изменена в недалеком будущем, даже если индивидуальная игра не реформируема в корне. Производство открытого доступа информационных статей является моделью, которая очень отличается от общепринятой коммерческой модели. Все еще существуют значительные затраты, связанные с производством открытого доступа, но эти затраты ниже, чем для большинства журналов, особенно если выгода не является переменной в подсчете затрат. Сложности открытого доступа в данной статье не рассматриваются, будут затронуты только некоторые моменты. Если возможно некоторое применение теории игр к анализу отношения библиотек к производителям открытого доступа, тогда должна быть возможность для изучения общих интересов и предпочтений. Идеал открытого доступа включает более тесное сотрудничество с авторами/учеными. Такой идеал, поскольку он действительно достижим, трансформирует игру путем изменения выбора отношений, сокращения кооптации доминирующего игрока. Утверждается, что вероятно существуют напряжения, если не конфликты между отдельными исследователями и вторичными производителями журналов. Сокращение, если не устранение, напряжения может привести к коалиционному играм, где возможен некоторый баланс. Также должна быть возможность изучения актуальных согласований и сделок, которые может облегчить сотрудничество. Эти факторы показывают, что сдвиг, который открытый доступ может (и акцент должен быть на *может*) помочь вызвать, будет в основном экономическим. Подробнее об открытом доступе и затратах см. в работе [22].

## ЗАКЛЮЧЕНИЕ И ДАЛЬНЕЙШЕЕ ИССЛЕДОВАНИЕ

Приводит ли динамичная система научной коммуникации к самой этической моральной экономии для высшего образования? В действительности, она не изучает этические проблемы совсем. Выражение вероятности,

присущее теории игр, не делает ничего, а только показывает, какой набор данных вероятнее всего указывает на что-то, подобное решению относительно постоянной должности, а какой нет. Что касается простой прагматики, то это выражение вероятности может быть полезным; преподавателям показывают, в какие действия они должны быть вовлечены, если хотят получить постоянную должность. Более того, это выражение указывает на эти данные профессорско-преподавательскому составу, когда индивидуумы находятся в начале своей карьеры. Бинмор [5] поместил бы это выражение в рамки сферы утилитарности; это является хорошим результатом собрания данных и подсчета вероятности того, что факультет проинформирован о способах, какими решения о постоянной должности должны приниматься в университетах. Бинмор [23] представляет отношение предпочтения символом “ $<$ ” в терминах реально значимой функции полезности:  $u(a) \leq u(b)$  если и только если  $a < b$ . Другими словами, полезность  $a$  меньше или равна полезности  $b$ , если и только если  $a$  менее предпочтительно, чем  $b$  для тех, кто делает выбор. Это представление является случаем теории игр и является настолько фундаментальным, чтобы быть предметом для вводных учебников. Однако что касается библиотечного дела, то эта основа должна стать явной. Если близко придерживаться точки зрения Бинмора, когда дело касается решений о постоянной должности и продвижения, то должны быть сделаны некоторые предположения, и некоторые из этих предположений будут проблематичными. Непроблематичным является то, что индивидуумы, выбирающие стремление к постоянной должности и продвижению в университете и обеспокоенные этой возможностью, вероятно, должны решить публиковаться стратегически такими способами, которые будут действовать в их интересах (как, например, публикация в течение шестилетнего периода как можно большего числа статей в рецензируемых журналах, или публикация лучших работ в признанных ведущих журналах). В таком случае, эти индивидуумы, вероятно, в таком случае должны быть в штате профессорско-преподавательского состава в конце шестилетнего периода обучения.

Здесь имеется возможная трудность со взглядом Бинмора [5] на не телеологический утилитаризм (польза не направленная априорной целью). Не телеологическая позиция подразумевает, что утилитаризм, как он его пропагандирует, является продуктом разума. Он получает поддержку своей позиции у Хьюма. Хотя Хьюм [17] (как указано выше) сказал, что «разум есть и должен быть рабом страстей». Хьюм подразумевает, что воля не вступает в рациональность; разум относится только к внешнему миру (следовательно отсутствие цели, которую только можно желать). Университет имеет телеологическое предпочтение; оно в первую очередь основывает решения на числе публикаций; престиж через публикацию является целевым решением университетских администраторов. Это предпочтение происходит по ряду причин: большее число публикаций приводит к большей обзримости учреждения (на основании признания принадлежности преподавателей); публикации поднимают престиж (большие показатели числа публикаций ведут к более высокому ранжированию при подсчетах голов и в обзорах); публикации показывают кардинальное (количественное) и порядковое ранжирования среди коллег (подсчитываются все публикации или число публикаций на человека); или публикации обеспечивают косвенный показатель научной интенсивности и/или важности университета. Многие преподаватели также имеют телеологические предпочтения. Некоторые хо-

тят получить постоянную должность и продвигаться в университетах, возможно, чтобы обеспечить гарантии. Интересы некоторых преподавателей могут быть направлены на пополнение общей интеллектуальной базы своих областей. Те преподаватели, которые оставляют университет до решения о назначении на постоянную должность и дальнейшем продвижении, могут выражать телеологическое предпочтение в пользу обучения и/или службы. Учитывая только что описанные предпочтения, телеологический утилитаризм, в лучшем случае, можно обсуждать. Библиотекари также имеют телеологические предпочтения, которые являются менее эгоцентричными, чем таковые администраторов или преподавателей и, значит, существует вероятность напряженности.

Вышеупомянутое представляет собой изучение возможного применения теории игр к экономикам научных библиотек, высшего образования и производства информации. Эти экономики, безусловно, являются сложными, но есть некоторые проблемы, связанные с применением классической теории игр. Некоторые трудности являются скорее подчиненными по отношению к истинной теории игр. Например, существует относительно небольшое число «ведущих» или высоко престижных журналов в определенной области. Преподаватели - те, кто стремится занять постоянную должность или получить звание старшего преподавателя - сталкиваются со значительной конкуренцией, когда дело касается публикации в этих журналах. Проблема, с которой сталкивается младший преподаватель, состоит в том, что Роберт Мертон [24] назвал «эффектом Матфея»: тот, кто однажды оставил данные (след) в публикации, наиболее вероятно будет успешнее других публиковаться в будущем. К тому же, а это касается почти всех журналов, особенно в социологии и гуманитарных дисциплинах, существует временной интервал между принятием статьи и ее реальной публикацией. Такая задержка может сработать против преподавателей, стремящихся к постоянной должности и продвижению. Еще одной трудностью является то, что некоторые редакторы журналов сообщают, что становится все сложнее найти рецензентов и получить своевременные и полезные рецензии на представленные рукописи. Наоборот, те преподаватели, которые пишут своевременные и полезные рецензии на рукописи, считают себя загруженными требованием рецензировать все больше и больше статей [25]. Все эти и другие факторы влияют на экономики, частью которых являются высшее образование и его компоненты.

Принятие рациональности и общего знания (среди прочего) может быть проблематичным; не все игроки осознают, что принимают участие в форме игры, когда дело касается, например, пребывания в постоянной должности и решений о продвижении по службе. Ценовая политика производства информации и оплата экономики имеют даже более сложные, связанные с этим, проблемы. Могут существовать альтернативы классической теории игр; некоторая теоретическая работа была недавно проделана с эволюционной альтернативой (см., например, [26]). Классическая теория игр стремится быть основанной на принципах физики, но некоторые экономисты предполагают, что биология может быть более подходящим источником для основных принципов (см. [27] на предмет более раннего понятия этой связи). Если теория игр подразумевает моральный аспект, то трудности с применением здесь включают два момента: 1) теория игр в том виде, как она развивалась до сих пор, ограничена в возможности охватить эти моральные аспекты и/или 2) высшее образование, библиотеки и производство информации не включают моральное поведение в свои

действия (по крайней мере, в том смысле, что теоретики игры формулируют идеалы морального аспекта этой теории). Дальнейшее изучение применения теории игр к своего рода проблемам, исследуемым здесь, должно включать появляющееся эволюционное изменение. Следует отметить, что сами экономисты работают над (с) применениями биологической модели, поэтому стабильный теоретический подход к настоящему времени еще не разработан. Гал [28] отмечает, что «имеется эмпирическое свидетельство того, что экономические агенты заботятся не только о своих собственных физических результатах, но и о результатах своих оппонентов и о том, как их сравнить». Этот тип мышления привносит вклад в полноту действий, окружающих сокращения, и может стать основой для дальнейшего исследования взаимосвязей между библиотеками и производителями информации. Возможно, наиболее важно то, что изучение преобладания баз данных и других ресурсов, которые предоставляют доступ к журнальной литературе, должно включать эволюционный подход к обеспечению информацией. Если, скажем, открытый доступ подразумевает большую степень сотрудничества среди многочисленных игроков (редакторы, авторы и библиотекари), тогда возможные выгоды, включая подход переносной полезности, могут быть изучены. Даже с такого рода аналитическим подходом некоторые значимо-перегруженные переменные остаются, как, например, стоимость страницы и способность автора оплатить ее, должны быть компонентами. Сагден [4], который видит значительный потенциал эволюционной альтернативы, делает акцент на том, что эта альтернатива должна быть чувствительной к историческим обстоятельствам. Исследование эволюции является историческим; оно изучает изменения, происходящие во времени. Это, рассматриваемое здесь, исследование экономик должно в будущем быть наподобие исторического.

## ЛИТЕРАТУРА

1. *von Neumann J., Morgenstern. O. Theory of games and economic behavior.* — Princeton: Princeton University Press, 1944.
2. *Nash J. The bargaining problem// Econometrica.*— 1950.— Vol. 18, No.2.— P. 155-162.
3. *Gintis H. Game theory evolving: A problem centered introduction to modeling strategic interaction.* — Princeton: Princeton University Press, 2009.
4. *Sugden R. The evolutionary turn in game theory// Journal of Economic Methodology.*— 2001.— Vol. 8, No.1.— P. 113-130.
5. *Binmore K. Game theory and the social contract, volume 2: Just playing.*— Cambridge: MA, MIT Press, 1998.
6. *Binmore K. Game theory: A very short introduction.* — Oxford: Oxford University Press, 2008.
7. *Piramuthu S., Shaw M. J. Learning-enhanced adaptive DSS: A design science perspective// Information Technology & Management.* — 2009.— Vol. 10, No.1.— P. 41-54.
8. *Ye D. W., Forgy G. Testing a decision-theoretic approach to the evaluation of information retrieval systems// Journal of Information Science.* — 2008.— Vol. 34, No.6. — P. 861-876.
9. *Jolly R., Wakeland W. Using agent based simulation and game theory analysis to study information flow in organizations: The KMscape// International Journal of Knowledge Management.*— 2009.— Vol. 5, No.1.— P. 17-28.
10. *Dickinson D. L. A bargaining experiment to motivate a discussion on fairness// Journal of Economic Education.*— 2002.— Vol. 33, No.2.— P. 136-151.

11. *Bergstrom T. C.* Librarians and the terrible fix: Economics of the big deal// *Serials*.— 2010.— Vol. 23, No.2.— P. 77-82.
12. *Hart S., Mansour Y.* How long to equilibrium? The communication complexity of uncoupled equilibrium procedures// *Games and Economic Behavior*.— 2010.— Vol. 69.— P. 107-126.
13. *Allingham M.* Choice theory: A very short introduction. — Oxford: Oxford University Press, 2002.
14. *Cho M., McKee M.* Authorship in biomedical research: Realities and expectations// *Science*. — 2002. — [http://sciencecareers.sciencemag.org/career\\_magazine/previous\\_issues/articles/2002\\_03\\_01/noDOI.1847997411683997393](http://sciencecareers.sciencemag.org/career_magazine/previous_issues/articles/2002_03_01/noDOI.1847997411683997393).
15. *Rapoport A.* Two-person game theory. — Mineola, NY: Dover, 1966.
16. *Kalai E.* Games in coalitional form /S. N. Durlauf, L. E. Blume (Eds.)// *Game theory*.— P. 127-143.— New York: Palgrave MacMillan, 2010.
17. *Hume D.* A treatise of human nature/ D. F. Norton, M. J. Norton (Eds.)— Oxford: Clarendon Press 2.3.3.4., 2000.
18. *Kyrillidou M., Bland L.* ARL statistics, 2007-2008.— Washington, DC: Association of Research Libraries, 2009.— <http://www.arl.org/bm~doc/arlstat08.pdf>
19. *Hausman D. M., McPherson M. S.* Economic analysis and moral philosophy. — Cambridge: Cambridge University Press, 1996.
20. *Winkler R. L.* An introduction to Bayesian inference and decision, 2<sup>nd</sup> ed. —Gainesville, FL: Probabilistic Publishing, 2003.
21. *Owen W. J.* In defense of the least publishable unit// *Chronicle of Higher Education*.— 2004.— Vol. 50, No. 23. — C1-C4.
22. *Beaudoin-Lafon M.* Open access to scientific journals// *Communications of the ACM*. — 2010. — Vol. 53, No. 2.— P. 32-34.
23. *Binmore K.* Rational decisions. — Princeton: Princeton University Press, 2009.
24. *Merton R. K.* The Matthew effect in science: The reward and communication systems in science are considered// *Science*. —1968. — Vol. 159, No.3810. — P. 56-63.
25. *Myers D. J.* The peer-review system is broken// *Chronicle of Higher Education*.— 2009.— Vol. 56, No.2.— B4.
26. *Sandholm W. H.* Deterministic evolutionary dynamics/ S. N. Durlauf, L. E. Blume (Eds.)// *Game theory*. — P. 78-86. — New York: Palgrave MacMillan, 2010.
27. *Hammerstein P., Selten R.* Game theory and evolutionary biology/ *Handbook of game theory with economic applications*.— Vol. 2.— P. 929-993.— Amsterdam: Elsevier, 1994.
28. *Gul F.* Behavioural economics and game theory/ S. N. Durlauf, L. E. Blume (Eds.)// *Game theory*.— P. 20-28. —New York: Palgrave MacMillan, 2010.
29. *McCain R.* Commitment and weakness of will in game theory and neoclassical economics// *Journal of Socio-Economics*. — 2009. — Vol. 38, No.4.— P. 549-556.

## Оценка влиятельности журналов: вебометрическая перспектива\*

Майк ТЕЛВОЛЛ  
(Mike THELWALL)

Исследовательская группа статистической  
киберметрии, Школа технологии, Университет  
г. Вулвергемптона, Великобритания

*В теории сеть имеет потенциал для обеспечения информации о более широком влиянии академического исследования за рамками традиционного научного влияния. Это потому, что сеть может отражать ненаучное использование исследования, такое как в правительственных документах в режиме онлайн, освещении в печати или публичных обсуждениях. Тем не менее, существуют практические проблемы с созданием метрик для журналов, основанных на сетевых данных: в основном такие, что большинство таких метрик должны быть легкими для редакторов или издателей журналов с целью манипуляции. Однако две альтернативы, кажется, должны быть перспективными и важными: ссылки, полученные из оцифрованных книг, и загружаемые подсчеты для журналов в рамках особых платформ доставки.*

### ВВЕДЕНИЕ

Широко используемые импакт факторы журналов являются показателями интеллектуального влияния, основанными на среднем числе ссылок на недавно опубликованные статьи любого выбранного журнала. Точнее, импакт фактор журнала за определенный год представляет собой число ссылок, данных в этом году на статьи, опубликованные в последние два года, — деленное на число «цитируемых статей» в журнале (как оценивается системой Thomson Reuters), опубликованных в предыдущие два года. Ссылки считаются только в том случае, если они имеются в журналах, сериальных выпусках или других публикациях, проиндексированных Thomson Reuters, и определяются как указывающие на статью в данном журнале в предыдущие два года. Несмотря на значительные технические и теоретические ограничения импакт фактора журнала [1-3], многие из которых, кажется, должны быть широко известными и признанными в наукометрии долгое время [4-5], он продолжает использоваться как подходящая оценка качества журнала. Однако эта статья фокусируется не на технических ограничениях, а на более широком ограничении импакт фактора журнала из-за того, что он основывается только на опубликованной научной работе. В частности, эта статья оценивает возможность замены или дополнения импакт фактора журнала с помощью информации, полученной из сведений сети. Это своевременно в силу недавнего появления движения *альтметрик*, стремящегося развить ряд новых показателей научного исследования,

которые изучают потенциал информации, открыто предоставляемой в сети [6-7].

Использование подсчетов ссылок в качестве показателей научного влияния вытекает из веры Мертона [8] в науку как нормативную и совокупную предметную область. Когда новые вклады делаются в науку, они, как правило, формализуются в опубликованных журнальных статьях. Совокупный характер науки признан учеными, цитирующими работу, которую они использовали для развития своего нового знания. Кумулятивный характер науки признается во фразе, написанной Ньютоном, «Если я видел дальше других, то потому, что стоял на плечах гигантов» [9]. Нормативная теория Мертона была основана одним из его студентов, явно использующим его анализ цитирования. Ранняя работа была весьма перспективной, обнаруживающей эмпирическое свидетельство того, что подсчеты ссылок соотносятся с экспертной оценкой лучше, чем любая другая легкодоступная количественная мера [4]. Тем не менее, имеются различные суждения, которые могут возникать из этой теории ссылок. Во-первых, авторы могут выбирать ссылки по причинам, отличным от научного качества и релевантности [10-12]. Но существеннее то, что ценность науки лежит в ее способности позволять нам понимать или контролировать физическую среду, и поэтому представляется возможным предвидеть научный успех, который получит редкоцитируемая публикация вопреки теории цитирования. Например, явно важная область исследования, такая как холодное слияние, по этой теории должна бы быть ложной или бесполезной [13]. В противном случае, большинство научных вопросов может быть решено таким образом, что сама проблема больше не является безотлагательной, а рассматриваемая научная область может стать бездеятельной. Простым примером этого является открытие того, что мытье рук в антисептических целях

\* Перевод Thelwall M. Journal impact evaluation: A webometric perspective. — 2012.— [http://www.scit.wlv.ac.uk/~cm1993/papers/Journal\\_impact\\_evaluation\\_Webometrics\\_preprint.pdf](http://www.scit.wlv.ac.uk/~cm1993/papers/Journal_impact_evaluation_Webometrics_preprint.pdf)

останавливает распространение родильной горячки, хотя в таком случае данное открытие потребовало времени, чтобы быть принятым [14].

В общем смысле публикуемые статьи могут внести значительный вклад в науку в целом таким образом, что маловероятно, что они отразятся в ссылках, потому что они будут использованы для решения проблем, не требующих дальнейшего исследования. Например, область анализа патентных ссылок появилась с целью признать и количественно оценить важность некоторого научного исследования, которое приводит к коммерческому использованию [15-16]. Однако данный подход срабатывает только для областей прикладного исследования, где патентование общепринято. Существуют также многие некоммерческие применения исследования, такие как улучшение образования и обучения или помощь принятия различного рода решений, таких как экономические, социальные или медицинские. В знак признания этого финансирование правительством Великобритании будет предоставляться тем исследованиям, которые могут показать «научное влияние» или «экономическое или общественное влияние» как часть учета регулярной оценки исследования. Эта часть экономического или общественного влияния включает, помимо других оценок, как «улучшение культурного обогащения, качества жизни, здоровья и материального благополучия», так и «содействие в доказательности решений и государственной политики и законодательства на местном, региональном, национальном и международном уровнях» [17]. Исследование может быть также ориентировано на профессию, такую как библиотечная профессия, социальная работа, медсестринское дело и юриспруденция, и таким образом вносить вклад в основном в профессиональную практику. Ученые поэтому могут влиять на мир, который недостаточно представлен ссылками, которые получают их публикации. Это мнение может быть также применено к самим журналам, таким, которые сосредоточены на политических или профессиональных вопросах или образовании. Подходящим примером является журнал *Journal of Education for Library and Information Science* (JELIS), который исчез из Science Citation Index из-за постоянно низкого импакт фактора, несмотря на хорошую репутацию внутри сообщества по библиотечному делу и информатике [18].

Другой фундаментальной проблемой с точки зрения Мертоновской перспективы является то, что, возможно, исследования по социальным и гуманитарным наукам могут первоначально не содержать обобщений, но могут вносить вклад, который гораздо важнее, чем сообщение фактов, а именно предоставление интересных интерпретаций или менее четких перспектив. Кажется, это должно быть типичным для областей, в которых исследуемые феномены являются трудно контролируемыми, поскольку знание может быть более незначительным, и в результате менее универсально согласованным. Это может быть описано как высшая «техническая задача неопределенности», по словам Вайтла [19], или эти дисциплины могут быть описаны как «мягкие» [20].

Несмотря на все ограничения анализа ссылок, он все еще является полезным инструментом для оценки исследователей, при условии осторожного использования с гарантиями, соответствующими уровню примененного конгломерата. Это существенно, так как некоторые выше обсуждаемые факторы стремятся усреднять большие группы исследователей. На уровне отдельных журналов, тем не менее, могут быть четкие и систематические противоречия, которые будут приводить к неверным импакт факторам журналов, даже если последние правильно подсчитаны. В частности, любые журналы, специализи-

рующиеся на профессиональном, образовательном или коммерческом применении, могут, по-видимому, иметь низкое значение фактора влияния по сравнению с теоретическими журналами, которые могут не иметь большого влияния вне научной литературы. Опора на импакт фактор журнала, вероятно, должна привести к систематическому смещению репутации в пользу журналов с теоретической ориентацией. Говоря простым языком, в смешанном порядке ранжирования списка журналов из одной и той же предметной области, таком как категории Journal Citation Reports, журналы с профессиональной, прикладной или образовательной ориентацией могут располагаться ниже, чем должно определять их значение. Безусловно, другие факторы могут также влиять на импакт факторы журналов, которые могут его усиливать или уменьшать, такие как нормы области по числу ссылок и любые перекрестные влияния в прикладном или образовательном исследовании. Влияние этих факторов может быть смягчено до некоторой степени путем ввода более сложного подсчета для новых факторов влиятельности журналов [2], но данная статья вместо этого концентрируется на новых источниках достоверности фактов для рассмотрения этих проблем.

Сеть и оцифровка журналов сделали возможным подсчет новых метрик для журналов и статей, и это может иметь потенциал для исправления некоторых недостатков фактора влиятельности журнала. Это может происходить двумя основными путями. Во-первых, поскольку сеть содержит не только научные статьи, но и более широкое разнообразие публичного контента, такого как интерактивные газеты и списки лекций университетских курсов, то возможно собрать ссылки из большего числа типов документов и, следовательно, охватить более широкие типы влияния, включая некоторые, которые недостаточно охвачены импакт фактором журнала. Во-вторых, можно подсчитать показатели того, как часто читается статья, на основе доступа к статистике для электронных версий. Это дает возможность подсчитать влияние в форме круга читателей, который может включать чтение для информирования о новом исследовании, а также обучающих, коммерческих, образовательных и других целей. Эта статья рассматривает исследования, изучающие интерактивные метрики такого рода, как для отдельных статей или целых журналов, и затем обсуждает возможность такого рода свидетельства, чтобы помочь в оценке научных журналов.

## ПОКАЗАТЕЛИ ИНТЕРАКТИВНОЙ ЧИТАЕМОСТИ

Возможно, наиболее очевидным источником новой информации о журнальных статьях в электронную эпоху является использование информации: как часто они читаются. Издатели журналов, хранящихся в режиме прямого доступа в цифровых библиотеках, будут иметь возможность регистрировать, как часто каждая статья запрашивается или загружается. Эта информация иногда предоставляется издателям или редакционным коллегиям и может быть использована для создания списков наиболее доступных/загружаемых статей или может быть представлена на информационных страницах, связанных с каждой статьей. Журналы, которые используют, по крайней мере, один из этих подходов (по данным на январь 2012 г.), включают *PLoS ONE* (например, «наиболее просматриваемые» на домашней странице), *Journal of Informetrics* (например, статистика доступа издательства Elsevier, предоставляемая ежегодно редакционной коллегии) и *Journal of Medical Information Research* (например, «Главные статьи» на домашней странице). Издатели уже могут подсчитывать факторы влиятельности на основе

загрузки для журналов в их массиве и использовать эти статистики для их оценки. Это имеет смысл в эпоху больших сделок для журналов: скорее продажа групп названий, а не отдельных названий. Без больших сделок пороговая линия для журналов должна представлять просто продажу или прибыль. Информация о загрузке может также использоваться, чтобы доказать ценность публикации покупателю, особенно если учреждение может потерпеть неудачу в подсчетах загрузки. В принципе, должно быть простым подсчитать среднее число загрузок на статью для каждого журнала, фактор влиятельности загрузки. На практике, однако, имеются три важные проблемы.

Первая проблема – подсчеты доступа не могут отражать число читателей. Статьи могут быть загружены, но не прочитаны; они могут быть напечатаны и переданы многим людям, таким как классу колледжа; они могут быть прочитаны в печатной версии журнала; версии могут быть также прочитаны через авторский учрежденческий архив или копию, хранящуюся на его сетевом сайте.

Вторая проблема состоит в том, что если цифровые данные используются для сравнения журналов, произведенных различными издателями, тогда должны быть согласованы надежные средства, чтобы сделать данные сравнимыми. Без этого каждый издатель будет свободен зависеть свои данные, как ему понравится. Следовательно, необходим стандарт индустрии, чтобы сделать факторы влиятельности загрузки возможными как общее дополнение к импакт фактору журналов. Двумя такими стандартами являются инициатива сбора статистики нормированного использования (SUSHI), которая является формализованным стандартом ANSI/NISO Z 39.93 – 2007 [21] и инициатива подсчета интерактивного использования сетевых электронных ресурсов (COUNTER) [22]. Последний сертифицирует особые типы отчетов как «COUNTER-compliant», если они отвечают его стандартам. Производители COUNTER-compliant являются также предметом периодического аудита. Список соответствующих производителей для статистики журнальной загрузки на момент написания этой статьи включал много крупных издателей, таких как Elsevier, John Wiley and Sons и Springer Verlag, и, таким образом, это кажется многообещающей инициативой [23]. В принципе использование факторов влиятельности для любого из перечисленных издателей будет сравнимым. Последующий проект начал изучать потенциал глобального использования факторов влиятельности для журналов [24]. Он рекомендовал, помимо других вещей, использование скорее медианного, чем среднего значения загрузок статей для журнала из-за асимметричных распределений загрузки. Тем не менее, необходимость одобрения издателем является все еще важным барьером на пути к этому.

Наконец, все загрузки не могут иметь равной ценности. Приведем крайне противоположный пример, статья, упомянутая в газете, часто может быть доступной обычным читателям, которые никогда не читают полного текста после обращения к ней, так как это слишком трудно для понимания несведущего. Наоборот, другая статья может быть доступна человеку, который использует информацию для производства чего-то важного, такого как доработка исследования, коммерческое применение или составления профессиональных рекомендаций. Более серьезно, цифровые данные могут быть довольно легкими для манипулирования ими, пока не будут разработаны эффективные защитные меры. Автор может неоднократно загружать свои собственные статьи или обязать своих студентов обращаться к ним, даже если они не являются ключевыми в каком-либо курсе.

В ряде исследований изучается использование подсчетов доступа для наукометрии путем оценки того, могут ли

они использоваться для предсказания будущих цитирований. Эти исследования показывают, что подсчеты загрузки могут предсказать будущие цитирования, поэтому подсчеты загрузки могут быть использованы как ранние свидетельства вероятного влияния статей [25-26] и служить соответствующими средствами гарантии, возможно, также для журналов. Значительные корреляции между подсчетами загрузок и ссылок также свидетельствуют о том, что загрузки стремятся быть научно-связанными в некотором роде, нежели просто случайными. Это делается, чтобы как-то объяснить их использование для наукометрических показателей. Однако кажется нет систематических исследований статей, которые привлекают относительно много загрузок по сравнению со ссылками. Это должно служить доказательством, что имелись положительные причины того, почему статьи привлекли относительно больше ссылок, чем загрузок, чтобы подтвердить утверждение, что показатели загрузки будут полезны, чтобы дополнять цитирования, основанные на показателях для отдельных статей или целых журналов.

Несмотря на значительные корреляции, найденные между ссылками и доступом к отдельным статьям, на уровне журналов этой корреляции кажется не существует. Используя статистику загрузки из девяти институтов Калифорнийского университета, применение факторов влиятельности к данным, собранным в 2004 г, обнаружило отсутствие корреляции между ними и импакт факторами журналов в целом [27]. На уровне отдельных дисциплин только одна из 17 изученных (образование) проявила значительную корреляцию между использованием факторов влиятельности и факторов загрузки журналов. Возможным объяснением отсутствия значительных корреляций является то, что интересы специалистов общества, представленного в этом примере, полностью не отражают их глобальные интересы. Положительные корреляции могли быть найдены в дисциплинах с относительно большими популяциями *выпускников*, а отрицательные корреляции - в дисциплинах с относительно большими популяциями *студентов* [27]. Возможным выводом из этого является то, что журналы с низкими факторами влиятельности стремятся быть *более* полезными для обучения студентов, чем журналы с высокими факторами влияния. Проиллюстрируем другие возможные объяснения. Журналы *Science* и *Nature* имели низкое использование (0,3) по сравнению с высокими факторами влияния журналов (31,9 и 32,2 соответственно), но оба продаются в значительной степени по печатной подписке индивидуальным пользователям, и, таким образом, их читаемость может быть плохо представлена посредством использования факторов влиятельности. Для сравнения – пример, где самое высокое использование журнала в двадцать раз больше его фактора влияния: *Topics in Early Childhood Special Education* (использование фактора влияния – 6,8; фактор использования – 0,9).

Новый метод подсчета использования журналов основан на их доступе в рамках отдельного сообщества, такого как группа университетов [27]. Это кажется целесообразным для больших сообществ, которые имеют общую библиотечную службу, дающую им доступ к журналам. Если может быть разработано программное обеспечение для регистрации доступа наряду с публикацией информации о числе статей на журнал, тогда это позволит подсчитать локализованные импакт факторы журналов. Они будут иметь преимущество адаптации к локальным условиям, но не будут служить рыночным потребностям издателей таким образом, как фактор влияния журнала. Предположительно факторы локализованного использования будут менее устремлены на манипулирование,

но издатели все еще могут нуждаться в достижении согласия по обеспечению необходимой информации, позволяющей осуществлять подсчет доступа к статьям унифицированным образом.

## ПОКАЗАТЕЛИ СОЦИАЛЬНОЙ УСТАНОВКИ ЗАКЛАДКИ

Альтернативой показателем интерактивной читаемости является подсчет того, как много людей добавило данную статью в свое программное обеспечение управления онлайн ссылками или архив. Такими людьми, по-видимому, являются те, кто читают статью и оценивают ее настолько хорошо, что хотят ее записать, а также те, кто стремится прочитать статью в будущем – предположительно потому, что они оценили ее релевантной по названию статьи или реферату - и записывают ее для этой цели. Программы управляющие онлайн ссылками, такие как Zotero, CiteULike [28], Mendeley [29], BibSonomy и Connotea, являются поэтому логичными местами для поиска публичных доказательств читаемости статей. Они предлагают свободную электронную услугу установки закладки статьи с дополнительной стоимостью средств для социального взаимодействия [30]. Тем не менее, этот источник данных, по-видимому, должен страдать от подобных смещений в сторону цифр онлайн читаемости. В частности закладки могут первоначально создаваться студентами и отражать списки для курсового чтения. Более того, статьи, связанные с социальной установкой закладок и социальной сетью, по-видимому, должны привлекать диспропорционально больше социальных закладок.

Имеется некоторое подтверждение важности подсчетов закладок как показателя научной ценности, так как одно исследование сделало вывод, что подсчет ссылок соотносится с подсчетом закладок Mendeley для массива из 1163 статей из журналов *Science* и *Nature* в 2007 г. [31], хотя это же исследование 2011 г. нашло социальные системы закладок слишком редко используемыми для предоставления вообще полезных данных для целей оценки влияния отдельных статей. Однако это может быть несправедливо для целых журналов. Изучение 45 журналов по физике с помощью CiteULike, BibSonomy и Connotea сравнило фактор влияния с рядом метрик, полученных от этих служб, таких как число пользователей на журнал и число статей с закладками на журнал, найдя значительные корреляции с большинством из них [32]. Ряд использованных показателей предполагал, что услуги социальных закладок могут использоваться для показателей ряда различных типов журнального влияния. Эти и другие исследования также обнаружили значительные практические проблемы с неточными и дублирующими данными [33]. Тем не менее, имеется некоторое подтверждение значимости подсчетов социальных закладок.

Несмотря на корреляции между показателями закладок и подсчетами ссылок, некоторые статьи являются высокоцитируемыми, но не с высоким числом закладок, и наоборот [33], предполагая, что закладки отражают другой аспект влияния по отношению к влиянию ссылок.

Одним решающим преимуществом системы социальных закладок является то, что они кажутся открытыми и прозрачными перед тем, что любая манипуляция ими может легко быть обнаружена. Тем не менее, так как имеется незначительный контроль или отсутствует контроль их качества, то легко представить законные способы их использования, которые будут манипулировать результатами любой статистики, порожденной ими. Например, почти любой лектор может поставить своему классу задачу по изучению того, как использовать систему соци-

альных закладок и оценить их присвоение закладок в массиве статей по выбору лектора. Таким образом, этими результатами будет вполне легко оправданно манипулировать.

## АНАЛИЗ ССЫЛОК

Гиперссылки в сетевых страницах являются техническим механизмом, позволяющим пользователю передвигаться от страницы, которую он читает, к другой странице с помощью клика на эту ссылку. Информационные специалисты определили на ранних этапах становления сети, что подобно библиографическим ссылкам гиперссылки являлись междокументальными связями [34-35] и могли, вероятно, использоваться для целей анализа цитирований. Когда коммерческая поисковая система ввела возможность поиска ссылок, тогда это превратило ее в показатель цитируемости и саму сеть в возможный источник доказательства влияния [36-37]. В ответ на это ряд исследователей начал изучать обоснованность использования подсчетов гиперссылок для подобных цитированию метрик для журналов, индивидуумов, исследовательских групп и учреждений, как описано ниже.

За рамками наукометрии гиперссылки широко используются как подтверждение важности документов или сетевых сайтов. Например, алгоритм Google's key PageRank algorithm [38] явно мотивирован анализом ссылок и разработан в помощь ранжированию сетевых страниц, соответствующих поиску. Другие алгоритмы ранжирования на основе гиперссылок также разрабатываются [39]. Гиперссылки также используются как подтверждения тем сетевых страниц [40-42] среди других применений. Поэтому значимость гиперссылок широко используется. Последствием важности гиперссылок для ранжирования поисковых систем и коммерческой ценности местоположения поисковых систем является то, что имеются широко распространенные попытки создания ссылок спама, единственной целью которых является улучшение ранжирования целевых сетевых сайтов. В свою очередь это приводит к продолжающемуся развитию алгоритмов, определяющих спам, использующих различные свойства гиперссылок как показатели спама или явно нацеленные на выявление спама среди ссылок [43]. Например, алгоритм TrustRank [44] использует структуру ссылок сети, чтобы присвоить сетевым сайтам значение доверия: по существу самыми заслуживающими доверия страницами являются те, на которые ссылаются из других надежных страниц. Он (TrustRank) предполагает, что научные сетевые сайты являются относительно заслуживающими доверия, и, таким образом, использует их, чтобы запустить рекурсивный алгоритм.

В рамках информатики гиперссылки оцениваются как показатели влияния для университетов внутри страны [45-49], отделений внутри страны [50-51] и отдельных исследовательских групп [52]. Эти исследования обнаружили значительные корреляции между подсчетами ссылок на университеты или отделения и мерами их исследовательской продуктивности. Тем не менее, эти корреляции были сильнее для более крупных единиц – т.е. сильные для целых университетов, слабые для отделений внутри области и незначительные для исследовательских групп внутри области. Это предполагает, что уровень агрегации важен, а поскольку журналы относительно малы, подсчеты гиперссылок, кажется, должны слабо работать для них [47]. Это направление исследования также открыло, что основной точкой притяжения научных гиперссылок был скорее размер сетевого издания, а не его качество – с помощью ученых, производящих больше и лучше исследований, а также создающих больше сетевых страниц

и привлекающих больше гиперссылок [53]. Оно (направление) также обнаружило, что лучше подсчитывать внутрисвязанные сетевые сайты, чем индивидуальные гиперссылки [54].

Имеется два логических способа использования гиперссылок для замены импакт факторов журналов: подсчитать общее число гиперссылок на сетевой сайт журнала, деленное на количество, представляющее размер журнала или его сетевого сайта, или подсчитать среднее число гиперссылок, полученных каждой опубликованной статьей. Первое гораздо проще, но неясно, какой должна быть лучшая цифра, чтобы разделить подсчеты ссылок и, таким образом, все исследование, кажется, должно только подсчитать общие ссылки на сетевые сайты журналов. Ранние исследования показали, что внутри одной предметной области и относительно однородного массива журналов число гиперссылок на сетевой сайт журнала соотносится с его импакт фактором [55]. Это не может встречаться у неоднородных массивов журналов, например относительно некоторых с открытым доступом и различной тематикой [47]. Возраст сетевого сайта является другим важным фактором [56].

Несмотря на положительные результаты, обсужденные выше, кажется ясным, что ранние предсказания относительно того, что ссылки не будут заменены гиперссылками [57], подтвердились в том смысле, что до сих пор нет фактора влияния журналов на основе гиперссылок. Главной причиной, почему не предпринято серьезной попытки создания этого, является, вероятно, легкость, с которой можно манипулировать гиперссылками, и предположение, что эта манипуляция будет происходить, как только ссылки реально используются для помощи определения репутации журнала. В некоторой степени они сегодня используются для репутаций университетов, не только в ранжировании поисковых систем, но и вебметрическом ранжировании мировых университетов [58]. Тем не менее, практические проблемы с подсчетом ссылок на журналы растут из-за превращения многих журналов в сетевые сайты издателей и их непосредственной доступности в рамках внутренних услуг на основе подписки. В результате этого, по-видимому, имеется небольшая потребность в независимых сетевых сайтах для журналов и для ученых, чтобы ссылаться на такие сетевые сайты.

## СЕТЕВЫЕ ССЫАКИ

Одним способом построения фактора влияния с широким кругом охвата типов влияния должен быть подсчет ссылок не только из научных журналов, но и сети целиком. Так как сеть содержит коммерческую, научную, правительственную и другую информацию, такой фактор влияния должен охватывать типы влияния, игнорируемые фактором влияния журналов на основе науки. Некоторые исследования пытаются воспроизвести импакт фактор журналов с использованием сетевых данных таким образом, который приводит к многообещающим результатам. Термин сетевое цитирование был задуман, чтобы обозначать ссылку на определенную научную статью, взятую из сетевой страницы. Сетевые ссылки могут обычно отыскиваться путем построения соответствующих запросов и представления их поисковым системам. Например, такой запрос может содержать название статьи как поисковую фразу вместе с годом публикации и именем первого автора.

Сетевые ссылки для статей в отдельных журналах, как описано выше, показаны для значительной корреляции с традиционными ссылками Web of Science для многих дисциплин [59-61]. Сетевая ссылка - гиперсвязная ги-

бридная мера, унифицированный указатель ресурса ссылки, также показывается для значительной корреляции со ссылками [59]. Некоторые исследования пошли дальше и подсчитали только сетевые ссылки из определенных типов документов, таких как презентации прямого доступа [62], интерактивные программы [63], оцифрованные книги [64] и блоги [65], – всех, показывающих существенные корреляции с подсчетами ссылок. Значительные корреляции дают подтверждение тому, что различные типы подсчетов сетевых ссылок относятся к научному влиянию, но эти корреляции, как правило, не высокие и поэтому также обоснованно предположить, что различные источники могут давать показатели для отчасти различных типов влияния. Например, исследование электронных ссылок на журнальные статьи из оцифрованных книг через поиск в Google Book Search показало, что некоторые статьи имели высокое влияние, основанное на книгах, несмотря на среднее влияние, основанное на журнале [64]. Это может быть особенно справедливо для статей, ориентированных на гуманитарные области в силу важности книг в области гуманитарных наук.

Многие исследователи применили скрытый тип сетевых ссылок используя Google Scholar для подсчета ссылок на статьи [66-69]. Эти результаты значительно соотносятся со ссылками Web of Science [59] и объединяют ряд источников документов, включая, но не ограничиваясь, цифровыми библиотеками издателей. Это эффективно является гибридным сетевым/журнальным источником базы данных, но, кажется, сильно сфокусированным на научных источниках, и не позволяет автоматического поиска таким образом, который необходим для автоматических подсчетов фактора влияния для журналов, если только не обеспеченных Google.

Внешне различные типы подсчетов ссылок на основе сети имеют больший потенциал для использования в качестве альтернатив импакт фактору журналов. Тем не менее, существуют две проблемы. Во-первых, за исключением подсчетов ссылок Google Book Search, все выше указанные источники могут быть легко манипулируемыми, если адаптировать их к широко распространенным факторам влияния. Во-вторых, крупномасштабный подсчет сетевых ссылок будет поглощать время, чтобы сделать это хорошо, так как кажется невозможным автоматически построить запросы к поисковым системам, которые будут необходимы. Например, невозможно построить эффективный запрос для журнальной статьи с коротким, общим названием (например, «Атомь») и распространенной фамилией автора (например, J. Smith), особенно учитывая, что любая ссылка может быть написана группой различных стандартных форматов ссылок.

## ТВИТТЕР

Твиттер кажется логичным местом для определения научного влияния, так как он используется, чтобы рекомендовать статьи, которые читают, и некоторые ученые считают, что такие ссылки отражают научное влияние [70]. Кроме того, есть эмпирическое свидетельство того, что ссылки в Твиттере соотносятся с более поздним влиянием ссылок для одного журнала [71]. Тем не менее, имеются две практические проблемы с использованием Твиттера для метрик влияния. Во-первых, что касается большинства сетевых данных, они имеют возможность быть манипулируемыми и могут находиться под влиянием спама. Во-вторых, и что более характерно для Твиттера, ограничение длины текстовых сообщений до 140 символов является недостаточным для полных ссылок в большинстве случаев, и даже определение гиперссылок не является простым из-за использования корот-

кого закодированного сервиса сокращений ссылок (bit.ly) или других унифицированных указателей ресурсов.

Можно некоторым образом обойти практически проблемы со ссылками Твиттера, но некоторые из них не являются масштабными. Например, проблемы спама и определения релевантных сообщений в Твиттере (твитов) можно обойти с помощью ручного анализа контента твитов выбранных отдельных ученых [70]. Релевантные по ученым и предположительно свободные от спама твиты могут быть также собраны путем мониторинга отдельных научных хэштегов, таких как и для конференций [72]. Более обобщенно, наборы данных твитов ученых могут теоретически собираться через запросы для списков научных пользователей твиттера, научных хэштегов или общих поисковых терминов [73]. Из них первые два кажутся наиболее практичными, но оба будут полагаться на большие списки пользователей или хэштеги, которые, кажется, невозможно создать автоматически [72] и может потребоваться слишком много времени для их создания.

Полагая, что релевантные твиты определены, ссылки необходимо извлекать. Рекомендация для отдельной статьи вероятно будет в форме URL или непрямого короткого описания, такого как «статья, написанная X в текущем журнале *Nature*» или «статья о биофлавоноидах в №12 (т. 4) выпуска *Food Hygiene*» или «library searching article! #jasis». Самым простым решением будет сбор только тех твитов, которые упоминают хэштег определенного журнала и игнорируют все другие ссылки Твиттера. Это, однако, даст явное преимущество журналам с простыми и хорошо используемыми хэштегами и также не остановит проблему спама. Альтернативно могут быть подсчитаны только твиты, содержащие URL опубликованных статей. Так как URL, как правило, закодированы, этот подход необходимо сочетать с крупномасштабным сбором данных из Твиттера. Это необходимо, потому что URL могут быть найдены только декодированием коротких URL во всех твитах, чтобы найти твиты, релевантные статьям журнала. Этот подход успешно использовался для одного электронного журнала [71], но для крупномасштабного анализа значительную ручную работу необходимо начать с определения структур URL всех журналов для оценки. Этот метод предположительно благоприятствует электронным журналам и журналам с простыми структурами URL, которые читатели сочтут удобными для дальнейшего продвижения. Что касается других журналов, то читатели, более вероятно, могут сослаться на статью иными способами, чем включение URL.

Подсчет ссылок Твиттера кажется практичным, учитывая источник финансирования для определения структур журнальных URL, но произведенная метрика, вероятно, будет легкой для манипуляций, например, через неправильные бюджеты Твиттера, и ее возможный уклон в сторону электронных журналов кажется неизбежным.

## ОБСУЖДЕНИЕ И ЗАКЛЮЧЕНИЕ

Как упоминалось выше, важным принципом в развитии показателя является то, что, когда показатель становится широко признанным, вероятно, будут попытки им манипулировать. Когда манипуляция импакт факторами журналов производится издателями журналов, поощряющими самоцитирование журнала, она может быть обнаружена подсчетом долей самоцитирования журнала и тогда могут быть возбуждены санкции [74]. Наоборот любые данные на основе сети, вероятно, являются легко манипулируемыми и это может быть даже дешево, как показано индустрией оптимизации поисковых систем. Эта индустрия является также опытной в избегании обнаружения и, таким образом, кажется, что попытки выявить

манипуляцию не могут быть легко автоматизированы и поэтому должны быть дорогими для опробывания. Это, кажется, исключает все факторы влияния, использующие открытые данные сети, включая те, которые основаны на подсчетах гиперссылок и различных типах подсчетов сетевых ссылок. Двумя возможными исключениями являются факторы влияния загрузки и факторы влияния, использующие книжные ссылки в Google Book Search.

Любой фактор влияния, полагающийся на Google Book Search, будет, вероятно, нуждаться в поддержке от Google, чтобы реализоваться и должен быть основан на большом числе запросов с целью проверки книжных ссылок на статьи - один на опубликованную статью. Эти запросы также нуждаются в поддержке со стороны проверки людьми в случаях, когда соответствующие запросы не могут быть построены. Результаты, по-видимому, должны иметь особое значение для журналов, ориентированных на гуманитарные области, в предметных областях, сочетающих гуманитарные и социальные исследования, такие как дисциплина библиотекосведения и информатики. Это потому, что они должны помогать компенсировать преимущество импакт фактора социальных журналов по сравнению с гуманитарно-ориентированными журналами. Ограничение факторов влияния на основе книг, тем не менее, является таким, что книги могут медленно писаться и публиковаться по сравнению с журнальными статьями, и так называемый фактор влияния журнала на основе книг будет нуждаться в охвате более длинного публикационного периода - возможно пяти лет по умолчанию вместо двух лет - и будет менее подверженным изменениям со временем как результат. Если этот подход будет принят, тогда проверки в отношении манипуляции также будут необходимы. В частности, ссылки из издательств, выпускающих книги на средства авторов, и университетских внутренних издательств могут нуждаться в более пристальном изучении или исключении. Эта идея будет нуждаться в поддержке вероятного выгодоприобретателя, чтобы иметь успех - возможно в лице Google, получающего выгоду от публичности, или национального совета по финансированию гуманитарных областей, стремящегося защитить свою научную область.

Факторы влияния загрузки также могут быть манипулируемыми в некоторой степени. Эта манипуляция может быть прямой в форме авторов или редакторов, повторно загружающих свои собственные статьи, или в виде инструкций студентам или другим лицам, делающим то же самое для них. Она может быть не прямой с помощью обмана, используемого вирусами [75]: использование цифровых видеоформатов iFrames в несвязанных сетевых страницах для подготовки браузеров к загрузке отдельных статей без осознания этого пользователем. Защитой для загружаемых статистических данных является то, что они являются более многочисленными, чем ссылки, и, следовательно, больше нуждаются в корректировке, чтобы иметь влияние. С другой стороны, корректировка загрузки является слишком простой и менее отслеживаемой, чем создание ссылки в опубликованной журнальной статье. Другой практической проблемой для создания общего загружаемого фактора влияния для науки является потребность в сотрудничестве издателя и распространителя, чтобы гарантировать обоснованно точную и сравнимую статистику загрузки. Описанные выше проекты, такие как COUNTER и SUSHI, по-видимому, сделали это отдельной возможностью для будущего. Интуитивно, использование факторов влияния кажется, тем не менее, проблематичным из-за

вовлеченного коммерческого соперничества и потребности издателя в сотрудничестве, чтобы производить сравнительную статистику. Другим возможным путем является развитие локализованных, основанных на загрузке, факторов влияния для журналов, подсчитанных для специальных систем цифровых библиотек. Это представляется возможным, так как локализованный характер делает их менее соблазнительными для манипуляции, и существование таких средств может быть маркетинговой поддержкой для цифровой библиотеки, с издателями, следовательно, поддерживающими создание программного обеспечения, необходимого для их производства. Такие региональные факторы влияния могут даже сочетаться с множественными похожими сайтами (например, университетов), чтобы создавать национальные или даже международные общие версии, хотя и ограниченные только сравнением между журналами в рамках отдельного поставщика.

Наконец, альтернативный подход должен применять обосновано большое множество оценок альтерметрик с уверенностью, что будет гораздо труднее манипулировать множественными метриками, если их будет достаточно. Кажется, имеется широкий потенциал для создания различных альтерметрик и, таким образом, это является возможностью, но кажется нежелательным по некоторым причинам. Создание большого числа метрик, кажется, увеличит целую творческую попытку, сокращая пока время, потраченное на каждую индивидуальную метрику. В результате каждая индивидуальная метрика может поэтому быть слабее и легче для манипулирования до тех пор, пока творческое усилие не будет распределено некоторым образом.

В заключение можно сказать, что пока имеется ясная ситуация важности метрик на основе сети в создании факторов влияния для журналов, которые должны охватить более широкие типы влияния, чем текущие импакт факторы, существуют практические проблемы, которые, кажется, исключают большинство инициатив. Два подхода, которые кажутся возможными, хотя и с ограниченной областью применения для обоих и с финансовой поддержкой, необходимой для их реализации, представляют книжные ссылки через Google Book Search и локализованные факторы влияния загрузки.

## ЛИТЕРАТУРА

1. *Garfield E.* The agony and the ecstasy: The history and the meaning of the journal impact factor// Fifth International Congress on Peer Review in Biomedical Publication, in Chicago, USA, 2005. - <http://garfield.library.upenn.edu/papers/jifchicago2005.pdf>.
2. *Moed H. F.* Measuring contextual citation impact of scientific journals// *Journal of Informetrics*. — 2010. — Vol. 4, No.3. — P. 265-277.
3. *Vanclay J. K.* Impact Factor: Outdated artefact or stepping-stone to journal certification?// *Scientometrics*. (в печати)
4. *Cole S., Cole J. R.* Scientific output and recognition: A study in the operation of the reward system in science// *American Sociological Review*. — 1967. — Vol. 32, No.3. — P. 377-390.
5. *Garfield E.* Citation analysis as a tool in journal evaluation// *Science*. — 1972. — Vol.178, No.4060. — P. 471-479.
6. *Priem J., Hemminger B. M.* Scientometrics 2.0: Toward new metrics of scholarly impact on the social

Web// *First Monday*.— 2010.— Vol. 15, No.7. - <http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2874/2570>.

7. *Priem J., Taraborelli D., Groth P., Neylon C.* Altmetrics: A manifesto. — 2011. - <http://altmetrics.org/manifesto/>.
8. *Merton R. K.* The sociology of science. Theoretical and empirical investigations. — Chicago: University of Chicago Press, 1973.
9. *Newton I.* (1676; 1992 republication). Letter from Isaac Newton to Robert Hooke, 5 February 1676. / J.-P. Maury (Ed.)// *Newton: Understanding the Cosmos*. New York: New Horizons.
10. *Bornmann L., Daniel H.-D.* What do citation counts measure? A review of studies on citing behavior// *Journal of Documentation*. — 2008. — Vol. 64, No.1. — P. 45-80.
11. *Brooks T. A.* Evidence of complex citer motivations// *Journal of the American Society for Information Science*. — 1986. — Vol.37. — P. 34-36.
12. *Cronin B.* The citation process: The role and significance of citations in scientific communication. — London: Taylor Graham, 1984.
13. *Huizenga J. R.* Cold fusion: The scientific fiasco of the century. — Rochester, NY: University of Rochester Press, 1992.
14. *De Costa C. M.* "The contagiousness of childbed fever": A short history of puerperal sepsis and its treatment// *The Medical Journal of Australia*. — 2002.— Vol. 177, No. 11/12.— P. 668-671.
15. *Meyer M.* Academic patents as an indicator of useful research? A new approach to measure academic inventiveness// *Research Evaluation*. — 2003.— Vol. 12, No.1.— P. 17-27.
16. *Oppenheim C.* Do patent citations count? / B. Cronin & H. B. Atkins (Eds.)// *The web of knowledge: A festschrift in honor of Eugene Garfield* (pp. 405-432). — Metford, NJ: Inormation Today Inc. ASIS Monograph Series, 2000.
17. *RCUK.* Types of Impact. — 2011. - <http://www.rcuk.ac.uk/documents/impacts/TypologyofResearchImpacts.pdf>.
18. *Coleman A.* Assessing the value of a journal beyond the impact factor// *Journal of the American Society for Information Science & Technology*. — 2007. — Vol. 58, No.8. — P. 1148-1161.
19. *Whitley R.* The intellectual and social organization of the sciences (2 ed.). — Oxford: Oxford University Press, 2000.
20. *Becher T., Trowler P.* Academic tribes and territories (2ed). — Milton Keynes, UK: Open University Press, 2001.
21. *NISO.* ANSI/NISO Z39.93-2007 The Standardized Usage Statistics Harvesting Initiative (SUSHI) protocol. — 2011. - [http://www.niso.org/apps/group\\_public/project/details.php?project\\_id=2021](http://www.niso.org/apps/group_public/project/details.php?project_id=2021).
22. *COUNTER.* Counting Online Usage of Networked Electronic Resources. — 2011. - <http://www.projectcounter.org/>.
23. *COUNTER.* Register of vendors providing usage reports compliant with Release 3 of the Code of Practice for Journals and Databases. — 2011. - [http://www.projectcounter.org/r2013/R2013CV\\_Dec2011.pdf](http://www.projectcounter.org/r2013/R2013CV_Dec2011.pdf).
24. *Shepherd P.* Journal Usage Factor: Results, recommendations and next steps. — 2011. - <http://www.uksg.org/sites/uksg.org/files/JournalUsageFactorReport080711.pdf>.
25. *Brody T., Harnad S., Carr L.* Earlier Web usage statistics as predictors of later citation impact// *Journal*

of the American Society for Information Science and Technology. — 2006. — Vol.57, No.8.— P. 1060-1072.

26. *Moed H. F.* Statistical relationships between downloads and citations at the level of individual documents within a single journal// Journal of the American Society for Information Science & Technology. — 2005. — Vol. 56, No.10. — P. 1088-1097.

27. *Bollen J., Van de Sompel H.* Usage impact factor: The effects of sample characteristics on usage-based impact metrics// Journal of the American Society for Information Science and Technology. — 2008. — Vol. 59, No.1. — P. 136-149.

28. *Bogers T., Bosch A. v. d.* Recommending scientific articles using citeulike// Proceedings of the 2008 ACM conference on Recommender Systems (RecSys '08) (pp. 287-290). — New York, NY: ACM, 2008.

29. *Henning V., Reichelt J.* Mendeley - A last.fm for research?// IEEE Fourth International Conference on eScience (eScience '08) (pp. 327-328).— Los Alamitos: IEEE, 2008.

30. *Maxmen A.* Science networking gets serious// Cell. — 2010. — Vol. 141, No.3. — P. 387-389.

31. *Li X., Thelwall M., Giustini D.* Validating online reference managers for scholarly impact measurement// Scientometrics. (в печати)

32. *Haustein S., Siebenlist T.* Applying social bookmarking data to evaluate journal usage// Journal of Informetrics. — 2011. — Vol.5, No.3. — P. 446-457.

33. *Bar-Ilan J.* Articles tagged by 'bibliometrics' on Mendeley and CiteULike// Metrics 2011 Symposium on Informetric and Scientometric Research. — 2011.

34. *Larson R. R.* Bibliometrics of the World Wide Web: An exploratory analysis of the intellectual structure of cyberspace// ASIS 59th annual meeting. Baltimore, MD. — 1996.

35. *Rousseau R.* Sitations: an exploratory study// Cybermetrics. — 1997. — Vol. 1, No.1.- <http://www.cindoc.csic.es/cybermetrics/articles/v2001i2001p2001.html>.

36. *Ingversen P.* The calculation of Web Impact Factors// Journal of Documentation. — 1998. — Vol. 54, No.2. — P. 236-243.

37. *Rodríguez i Gairín J. M.* Valorando el impacto de la información en Internet: AltaVista, el "Citation Index" de la Red (Evaluating the impact of Internet information: Altavista, the "Citation Index" of the Web)// Revista Española de Documentación Científica. — 1997. — Vol. 20, No.2. — P. 175- 181.

38. *Brin S., Page L.* The anatomy of a large scale hypertextual Web search engine// Computer Networks and ISDN Systems. — 1998. — Vol. 30, No.1-7. — P. 107-117.

39. *Kleinberg J. M.* Authoritative sources in a hyperlinked environment// Journal of the ACM. — 1999. — Vo. 46, No.5. — P. 604-632.

40. *Chakrabarti S., Dom B., Raghavan P., Rajagopalan S., Gibson D., Kleinberg J. M.* Automatic resource compilation by analyzing hyperlink structure and associated text.// 7th International World Wide Web Conference.— 1998 (April).

41. *Chakrabarti S., Joshi M. M., Punera K., Pennock D. M.* The structure of broad topics on the Web. — 2002. - <http://www2002.org/CDROM/refereed/338>

42. *Chakrabarti S., VanDen Berg M., Dom B.* Focused crawling: A new approach to topic-specific Web

resource discovery// 8th International World Wide Web Conference. — 1999 (May).

43. *Han S., Ahn Y.-y., Moon S., Jeong H.* Collaborative blog spam filtering using adaptive percolation search// WWW2006Workshop.—2006.-<http://www.blogpulse.com/www2006-workshop/papers/collaborative-blogspam-filtering.pdf>.

44. *Gyongyi Z., Garcia-Molina H., Pedersen J.* Combating web spam with TrustRank// Proceedings of the thirtieth international conference on very large data bases. — 2004. — Vol. 30. — P. 576-587.

45. *Kousha K., Horri A.* The relationship between scholarly publishing and the counts of academic inlinks to Iranian university web sites: Exploring academic link creation motivations// Journal of Information Management and Scientometrics. —2004. — Vol. 1, No.2. — P. 13-22.

46. *Qiu J. P., Chen J. Q., Wang Z.* An analysis of backlink counts and Web Impact Factors for Chinese university websites// Scientometrics. — 2004. — Vol. 60, No.3.— P. 463-473.

47. *Smith A. G.* A tale of two web spaces; comparing sites using Web Impact Factors// Journal of Documentation. — 1999. — Vol. 55, No.5. — P. 577-592.

48. *Thelwall M.* Extracting macroscopic information from web links// Journal of American Society for Information Science and Technology. — 2001. — Vol. 52, No.13. — P. 1157-1168.

49. *Vaughan L., Thelwall M.* A modeling approach to uncover hyperlink patterns: The case of Canadian universities// Information Processing & Management.— 2005.— Vol. 41, No.2.— P. 347-359.

50. *Li X., Thelwall M., Wilkinson D., Musgrove P. B.* National and international university departmental web site interlinking, part 2: Link patterns// Scientometrics. — 2005. — Vol. 64, No.2. — P. 187-208.

51. *Thomas O., Willett P.* Webometric analysis of departments of librarianship and information science// Journal of Information Science. — 2000. — Vol. 26, No.6. — P. 421-428.

52. *Barjak. F., Thelwall M.* A statistical analysis of the web presences of European life sciences research teams// Journal of the American Society for Information Science and Technology. — 2008. — Vol. 59, No.4. — P. 628-643.

53. *Thelwall M., Harries G.* Do the Web sites of higher rated scholars have significantly more online impact? // Journal of the American Society for Information Science and Technology. — 2004. — Vol.55, No 2.— P. 149-159.

54. *Thelwall M.* Conceptualizing documentation on the Web: An evaluation of different heuristic-based models for counting links between university web sites// Journal of American Society for Information Science and Technology. — 2002. — Vol. 53, No.12. — P. 995-1005.

55. *Vaughan L., Hysen K.* Relationship between links to journal Web sites and impact factors// ASLIB Proceedings. — 2002.—Vol. 54, No.6.— P. 356-361.

56. *Vaughan L., Thelwall M.* Scholarly use of the Web: What are the key inducers of links to journal Web sites?// Journal of American Society for Information Science and Technology.— 2003.— Vol. 54, No.1.— P. 29-38.

57. *van Raan A. F. J.* Bibliometrics and Internet: Some observations and expectations// Scientometrics.— 2001.— Vol. 50, No.1.— P. 59-63.

58. *Aguillo I. F.* Measuring the institution's footprint in the web// Library Hi Tech. — 2009. — Vol. 27, No.4.— P. 540- 556.

59. *Kousha K., Thelwall M.* Google Scholar citations and Google Web/URL citations: A multi-discipline exploratory

- analysis// Journal of the American Society for Information Science and Technology. — 2007. — Vol. 58, No.7. — P. 1055-1065.
60. *Vaughan L., Shaw D.* Bibliographic and Web citations: What is the difference? // Journal of the American Society for Information Science and Technology. — 2003. — Vol. 54, No.14. — P. 1313-1322.
61. *Vaughan L., Shaw D.* Web citation data for impact assessment: A comparison of four science disciplines// Journal of the American Society for Information Science & Technology. — 2005. — Vol.56, No.10. — P. 1075-1087.
62. *Thehwall M., Kousha K.* Online presentations as a source of scientific impact?: An analysis of PowerPoint files citing academic journals// Journal of the American Society for Information Science and Technology.— 2008.— Vol. 59, No. 5.— P. 805-815.
63. *Kousha K., Thehwall M.* Assessing the impact of disciplinary research on teaching: An automatic analysis of online syllabuses// Journal of the American Society for Information Science and Technology. — 2008. — Vol. 59, No. 13. — P. 2060-2069.
64. *Kousha K., Thehwall M.* Google Book Search: Citation analysis for social science and the humanities// Journal of the American Society for Information Science and Technology. — 2009. — Vol. 60, No. 8. — P. 1537-1549.
65. *Kousha K., Thehwall M., Rezaie S.* Using the Web for research evaluation: The Integrated Online Impactindicator// Journal of Informetrics. — 2010. — Vol. 4, No. 1. — P. 124-135.
66. *Jasó P.* Google Scholar: the pros and the cons// Online Information Review. — 2005. — Vol., 29, No.2. — P. 208-214.
67. *Mayr P., Walter A. K.* An exploratory study of Google Scholar// Online Information Review. — 2007. — Vol. 31, No.6. — P. 814-830.
68. *Mebo L. I., Yang K.* Impact of data sources on citation counts and rankings of LIS faculty: Web of Science vs. Scopus and Google Scholar// Journal of the American Society for Information Science and Technology. — 2007. — Vol. 58, No.13. — P. 2105-2125.
69. *Vaughan L., Shaw D.* A new look at evidence of scholarly citation in citation indexes and from web sources// Scientometrics. — 2008. — Vol.74, No.2. — P. 317-330.
70. *Priem J., Costello K. L.* How and why scholars cite on Twitter// Proceedings of the American Society for Information Science and Technology (ASIST 2010). — 2010 — Vol. 47. — P. 1-4.
71. *Eysenbach G.* Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact// Journal of Medical Internet Research.— 2011.— Vol.13, No.4, e123.
72. *Weller K., Dröge E., Puschmann C.* Citation analysis in Twitter: Approaches for defining and measuring information flows within tweets during scientific conferences// 1st Workshop on Making Sense of Microposts.— 2011. - [http://sunsite.informatik.rwthachen.de/Publications/CEUR-WS/Vol-2718/paper\\_2004.pdf](http://sunsite.informatik.rwthachen.de/Publications/CEUR-WS/Vol-2718/paper_2004.pdf).
73. *Weller K., Puschmann C.* Twitter for scientific communication: How can citations/references be identified and measured?// Proceedings of the Poster Session at the Web Science Conference 2011. - <http://journal.webscience.org/2500/>.— Koblenz, Germany.— 2011.
74. *Davis P.* Gaming the Impact Factor puts journal in time-out. — 2011. - <http://scholarlykitchen.sspnet.org/2011/2010/2017/gaming-the-impact-factor-putsjournal-in-time-out/>.
75. *Provos N., Mavrommatis P., Rajab M. A., Monroe F.* All your iFRAMEs point to us. — 2008. - [http://static.googleusercontent.com/external\\_content/untrusted\\_dlcp/research.google.com/en//archive/provos-2008a.pdf](http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en//archive/provos-2008a.pdf).

# Сущности и идентификаторы в исследовательских информационных системах\*

**Бригитта ЙЁРГ**  
(Brigitte JÖRG)

Немецкий научный центр искусственного  
интеллекта (DFK GmbH), г. Берлин, Германия;  
euroCRIS

**Торстен ХЁЛЛРИГЛ**  
(Thorsten HÖLLRIGL)

Фирма AVEDAS AG, г. Карлсруэ, Германия

**Мигель-Анжел СИСИЛИЯ**  
(Miguel-Angel SICILIA)

Университет Алькала, г. Алькала-де-Энарес,  
Испания; euroCRIS

*Поскольку наука всё больше становится глобальным бизнесом и к тому же с инициативами открытых данных и доступа, срочно необходимы средства для определения и, таким образом, для соединения внутренних и внешних для системы сущностей. Необходимость идентификаторов за пределами системы является, поэтому, глобальным требованием, а также соответствующим в границах организации, охватывающей множество систем. Различные инициативы идентификаторов и систем появляются в научной области и за ее пределами. Однако они пока не достигли необходимой межоперабельности. В этой статье мы намерены собрать дополнительные важные ингредиенты для разработки поддерживаемой форматом CERIF подмодели идентификатора, вносящей вклад в продолжающиеся дискуссии и поддерживающей проектные решения. Здесь мы, во-первых, представляем общие проблемы, связанные с идентификаторами в широкой научной области, а затем анализируем имеющиеся системы, технологии и современные инициативы для устранения пробела глобальной идентичности.*

## ВВЕДЕНИЕ

Сегодня большинство информационных систем все еще создается на основе реляционных систем управления базами данных (реляционных СУБД), где таблицы являются физическими контейнерами для описания объектов реального мира и их связей. Обычно они разрабатывались на основе моделей взаимосвязей сущностей для описания определенной области интереса, часто отражая концептуальные модели области [1]. «Концептуальное моделирование – это деятельность по формальному описанию некоторых областей окружающего нас физического и социального мира в целях понимания и коммуникации» [2]. Хотя концептуальные модели хорошо поддерживают описания мира интересов и, следовательно, объектов внутри него, они до сих пор не имеют многого, касающегося идентификации самих описываемых объектов [3]. Реляционные информационные системы определяют объекты через идентификаторы внутри системы, и они включают поддержку исходной целостности для сбора информации, содержащейся в распределенных таблицах, соотнося ее с объектом реального мира (например, организацией, личностью и т.п.). Объект реального мира может создаваться во время включе-

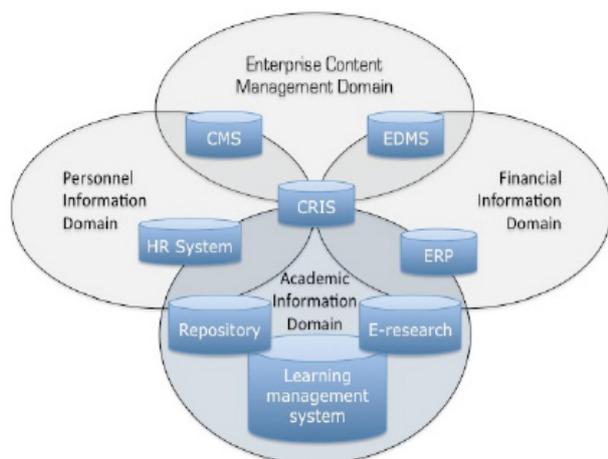
ния информации, либо обмена ею, во время запроса или посредством применения правил, ограничения системы, либо накопленных запросов в интерфейсах пользователей, либо предварительно выбранных видов.

Внутрисистемные идентификаторы хорошо работают в рамках системы для определения и накопления информации об объектах, но они не определяют масштаба использования систем. Поскольку наука все больше становится глобальным бизнесом, научные среды и, следовательно, научная информация становятся более открытыми и доступными; релевантные информационные объекты повторяются во множестве систем, поддерживающих совместно используемые описания.

Средства для описания и, следовательно, связи внутрисистемных и внешних сущностей являются, таким образом, необходимыми как в глобальном масштабе, так и организационных границах, охватывающих множество систем. Например, человек и первые имена этого человека или фамилии должны поддерживаться в системе человеческих ресурсов, а также человек должен быть известен в проектной управляющей системе, хранилище и финансовой системе (см. рис.). В каждой из этих систем один и тот же человек, наиболее вероятно, будет иметь свой внутрисистемный идентификатор, и в каждой системе описания имени могут иметь орфографические варианты записей в другом написании (например, B. Jörg, Brigitte Jörg, B. Joerg). Наиболее часто идентификация объектов в информационных системах зависит от инсти-

\* Перевод Jörg B., Höllrigl T., Sicilia M.-A. Entities and identities in research information systems. – [http://eprints.rclis.org/17176/1/Joerg-Entities and Identities CRIS 2012.pdf](http://eprints.rclis.org/17176/1/Joerg-Entities%20and%20Identities%20CRIS%202012.pdf)

туциональных и юридических установок и правил [4,5]. Следовательно, необходим механизм управления идентичностью, чтобы обеспечить, например, управление различными идентификаторами [6,7].



Расширенная модель исследовательской информационной области [8, с.49]

Современные исследовательские информационные системы (CRIS – Current Research Information Systems) были признаны центром структуры научно-информационной межоперабельности [8], выполняющим решающую роль в распределении информации и обеспечении необходимой гибкости для удовлетворения потребностей множества заинтересованных сторон.

Тем не менее, по технико-историческим причинам они допускали полноту информации в границах системы и все еще не четко учитывали различного рода идентификаторы. История CRIS в Европе тесно связана с общеевропейским научно-информационным форматом (CERIF – Common European Research Information Format). Сущность *eElectronicAddress* иногда применялась с форматом CERIF\* в качестве вспомогательного средства для идентификаторов. Тем не менее, в сообществе все еще продолжают дискуссии по поводу того, схожи семантический электронный адрес с понятием идентификатора, и, следовательно, может или не может он использоваться как таковой. Расширение модели CERIF до сущности *eFederateIdentifier* было согласовано, вызвав проектные решения, фактически, такой сущности, которая, будучи физически встроенной в закрытый мир информационной системы, является концептуально открытой для мира, т.е. идентификатором любой информационной системы. Недавно было предложено расширение модели CERIF для связи систем CRIS через сущности уровня детализации (LOD – level of detail) [9], хотя необходимы дальнейшие дискуссии относительно того, в какой степени расширение LOD CERIF объясняет общую подмодель идентификатора CERIF.

Необходимость многочисленных или глобальных идентификаторов объектов в научных системах очевидна. Масштабная и, следовательно, поддерживаемая межоперационная структура или модель, подобная CERIF, должна объяснять идентификацию объекта за рамками системы и через системы. Сейчас мы, в первую очередь, представляем общие проблемы идентификаторов в широкой научной области, прежде чем анализировать доступные системы, технологии и текущие инициативы, направленные на решение проблемы глобальной иден-

тичности. Поэтому мы исследуем системы идентификаторов и инициативы в широкой научной области.

## ПРОБЛЕМЫ ИДЕНТИФИКАТОРА В ИССЛЕДОВАТЕЛЬСКИХ ИНФОРМАЦИОННЫХ СИСТЕМАХ

Из раздела «Введение» мы узнали, что исследовательские информационные системы в этом контексте рассматриваются не в очень строгом смысле, а охватывают широкую сферу, которую авторы работы [8] называют *расширенной исследовательской информационной областью* (см. рис.). Далее мы продемонстрируем необходимость уникальной идентификации, представляя известные случаи использования.

### Случай использования для исследователя

Внутри учреждения исследователи обычно идентифицируются с помощью локальной уникальной строки, или номера. Однако большей частью этот идентификатор является уникальным только в определенной системе или службе, и часто каждая организационная единица, такая как вычислительный центр, библиотека или администрация университета, создает и поддерживает свои собственные уникальные идентификаторы для каждого работника с его собственными идентификационными признаками. Для совокупного представления, например, продукции исследователя, он должен единообразно идентифицироваться во всех системах, имеющих его публикации, и необходимо энергично браться за осуществление следующих задач:

- Во-первых, должен быть создан идентификатор, который будет уникальным не только в контексте одной системы, но и во всей организации или даже за пределами организации.

- Во-вторых, различные опознаватели (идентификаторы) одного лица должны быть связаны друг с другом.

Мотивации иметь уникальные персональные идентификаторы, обращающиеся в организации, различны. Понятие идентификационного номера, который однозначно определяет личность за пределами учреждения, является общим в правительственной сфере, т.е. национальным идентификационным номером. Правительства используют такие идентификационные номера для однозначного определения и слежения за гражданами, например, в целях работы [10]. Однако такие идентификаторы – например, цифровой авторский идентификатор (DAI – Digital Author Identifier) в Нидерландах или номер социальной защиты (SSN – Social Security Number) в Норвегии – часто неизвестны за пределами национальных систем и обычно не используются исследователями.

Известными технологиями, которые поддерживают использование уникальных идентификаторов, являются так называемые системы единственной записи (SSO – Single Sign-On). Они позволяют устанавливать персональную подлинность (аутентифицировать) однажды, и впоследствии использовать внешне- или внутриорганизационные услуги без повторных регистраций. Для связи или различного рода учета на *федеративных* началах необходимо достижение SSO. Открыто доступным средством для поддержки аутентификации, основанным на федеративных решениях SSO, является, например, Shibboleth (<http://shibboleth.internet2.edu/>).

Общая потребность в уникальной идентификации исследователей состоит в их роли как авторов. Отыскивание всех публикаций определенного ученого на основе его заданных имени собственного и фамилии не приводит к удовлетворительному и полному перечню всех его публикаций, поскольку они не являются уникальными

\* Общеевропейский научно-информационный формат (CERIF); Рекомендация для стран – членов ЕС: <http://cordis.europa.eu/cerif/>, <http://www.euroCRIS.org/>

для всех авторов в мире. Более того, весьма вероятно, что фамилия или даже имя собственное лица может измениться. Система идентификации автора – уникальная и устойчивая во времени – является одним из наиболее важных требований и долговременной проблемой. Уникальный идентификатор должен пройти через сравнительные записи на подверженность ошибкам, основанные на (библиографических) строках, и публикации должны быть снова соединены с идентификаторами авторов. Это будет весьма полезно также для массивов данных (например, CRISPool; <http://www.crispool.org/>) и позволит избегать создания новых идентификаторов при объединении записей с новыми массивами.

### **Случай использования для описания публикаций**

В научной области уникальная идентификация публикации является важной. Когда ученый публикует статью, он хочет сделать эту публикацию доступной для цитирования, поскольку число ссылок считается общей мерой признания его работы. Фактически, вся научная деятельность строится на опубликованных результатах, и поэтому предшествующая работа и, следовательно, публикация должны сделаться доступными посредством однозначной ссылки. Сегодня обычным считается делать публикации доступными через всемирную сеть WWW с дополнительным требованием долговременного хранения. Следовательно, от системы требуется хранить связи с публикациями постоянно, обеспечивая как уникальный указатель ресурсов URL, так и идентификатор, который является единым для издателя и независимой системы (т.е. постоянным идентификатором). Таким образом, это также необходимо потому, что URL, который используется для связи с определенной публикацией, должен функционировать, даже если местонахождение сервера, где хранится публикация, было изменено.

Уникальные идентификаторы публикаций сопровождаются не только ссылками и местом хранения, но, кроме того, запросами или обменом, а также системными перемещениями. Обычно различные публикационные порталы, такие как PubMed, бесплатная база данных Национальной медицинской библиотеки США (<http://www.ncbi.nlm.gov/pubmed/>), электронный портал Немецкой национальной библиотеки (DNB; <http://www.dnb.de/eng/index.htm>), база данных Web of Science корпорации Thomson Reuters (<http://wokinfo.com/productstools/multidisciplinary/webofscience>) или база данных Scopus издательства Elsevier (<http://www.info.sciverse.com/scopus>), хранят информацию о публикациях и обеспечивают интерфейсы для загрузки метаданных (в некоторых случаях с ограничениями). На основе этих метаданных ученые могут импортировать информацию о своих публикациях, а также ссылках. Чтобы однозначно идентифицировать одну и ту же публикацию в различных электронных порталах, создается единый идентификатор, поскольку иначе идентификация описаний зависит от сравнений названий публикаций или соответствующих авторских имен, которые подвержены ошибкам, бывают неоднозначными, и поэтому просто не обнаруживаются.

### **Случаи использования для организаций**

Приписывание уникальных и постоянных идентификаторов требуется не только публикациям или ученым, но также полезно для описания организаций. Финансирующие организации, а также научные проекты (например, анализ научной области [11]) или осуществление оценки определенно получают выгоду от глобального

уникального идентификатора организации, так как он должен значительно улучшить качество данных. Европейская комиссия ввела так называемые коды идентификации стран – членов ЕС (PIC – Participant Identification Codes; [http://cordis/europa/eu/fp7/pp-pic\\_en.html](http://cordis/europa/eu/fp7/pp-pic_en.html)) для седьмой рамочной программы (Seventh Framework Programme). Они все еще являются мандатными для всех организаций, готовящих представление документов в соответствии с предложением. Коды PIC позволяют однозначно идентифицировать организации и их подразделения, тогда как у представляющих документы единицы нет необходимости повторять связанную с организацией информацию при каждом представлении. Код PIC не является публичным.

Представленные случаи использования являются обычными в научной области, а выбранные сущности – лицо, публикация и организация, т.е. деятели и продукция – определенно являются самыми важными в смысле научных параметров, для которых однозначная идентификация – это, несомненно, вклад в качество. Важно отметить, что глобальные идентификационные системы требуют не просто технологических решений, но также необходимого управления. Впоследствии исследованные инициативы, главным образом, обеспечивают собственную технологию для реализации своей управляющей миссии.

## **ГЛОБАЛЬНО ОДНОЗНАЧНЫЕ ИДЕНТИФИКАТОРЫ – ИНИЦИАТИВЫ И СИСТЕМЫ**

В научном мире ликвидация глобального пробела в идентификаторах признается решающим моментом для улучшения качества информационных систем, и в то же время для того, чтобы дать возможность крупномасштабного распространения информации или многократного ее использования. Журнал *Science* посвятил идентификации ученых исчерпывающую статью «Готовы ли Вы стать номером?» [10], и в этом отношении были предприняты многочисленные международные инициативы.

Жизнь могла бы стать намного легче для вовлеченных заинтересованных сторон, если бы научные сущности, такие как публикации, организации и люди, имели бы уникальные идентификационные номера. Действия, направленные на глобальный идентификатор ученого, развиваются с большой скоростью. Идентификатор ученого (ResearcherID) появился первым [10]; более поздний подход представляет ORCID (см. ниже). Если идентификаторы авторов, сотрудников или ученых становятся наиболее интересными сейчас, то инициативы идентификатора публикаций начались больше десятилетия назад с работы CrossRef – агентства по регистрации официального идентификатора цифрового объекта (DOI – Digital Object Identifier). Примерно в то же время в библиотечной области был инициирован виртуальный международный нормативный файл (VIAF - Virtual International Authority File), а десятью годами раньше стартовала широко распространенная система идентификаторов, а именно Handle. За пределами научной области существуют всеобщие уникальные идентификаторы (UUID – universally unique identifiers), унифицированные указатели ресурсов (URL – Uniform Resource Locator) или OpenID. Наш перечень анализируемых инициатив и систем не претендует на полноту. Мы (авторы) пытаемся обеспечить обзор самых общеизвестных и популярных инициатив и систем для рассмотрения возможного использования в научной области.

## Открытый идентификатор ученого и автора (ORCID – Open Researcher & Contributor ID)

Наиболее выдающейся инициативой является, определенно, ORCID (<http://about.orcid.org/>) – открытый идентификатор ученого и автора (сотрудника), принятая в конце 2009 г. «для решения проблемы неоднозначности имени автора в научной коммуникации». В августе 2010 г. была создана ведущая некоммерческая организация, членство в которой быстро растет. Инициатива ORCID намеревается запустить свои услуги во втором квартале 2012 г. Она лицензировала код Thomson Reuters ResearcherID в августе 2011 г. и будет обеспечивать прикладной программный интерфейс запросов для:

- **Bio:** определенный автор, использующий имена и сопутствующие данные
- **Works:** определенный автор, использующий работы, которые были вкладом в ...
- **Work:** определенная работа, включающая авторов, ответственных за нее
- **Search:** определенные любые метаданные, содержащие ранжированный перечень потенциальных авторов, идентифицируемых по этим метаданным.

Если первая фаза касается создания службы, то вторая фаза нацелена на устранение двусмысленности обеспечиваемых данных и приведение их к единообразию с применением автоматических средств. Инициатива ORCID финансируется за счет организационных членских взносов и стремится быть бесплатной для отдельных ученых [12].

## Идентификатор ученого (ResearcherID) корпорации Thomson Reuters

Идентификатор ученого – Thomson Reuters's ResearcherID (<http://www.researcherid.com/>) – был официально запущен на указанном сайте в январе 2008 г. [10]. Его цель – приписывание уникального идентификатора каждому автору, который участвует и для которого он намерен создать индекс ученого. Идентификатор был одобрен для использования, поскольку был первой глобальной схемой, готовой и доступной для идентификации ученого. В настоящее время ResearcherID – бесплатная услуга для многодисциплинарного научного сообщества, но в сообществе появляется интерес к тому, как в дальнейшем будут определяться доступ и стоимость. Идентификатор ResearcherID открывает доступ к измерениям числа ссылок, поиску и связыванию с соавторами и выгрузке публикаций из базы данных Web of Knowledge или EndNote Web. Профили этого идентификатора содержат многочисленные варианты имен; эта публичная служба дает ответы на запросы по фамилии и заданным именам, по учреждению или стране, по ключевым словам. Идентификатор ResearcherID обозначается, например, как A-1026-2007.

## Инициатива CrossRef

Инициатива CrossRef (<http://www.crossref.org/>) начала действовать в 2000 г. благодаря некоммерческой независимой организации, созданной ведущими научными издателями мира и именуемой ассоциацией PILA (Publishers International Linking Association). Эта инициатива возникла в 1999 г. на основе комбинации элементов из двух проектов.

Если первая инициатива была больше услугой библиографической связи с использованием идентификатора цифрового объекта (DOI), то последняя инициатива была больше коалицией издателей с критической массой для запуска, роста и поддержки такой системы,

т.е. обеспечивающей бизнес-модель. Первоначальная миссия библиографической связи, или DOI, была позже расширена «для возможности легкой идентификации и использования надежного электронного контента путем содействия совместной разработке и применению поддерживающей инфраструктуры» [13].

DOI – это уникальная буквенно-цифровая строка, CrossRef связывает с каждым DOI множество основных метаданных, а URL указывает на полный текст в сети. Например, DOI 10.1007/978-3-642-05290-3\_91 через систему CrossRef (<http://dx.doi.org/>) разрешает доступ к URL, позволяющему доступ к полному тексту публикации: <http://www.springerlink.com/content/m574117014g75661/>.

## Система обработки Handle

Хранилища публикаций, такие как DSpace (<http://www.dspace.org/>), Eprints (<http://www.eprints.org/software/>), Fedora (<http://www.fedora/commons.org/>), используют систему обработки Handle (<http://www.handle.net/>) для однозначного опознавания описаний публикаций. Система обработки включает некоторые характеристики для разрешения обработки информации в целях размещения, доступа, контакта, опознавания личности и других случаев использования ресурса. Связанная с обработкой информация может быть изменена при необходимости без изменения самого идентификатора. Система связи с каждым новым фрагментом информации Handle разработана корпорацией CNRI (Corporation for National Research Initiatives), некоммерческой организацией, чтобы предпринимать, содействовать и продвигать исследование в интересах общественности.

Инфраструктура системы обработки обеспечивается предварительной регистрацией и гонорарами службе, основным вкладчиком является Международная организация DOI. Первоначально система Handle предназначалась для разработки проекта соответствующей инфраструктуры цифровых библиотек, а немного позже она была дополнительно профинансирована со стороны Агентства передовых оборонных исследовательских проектов США (DARPA) [14]. Системы обработки использовались в цифровых водоизмерительных приложениях, приложениях GRID, а также хранилищах, регистраторах и т.д. В именном пространстве обработки каждый идентификатор состоит из двух частей: префикс обработки (предварительно названный идентификатором личности) и суффикс, или уникальное «локальное имя» при префиксе, отделенное косой чертой: 10.1045/january2010-reilly.

## Виртуальный международный нормативный файл (VIAF)

Виртуальный международный нормативный файл (VIAF – Virtual International Authority File; <http://www.oclc.org/viaf/>) управляется OCLC (Online Computer Library Center, Inc.) – международной службой библиотек всего мира, объединенных для обеспечения доступа к крупным нормативным файлам мира. Файл VIAF был инициирован Библиотекой конгресса США, Немецкой национальной библиотекой, Национальной библиотекой Франции и OCLC. Сам файл выглядит как строительный блок семантической сети. Наиболее крупные библиотеки поддерживают перечни имен людей, корпораций, конференций и географических мест, а также перечни для контроля работ и других сущностей – они называются нормативными (авторитетными) файлами, которые VIAF предназначает для включения и содействия многократному использованию. Пример-

ная запись VIAF RDF (<http://www.oclc.org/developer/documentation/virtual-international-authority-file-viaf-rdf-example>) собирает и отражает многочисленные схемы наименований и словари, помимо VIAF. Мы пока еще не исследуем Международный стандартный идентификатор имен (ISNI – International Standard Name Identifier; <http://www.isni.org/>), который первоначально основывался на VIAF и ISO Standard (ISO 27729).

### Унифицированные идентификаторы ресурсов (URI)

Обозримость семантической сети по отношению к документам за пределами сети осуществляется через связь данных посредством применения основного ряда принципов, а именно использования унифицированных идентификаторов ресурсов (URI – Uniform Resource Identifiers) как названий вещей. Связанные данные отсылают к URI как глобальным идентификаторам [15]. Технически URI – это структурированная строка знаков, применяемая для опознания имени или ресурса. Синтаксис URI включает схему URI (например, `http`, `ftp`, `file`) с последующим знаком двоеточия и затем специальной частью схемы. Идентификатор URI может быть относительным или абсолютным, например: `resource.txt` или `http://example.org/resource.txt`. Фактически, мы создаем новые идентификаторы URI в связи с каждым новым фрагментом информации, который мы храним, давая ему имя.

### Универсально уникальные идентификаторы (UUID)

Организация OSF (Open Software Foundation) рекомендовала использование так называемых универсально уникальных идентификаторов (UUID – Universally Unique Identifiers) в виде программного обеспечения. «Любой человек может создать UUID и использовать его для идентификации чего-либо с полной уверенностью, что один и тот же идентификатор никогда не будет создан кем-либо неумышленно для идентификации чего-либо еще» (Википедия). Организация OSF была некоммерческой организацией, основанной в 1988 г. в США для создания открытого стандарта по внедрению операционной системы UNIX.

В 1994 г. OSF заявила новую организационную модель, знаменуя окончание своей разработки программного обеспечения. Однако UUID все еще широко используется. UUID – это 16-байтовый (128-битовый) номер. В своей канонической форме он представлен 31-й гексадесятичной цифрой, разделенной на пять групп знаком тире, всего из 36 знаков в форме 8-4-4-4-12, например: 550e8400-e29b-41d4-a716-446655440000 (Википедия).

### Открытый идентификатор (OpenID)

OpenID Foundation (OIDF) – это международная некоммерческая организация индивидуумов и компаний, отвечающая за предоставление, продвижение и защиту технологий OpenID. Основанная в июне 2007 г. эта организация служит как доверительная организация, представляющая открытое сообщество разработчиков, поставщиков и пользователей и содействующая сообществу обеспечением инфраструктуры и помощи в продвижении и поддержании расширенного применения OpenID. Идентификатор OpenID – это децентрализованный протокол аутентификации (установления подлинности личности), который делает его легким для людей при записи и учете доступа к сети. Среди спонсирующих членов – Google, Microsoft, PayPal. Идентификатор OpenID не является специально адресованным научной области, а действует в общем информационном пространстве.

## АНАЛИЗ И ПОДТВЕРЖДЕНИЕ

Представленные в статье инициативы имеют похожие цели – стремление к глобально уникальной системе идентификаторов и, следовательно, к сетевой научной информационной инфраструктуре или научному информационному пространству. Охват каждой инициативы незначительно отличается в зависимости от ее истории. CrossRef больше концентрируется на научном результате, хотя стремится к более общему охвату, тогда как ORCID и ResearcherID имеют в центре внимания ученого, связанного с результатом. Инициатива Handle исходит из библиотечной области в направлении поддерживаемой, доступной, надежной и богатой метаданными инфраструктуры, как и VIAF, обеспечивающая строительные блоки, т.е. нормативные файлы. Инициативы ORCID, ResearcherID и CrossRef с DOI четко направлены на *идентификацию* своих имен, однако отличаются от URL, UUID и OpenID, которые больше направлены на технологию, хотя первые скорее сосредотачиваются на самих научных сущностях.

## РЕЗЮМЕ И ЗАКЛЮЧЕНИЕ

Из описанных здесь случаев использования очевидно необходимость и, следовательно, мотивация глобально уникальной и постоянной идентификации научных сущностей. Специальная группа CERIF начала дискуссии по наименованию и концептуальной формализации, т.е. моделированию в CERIF сути федеративного идентификатора. Тем не менее, конструкция и концепция еще не полностью ясны. Здесь мы исследовали системы идентификаторов и инициативы в широкой научной области, чтобы лучше понять ключевые проблемы, связанные с идентификацией.

Идентификация понимается как важная сущность, и этот вклад должен быть направлен специальной группой CERIF для продолжающейся дискуссии и понимания выпуска предстоящей модели CERIF. Идея глобально уникальных идентификаторов возбуждает, но вопросы управления и защиты не должны упускаться из вида. Мы поддерживаем Клиффорда Линча, директора Коалиции сетевой информации (CNI; <http://www.cni.org/>) в том, что отсутствие взаимосвязи удивляет [10], но «нет ничего неправильного относительно допущения развития двух систем – говорит он, – если они всегда могут быть связаны или объединены позже».

*Благодарность.* Мы хотим поблагодарить специальную группу CERIF за стимулирование и проведенные на высоком уровне дискуссии и поддерживающие вклады. Эта работа была частично поддержана также Фондом ЕС по соглашению о грантах № 249119.

## ЛИТЕРАТУРА

1. Wand Y., Weber R. Research commentary: Information systems and conceptual modeling – A research agenda // Information Systems Research. – 2002. – Vol.13, No.4. – P.363-376.
2. Mylopoulos J. Conceptual modeling and telos / P. Loucopoulos, R. Zicari (Eds.) // Conceptual modeling, databases and cases. – New York et. al.: John Wiley & Sons, Inc., 1992.
3. Evermann J., Wand Y. Towards ontologically based semantics for UML constructs / Kunii H.S, Jojodia S., Sölvberg A. (Eds.) // Proceedings: ER2001. – P. 354–367, 2001.
4. Hoellrigl T., Schell F., Hartenstein H. Federated and service-oriented identity management at a university // Proceedings of the 14th European University Information Systems (EUNIS 2008). – P. 59. – Denmark: Aarhus, June, 2008.

5. *Bischof C., Bunsen G., Müller J.* Die RWTH-Kundennummer – Eine einfache aber robuste Identitätsnummer, DFN Mitteilungen 72. – 2007. - P 19-21.
6. *Hoellrigl T., Schell F., Hartenstein H.* Federated identity management as a basis for integrated information management // Information Technology. – 2009. - Vol. 51, No.1. – P. 14-23.
7. *Hoellrigl T., Dinger J., Hartenstein H.* FedWare: Middleware services to cope with information consistency in federated identity management // Proceedings of the Fifth International Conference on Availability, Reliability and Security (ARES 2010). – Poland: Krakow, February, 2010.
8. *van Godtsenhoven K., Karstensen M.E., Sierman B., Bijsterbosch M., Hochstenbach P., Russell R. and Vanderfeesten M.* Emerging standards for enhanced publications and repository technology: Survey on technology. - Amsterdam: Univ. Press, SURF/EU Driver Series, 2009
9. *Jörg B., Ruiz-Rube I., Sicilia M-A., Dvorak J., Jeffery K., Höllrigl T., Rasmussen H.S., Engffer A. Vesterdam T., Garcia Barriocanal E.* Connecting closed world research information systems through the linked open data web // International Journal of Software Engineering and Knowledge Engineering (IJSEKE). – 2012. - Vol. 22 (Consuming and Producing Linked Data on Real World Applications. June, 2012.).
10. *Enserink M.* Are you ready to become a number? // Science. – 2009. – P.1662-1664.
11. *Jörg B., Ferlez F., Uszkoreit H., Jermol M.* Analysing European research competencies in IST: Results from a European SSA Project // Proceedings: CRIS 2008, Maribor, Slovenia.
12. *Fenner M.* ORCID unique identifiers for authors and contributors. Conference Slides The Value of Unique Scholarly Identifiers to Academics, Institutions and Countries. February 2012. - <http://about.orcid.org/content/orcid-unique-identifiers-authors-and-contributors>
13. *CrossRef: A short history.* - 2009.-<http://www.crossref.org/08downloads/CrossRef10Years.pdf>
14. *Kahn R., Wilensky R.* A framework for distributed digital object services // International Journal on Digital Libraries. - 2006. – Vol. 6, No. 2. – P. 115—123.
15. *Bizer C., Heath T., Berners-Lee T.* Linked data – the story so far // International Journal on Semantic Web and Information Systems. – 2009. – Vol. 5, No.3. – P. 1-22.

# Хранилища открытого доступа: сценарий для стран Азии с особым акцентом на библиотековедение и информатику\*

**Шанташри СЕНГУПТА**  
(Shantashree SENGUPTA)

Факультет библиотековедения и информатики,  
Колледж С.Т. Бора, г. Ширур, округ Пуна,  
Индия

*Во всем мире учреждения высшего образования испытывают потребность в более эффективном управлении образованием, научным исследованием и ресурсами. Институциональные хранилища открытого доступа представляют собой наилучший способ, с помощью которого институциональная продукция откроется всему миру, институциональные хранилища помогают в максимизации обозримости и влияния этой научной продукции, дающей в результате возможность осуществления и поощрения междисциплинарных подходов к научному исследованию. Благодаря разнообразным выгодам, получаемым от институциональных хранилищ, многие учреждения развивают свои собственные хранилища. Данное исследование дает картину роста и развития институциональных хранилищ области библиотековедения и информатики в Азии, приводя данные из базы данных OpenDOAR.*

## ВВЕДЕНИЕ

С начала XXI в. наблюдается большой рост числа цифровых библиотек и хранилищ по всему миру. Открытый доступ – это преимущество для учреждений, авторов и т.д., открытый доступ использует Интернет для бесплатного распространения различных типов литературы в мире. В прошлом из-за отсутствия продвинутой информационно-коммуникационных технологий научная литература, опубликованная в мире, не была доступна мировому научному сообществу. Движение открытого доступа получило большое значение и популярность после появления в 2002 г. Будапештской инициативы открытого доступа (BOAI – Budapest Open Access Initiative).

Будапештская инициатива открытого доступа [1] определяет открытый доступ как «бесплатную возможность публичного Интернета, позволяющую любым пользователям читать, выгружать, копировать, распространять и/или печатать, а также осуществлять поиск или связь с полными текстами этих статей, использовать их для индексирования, передавать в качестве данных для программного обеспечения или применять для любой другой законной цели без финансовых, правовых или технических барьеров, отличающихся от тех, которые являются обязательными для получения доступа к самому Интернету.

## ОБЗОР ЛИТЕРАТУРЫ

Бина и Аршана [2] описывают современный опыт управления и совместного использования интеллектуального богатства Университета г. Кочин по науке и технологии (CUSAT – Cochin University of Science and Technology), используя платформы открытых источников. Ученые обнаружили, что разнообразные средства открытого доступа, такие как DSpace, Bibapp, MENDELEY, Zotero используются для управления интеллектуальной информатикой.

Тагер и Оза [3] считают институциональные хранилища необходимым средством для процесса управления знанием. Институциональные хранилища восстанавливают философию д-ра Ш.П. Ранганатана относительно порождения знания, оценки и использования особенно возникающих микромыслей, и ее (философии) механизм на современной цифровой арене.

Раджачекар [4] обсуждает два компонента издательства открытого доступа в контексте Индии. Первый концентрируется на релевантности издательства открытого доступа в развивающихся странах, потенциале для издательства открытого доступа в Индии и ряде современных инициатив открытого доступа в Индии.

Уани, Гул и Рах [5] освещают рост и развитие хранилищ открытого доступа в мире. Далее исследование делает акцент на проводящихся в этом направлении работах в странах Азии и приводит подробные профили Азии.

Уанг и Су [6] констатируют, что хотя открытый доступ начался с развитых стран, он является привлека-

\* Перевод Sengupta Sh. Open access repositories: The Asian scenario with special reference to library & information science. – <http://eprints.rclis.org/18129/>

тельным для развивающихся стран и быстро распространяется по всему миру.

Браун и Аббас [7] в своем исследовании освещают растущее число учреждений в США, которые активно создают институциональные хранилища и множество специальных фондов, работающих в реальном времени, особенно в области биологических наук.

Кришнамурти и Кемпараджу [8] сообщают об исследовании институциональных хранилищ, которые используются в университетах и научных институтах Индии. Исследовались 20 хранилищ, охватывающих фонды различного типа. Большинство изученных институциональных хранилищ работают на платформе операционных систем UNIX/Linux. В то время как элитные учреждения, такие как IIS, IIT, и ряд элитных университетов быстро создали институциональные хранилища, очень большое число университетов, особенно университетов штатов, колледжей и учреждений высшего образования еще только должны инициировать какие-то основные действия по созданию институциональных хранилищ. По мнению ученых, Индия, кажется, все еще слишком отстает от входа в среду современных технологий институциональных хранилищ и их эффективного использования.

Лихиткар и Агаше [9] представили индийский сценарий развития институциональных хранилищ. Всего было определено 33 хранилища, которые анализировались на основе критериев выборочного исследования, таких как используемое программное обеспечение, объем документов, включаемые контент, языки, описания и страны.

Тонта [10] рассматривает текущее положение дел с институциональными хранилищами в Европе и приводит хронику развития институциональных хранилищ в Турции. Ученый рекомендует расширить осведомленность членов профессорско-преподавательского состава относительно открытого доступа и институциональных хранилищ. Предложенные научные гранты должны расширить открытый доступ для научных статей, финансируемых налогоплательщиками. Контент таких хранилищ должны быть организованы в соответствии с установленными стандартами (например, DAI-PMH и Open URL), чтобы они могли быть полезными для тех, кто осуществляет работу с помощью поискового механизма, таким образом увеличивая всеобщую обзорность исследования, осуществляемого турецкими учеными.

## ЦЕЛИ ИССЛЕДОВАНИЯ

Цели настоящего исследования заключались в том, чтобы раскрыть структуру и создать хранилища открытого доступа в области библиотекведения и информатики в Азии. Исследование проводилось по следующим аспектам:

- Изучить материалы, включенные в хранилища открытого доступа по библиотекведению и информатике на предмет географического представления.

- Определить другие дисциплины, охваченные хранилищами открытого доступа по библиотекведению и информатике.

- Определить разнообразие языков.

- Определить основные типы контента.

- Определить разные виды программного обеспечения, используемого для создания хранилищ открытого доступа.

- Определить различные типы хранилищ открытого доступа.

## СФЕРА ИССЛЕДОВАНИЯ

Данные для настоящего исследования были взяты исключительно из «Указателя хранилищ открытого доступа» (DOAR – Directory of Open Access Repositories), широко известного как OpenDOAR ([www.opendoar.org](http://www.opendoar.org)) [11], и касались хранилищ открытого доступа, созданных в странах Азии.

Из 388 хранилищ открытого доступа, идентифицированных в OpenDOAR за период с 25 октября по 7 ноября 2012 г, было отобрано всего 13 (см. Приложение), охватывающих соответствующие контент из области библиотекведения и информатики. Эти 13 хранилищ изучались в соответствии с целями данного исследования.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ

### Хранилища открытого доступа в Азии по странам

Исходя из распределения 13 хранилищ открытого доступа стран Азии в области библиотекведения и информатики по странам, Индия, имеющая 4 хранилища открытого доступа по контентам библиотекведения и информатики, занимает первое место, за ней следуют Тайвань (3 хранилища) и Малайзия (2 хранилища). Подробное представление дано в табл. 1.

Таблица 1

Распределение хранилищ открытого доступа по библиотекведению и информатике в Азии

№ п/п	Страна	Число хранилищ
1	Индия	4
2	Тайвань	3
3	Малайзия	2
4	Индонезия	1
5	Сингапур	1
6	Китай	1
7	Бангладеш	1
	<b>Всего</b>	<b>13</b>

### Распределения хранилищ открытого доступа стран Азии по предметному охвату

Из 13 хранилищ открытого доступа, охватывающих контент по библиотекведению и информатике, только одно хранилище (Цифровая библиотека для библиотечарей) полностью посвящено библиотекведению и информатике. Остальные 12 хранилищ включают также и другие дисциплины. Подробная картина собранных данных представлена в табл. 2.

Таблица 2

Хранилища открытого доступа стран Азии по предметному охвату

№ п/п	Предмет	Число хранилищ
1	Многодисциплинарность	1
2	Технология в целом	5
3	Химия	2
4	Наука в целом	2
5	Здравоохранение и медицина	1
6	Химическая технология	2

№ п/п	Предмет	Число хранилищ
7	Физика и астрономия	1
8	Сельское хозяйство	1
9	Продукты питания и ветеринария	1
10	Экология и окружающая среда	1
11	Компьютеры и информационные технологии	5
12	Бизнес и экономика	5
13	Машинная инженерия	2
14	Искусство и гуманитарные науки	1
15	Биология и биохимия	2
16	Общественные науки	3
17	География и региональные исследования	1
18	Язык и литература	3
19	Математика и статистика	1
20	Образование	3
21	Право и политика	2
22	Управление и планирование	7
23	История и археология	1
24	Электроника и электротехника	2
25	Психология	1
26	Гражданское строительство	2

### Хранилища открытого доступа в Азии по языковому распределению

Английский язык, будучи международным, является наиболее предпочтительным для хранилищ открытого доступа в Азии. Однако использование национальных языков и в ряде случаев даже региональных языков помогает делать хранилища открытого доступа более популярными среди научного сообщества определенного региона или страны, гарантируя максимальное использование фондов хранилища. Имеется много хранилищ, которые используют более одного языка в качестве языкового интерфейса. В табл. 3 подробно представлено языковое распределение, используемое в хранилищах открытого доступа в Азии.

Таблица 3

### Хранилища открытого доступа в Азии по языковому распределению

№ п/п	Язык	Число хранилищ
1	Английский	12
2	Китайский	4
3	Малайский	2
4	Хинди	1
5	Каннада	1

### Хранилища открытого доступа стран Азии по типу основного контента

Из 13 хранилищ открытого доступа, охватывающих область библиотекостроения и информатики, 10 имеют фонды статей, затем идут 8 хранилищ, содержащих книги, и 7 хранилищ, включающих диссертации. Наименее охваченным типом контента являются патенты. Подробная картина полученных данных представлена в табл. 4.

Таблица 4

### Хранилища открытого доступа стран Азии по типу основного контента

№ п/п	Тип контента	Число хранилищ
1	Статьи	10
2	Книги	8
3	Диссертации	7
4	Материалы конференций	6
5	Неопубликованные материалы	5
6	Ссылки	4
7	Мультимедиа	3
8	Специальные материалы	2
9	Учебные материалы	1
10	Патенты	1

### Хранилища открытого доступа стран Азии по типу программного обеспечения

Четыре программных обеспечения на основе открытого источника используются головными организациями или учреждениями для создания хранилищ диссертаций в электронном виде. 9 хранилищ открытого доступа в Азии используют программное обеспечение DSpace. Способность признавать и управлять большим числом форматов файлов и типов изображения делает DSpace широко используемым программным обеспечением на основе открытого источника. Подробная картина представляет следующее.

Таблица 5

### Хранилища открытого доступа стран Азии по типу программного обеспечения

№ п/п	Программное обеспечение	Число хранилищ
1	DSpace	9
2	Eprints	2
3	Greenstone	1
4	Digital Commons	1

### Хранилища открытого доступа в Азии по своему типу

Из 13 хранилищ открытого доступа 11 хранилищ являются институциональными и 2 – по дисциплинам. Ни одно из хранилищ открытого доступа не располагается в каких-либо правительственных организациях или в каком-то объединенном учреждении.

Таблица 6

## Хранилища открытого доступа стран Азии по своему типу

№ п/п	Тип хранилища	Число хранилищ
1	Институциональное	11
2	По дисциплинам	2
3	Объединенное	0
4	Правительственное	0
	<b>Всего</b>	<b>13</b>

## ЗАКЛЮЧЕНИЕ

Приведенная выше статистика относительно роста и разнообразия хранилищ открытого доступа по библиотековедению и информатике в Азии показывает, что во

всем мире ученые пытаются распространять свою научную работу или институциональные материалы с помощью Интернета среди группы людей. Хотя их число еще очень невелико.

Библиотековедение и информатика, являясь одной из самых старых дисциплин, должна иметь большой охват среди институциональных хранилищ, по крайней мере, через университеты и учреждения, где данная дисциплина преподается.

Движение открытого доступа, особенно хранилища открытого доступа, намереваются установить новые стандарты для совместного использования информации и управления. Тенденция создать хранилища открытого доступа по всему миру должна быть поддержана, особенно в таком мире, как Азия, которая имеет полное основание выделиться при наличии сильной основы информационно-коммуникационных технологий и быстро растущего качества академических и научных учреждений с высоким выходом научной продукции.

## ПРИЛОЖЕНИЕ

## Перечень хранилищ открытого доступа стран Азии в области библиотековедения и информатики

№ п/п	Название	URL
1	Faculty of Computer Science and Information Technology, University of Malaya <b>MyManuskrip: Digital Library of Malay Manuscripts (Pustaka Digital Manskrip Melayu)</b>	<a href="http://mymanuskrip.fsktm.um.edu.my/Greenstone/cgi-bin/library.exe">http://mymanuskrip.fsktm.um.edu.my/Greenstone/cgi-bin/library.exe</a>
2	<b>Binus University Repository</b>	<a href="http://eprints.binus.ac.id/">http://eprints.binus.ac.id/</a>
3	<b>Chaoyang University of Technology Institutional Repository (CYUTIR)</b>	<a href="http://ir.lib.cyut.edu.tw:8080/">http://ir.lib.cyut.edu.tw:8080/</a>
4	SDM College Of Engineering and Technology Dharwad <b>dspace @ sdmcet</b>	<a href="http://210.212.198.149:8080/jspui">http://210.212.198.149:8080/jspui</a>
5	ICFAI Business School <b>Dspace@IBSA</b>	<a href="http://202.131.96.59:8080/dspace">http://202.131.96.59:8080/dspace</a>
6	Hsiuping Institute of Technology Institutional Repository ( <b>HITIR</b> )	<a href="http://ir.hust.edu.tw/">http://ir.hust.edu.tw/</a>
7	Indian Institute of Petroleum Institutional Repository	<a href="http://library.iip.res.in:8080/dspace">http://library.iip.res.in:8080/dspace</a>
8	InK ( <b>Institutional Knowledge at Singapore Management University</b> )	<a href="http://ink.library.smu.edu.sg/">http://ink.library.smu.edu.sg/</a>
9	Institutional Repository of Institute of Geographic Sciences and Natural Resources Research, CAS ( <b>IGSNRR OpenIR</b> )	<a href="http://159.226.115.200/">http://159.226.115.200/</a>
10	International Centre for Diarrhoeal Disease Research Digital Repository, Bangladesh ( <b>ICDDR,B</b> )	<a href="http://dspace.icddr.org/dspace/">http://dspace.icddr.org/dspace/</a>
11	Documentation Research and Training Centre (DRTC), Indian Statistical Institute, Bangalore Centre (ISI), <b>Librarians' Digital Library (LDL)</b>	<a href="https://drtc.isibang.ac.in/">https://drtc.isibang.ac.in/</a>
12	<b>National Pingtung Institute of Commerce Institutional Repository</b>	<a href="http://irs.lib.ksu.edu.tw/NPIC/">http://irs.lib.ksu.edu.tw/NPIC/</a>
13	Multimedia University <b>SHDL@MMU Digital Repository</b>	<a href="http://shdl.mmu.edu.my/">http://shdl.mmu.edu.my/</a>

## ЛИТЕРАТУРА

1. *Budapest Open Access Initiative, 2002.* – <http://www.soros.org/openaccess>
2. *Beena C., Archana N.S.* Open access movement for managing intellectual informatics // *Library Herald.* – 2011. – Vol. 49. – P. 221-228.
3. *Thaker U., Oza N.* Institutional repository: An effective tool for knowledge management // *SPELS Journal of Information management.* – 2010. – Vol. 47, No. 5. – P. 507-516. – <http://iproxy.inflibnet.ac.in:2108/ijor.aspx?target=ijor:sjim&volume=47&issue=5&article=003&type=pdf>
4. *Rajashekhar T.* Open access initiatives in India / N.J. Deshpande, S.K. Patil (Eds.), *University and college librarianship in India in the 21<sup>st</sup> century.* – P. 201-207. – Pune: Prof. S.G. Mahajan Felicitation Committee, 2004.
5. *Wani Z.A., Gul S., Rab J.A.* Open access repositories: A global perspective with an emphasis on Asia // *Chinese Librarianship: An International Electronic Journal.* – 2009. – Vol. 27. – <http://www.iclc.us/cliej/c127WGR.htm>
6. *Wang X., Su Chang.* Open access-philosophy, policy and practice: A comparative study // *Chinese Librarianship: An International Electronic Journal.* – 2007. – Vol. 23. – <http://www.iclc.us/cliej/c123WangSu.htm>
7. *Brown C., Abbas J.M.* Institutional digital repositories for science and technology: A view from the laboratory // *Journal of Library Administration.* – 2010. – Vol. 50. – P. 181-215. – doi: 10.1080/01930821003634930
8. *Krishnamurthy M., Kemparaju T.D.* Institutional repositories in Indian universities and research institutes: A study // *Program: Electronic library and information systems.* – 2011. – Vol. 45. – P. 185-190. – doi: 10.1108/00330331111129723
9. *Lihitkar S.R., Lihitkar R.S., Agashe A.T.* A study of major institutional repositories in India. - 2009. – [http://eprints.rclis.org/bitstream/10760/14234/1/ETD\\_2009\\_IRshaliniIndia.pdf](http://eprints.rclis.org/bitstream/10760/14234/1/ETD_2009_IRshaliniIndia.pdf)
10. *Tonta Yasar.* Open access and institutional repositories: The Turkish landscape. – 2008. – <http://eprints.rclis.org/bitstream/10760/11890/1/tonta-liber-2008-06-16.pdf>
11. *OpenDOAR: The directory of open access repositories.* – <http://www.openoar.org/>

Приглашаем российских и зарубежных авторов к сотрудничеству  
в журнале «Международный форум по информации».  
Оригинальные статьи и другие материалы (рецензии, письма)  
можно присылать на русском или английском языке  
по почтовому адресу, указанному в «Памятке для авторов»  
или по электронной почте: [mfi@viniti.ru](mailto:mfi@viniti.ru).

Ответственный за выпуск *Л. В. Кобзева*

Компьютерная верстка *М. А. Филимонова*

ИД № 04689 от 28.04.2001 г.

Подписано в печать 17.06.2013 г.

Бумага «Хегох». Формат 60x841/8. Гарн. литер. Печать цифровая

Усл. печ. л. 4,00 Уч.-изд. л. 4,22 Тираж 55 экз.

Адрес редакции: 125190, Россия, г. Москва, ул. Усиевича, д. 20  
Тел. (499) 155-44-95