

УДК 001.812

ИНФОРМАЦИОННАЯ МОДЕЛЬ РЕЗУЛЬТАТОВ НАУЧНОГО ИССЛЕДОВАНИЯ

А.М. Беглер

Санкт-Петербургский государственный университет, Санкт-Петербург, Россия
alena.begler@gmail.com

Аннотация

Результаты научного исследования существуют в разных формах, появляются на разных этапах процесса исследования и описываются несколькими относящимися к ним понятиями. Во-первых, это этапы исследования, на которых они производятся. Выделяют несколько этапов – от поиска информации до распространения результатов. На каждом из них исследователь осуществляет деятельность, направленную на получение результата. Во-вторых, сами результаты исследования существуют в разных формах – от традиционных статей до наборов данных, получаемых в результате экспериментов или знаний исследователя, лежащих в основе гипотез. Их условно можно разделить на контент, данные и знания. Для производства результатов, в-третьих, используется набор инструментов, которые могут быть классифицированы по результату, который они обеспечивают. В-четвёртых, результатам может быть обеспечен разный уровень доступности, который складывается из степени их открытости научному сообществу и интегрированности в существующие системы организации исследовательской информации. В работе впервые предлагается информационная модель результатов исследования, соединяющая рассмотренные характеристики. Изложенный материал может быть полезен при организации научно-исследовательских работ, в особенности на уровне руководителей и администраторов научных групп.

Ключевые слова: результаты научного исследования, продукты исследования, жизненный цикл исследования, этапы исследования, инструменты исследования, организация исследования.

Цитирование: Беглер, А.М. Информационная модель результатов научного исследования / А.М. Беглер // Онтология проектирования. – 2017. – Т. 7, №2(24). – С. 160-171. – DOI: 10.18287/2223-9537-2017-7-2-160-171.

Введение

Основной результат научного исследования – увеличение знания об исследуемой области. Это знание существует в разных формах: оно описано в статьях, лежит в основе гипотез исследования, определяет используемые экспериментальные методы и так далее. Перечисленные формы результатов можно назвать продуктами научного исследования (*scholarly products*). Они представлены, по большей части, в цифровом виде, и их количество быстро растёт [1]. При этом доступность полученных ранее результатов со временем снижается [2].

Проблема снижения доступности результатов исследований при возрастании их общего количества отчасти решается появлением специализированного инструментального обеспечения. Например, сейчас функционирует более полутора тысяч хранилищ исследовательских данных и несколько десятков систем организации информации для лабораторий и научно-исследовательских организаций. Существующие решения хотя и обеспечивают сохранение продуктов научного исследования, но обладают двумя недостатками. Во-первых, они не универсальны и решают определённые частные проблемы. Во-вторых, они не позволяют определить, организация какой части результатов страдает. Так, многие инструменты и системы позволяют сохранить данные, получаемые в результате экспериментов, но не всегда

позволяют обозначить их взаимосвязи с другими продуктами исследования (например, статьями) и сохранить связанные с ними знания исследователей (например, о процессе сбора этих данных).

Решение обеих проблем могла бы обеспечить базовая, не привязанная к определённой предметной области (ПрО), модель продуктов научного исследования. Такая модель эксплицитно обозначит весь спектр имеющихся продуктов исследования, их взаимосвязи и характеристики. Подобный подход предлагается в работах [3, 4] для конструирования порталов научных знаний. Они используют базовые онтологии научной деятельности и научного знания, которые лежат в основе портала и могут быть расширены за счёт онтологий ПрО.

В работе предлагается базовая информационная модель результатов научного исследования, описывающая, что именно получается в результате исследования (продукт исследования), каким образом (на каком этапе) и с помощью чего (инструмент).

Предлагаемая модель позволит представить структуру продуктов научного исследования во всей полноте и выстроить единую и расширяемую систему, которая может быть использована несколькими способами.

- 1) Для организации информационных потоков исследования, в первую очередь – для руководителей и администраторов научных групп. При таком использовании она может выступать в качестве «каркаса», на котором будет построена система организации конкретной научной группы.
- 2) При подготовке инструментов для обеспечения научной информации – для выявления потребностей в инструментальном обеспечении.
- 3) При конструировании баз и порталов знаний ПрО. В этом случае она должна быть дополнена онтологиями продуктов исследования и инструментария ПрО.

Применение предлагаемой модели описано в заключении к статье. В четырёх разделах основной части статьи представлена каждая из входящих в модель характеристик продукта научного исследования: в первом кратко рассматривается процесс производства продукта исследования; во втором – виды продуктов; в третьем – их доступность, в четвёртом – инструментальное обеспечение. В каждом разделе приведён конкретный способ применения каждой из частей модели на примере подготовки текста научной статьи. Для большинства используемых терминов в скобках даны англоязычные аналоги.

1 Процесс научного исследования

Существует несколько классификаций научно-исследовательского процесса (*research lifecycle, scientific workflow*) по этапам. Kramer и Bosman [5] выделяют шесть этапов научного исследования: поиск информации, анализ, написание текста, публикация, распространение и оценка. Kubiilius [6] предлагает цикл исследовательских работ, характерных для наук о поведении и нейронаук: формулирование гипотезы, разработка программы эксперимента, сбор данных, анализ данных, подготовка рукописи, публикация материала. Nicholas и Rowlands [7] выделяют восемь стадий жизненного цикла исследования и соотносят их с использованием разных социальных медиа: определение возможностей для исследования, научное сотрудничество, получение финансирования, обзор литературы, сбор данных, анализ данных, распространение результатов, менеджмент процесса исследования.

Если рассматривать только этапы, относящиеся к производству непосредственно продукта научного исследования, и не брать в расчёт обеспечение работ, то видно, что описанные классификации совпадают в нескольких ключевых пунктах (рисунок 1). Сначала происходит поиск информации по теме исследования (например, анализ литературы), в результате которого оценивается уже имеющаяся информация и формулируется гипотеза исследования. По-

сле этого начинается сбор и анализ данных (в том числе планирование эксперимента), с помощью которого сформулированная гипотеза проверяется. Когда данные проанализированы, происходит подготовка текста, в котором полученная новая информация сформулирована в явной форме. Далее происходит публикация материалов исследования (не только статей по результатам, но и, например, массивов данных). Оценка эффекта от публикации (комментарии, цитирования) помогает исследователям оценить, насколько актуальна полученная информация среди коллег и общественности. К условно завершающему этапу, который в описанных классификациях назван «Распространение», добавлена «Интеграция»: распространение ориентировано на общественность и коллег – посты в социальных сетях (в том числе профессиональных) и блогах, интервью, выступления с лекциями и так далее, а интеграция касается скорее внедрения новой информации в существующее информационное поле, например, добавление в Википедию или тематические проекты, пополнение баз данных и знаний по изучаемому феномену.



Рисунок 1 – Этапы научного исследования

Отдельное исследование проходит через описанные этапы не линейно, а скорее итеративно. Например, неожиданный результат анализа данных может повлечь за собой сбор новых данных или дополнительный анализ литературы, а подготовка текста – идти параллельно со сбором и анализом данных.

Пример. Текст научной статьи готовится на этапе «Подготовка текста», но с ним связана часть продуктов, полученных на предшествующих этапах. Например, найденные на этапе поиска информации статьи могут быть процитированы в тексте. Зная это, можно подобрать инструмент (менеджер библиографии), который позволит не только сохранять тексты статей, но и в дальнейшем автоматически вставлять цитирования и формировать библиографию. Непосредственно на этапе подготовки текста стоит принимать во внимание последующие этапы. Например, то, в каком издании будет опубликован текст (этап «Публикация») повлияет как на его структуру, так и на особенности содержания. Такие связи существуют между всеми этапами, и заблаговременное их построение экономит время (текст претерпит меньше изменений в процессе публикации) и препятствует потерям информации (сохранённые тексты статей не будут потеряны).

Каждому из этапов присуща деятельность, направленная на производство продукта исследования (рисунок 2).

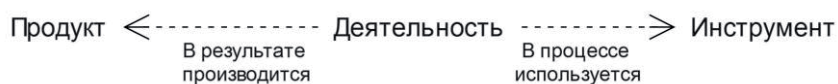


Рисунок 2 – Взаимосвязь деятельности, инструментов и продуктов исследования

Сделан ряд попыток обобщения отдельных направлений исследовательской деятельности. Unsworth [8] предлагает несколько базовых элементов научной деятельности (*scholarly primitives*): изучение, аннотирование, сравнение, проставление ссылок, измерение, пояснение и представление. Palmer с соавторами [9] расширяют количество базовых элементов и на их основе выделяют несколько форм деятельности, относящихся к работе с информацией (*scholarly information activities*) – они, как и базовые элементы, междисциплинарны и могут происходить на любом из описанных этапов. Эти виды деятельности распадаются на пять ключевых групп (рисунок 3): поиск, сбор, чтение, написание и взаимодействие.

Пример. При подготовке текста происходит поиск литературы (поиск), подготовка ссылок на источники (сбор), чтение найденных текстов (чтение), собственно подготовка текста (написание) и обсуждение с

соавторами (взаимодействие). Инструментальное обеспечение исследования должно предусматривать все описанные виды деятельности и получаемые в их результате продукты.

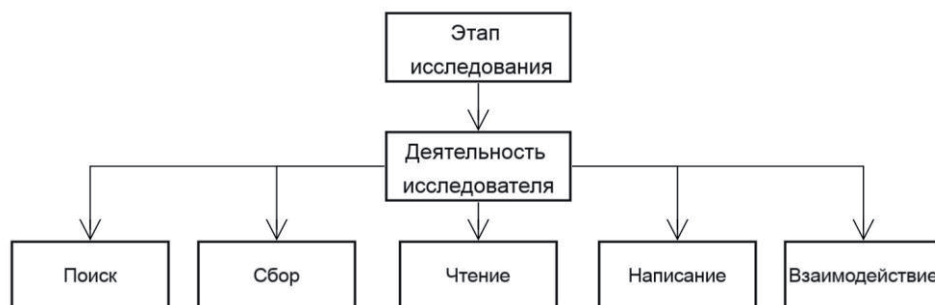


Рисунок 3 – Виды деятельности исследователя

2 Формы продуктов исследования

Если раньше продуктами исследования считались в основном научные статьи, то сейчас к их числу относятся также массивы данных, экспериментальные программы, протоколы исследований и т.п. Косвенно об изменениях продукта исследования можно судить по предлагаемым «альтернативным метрикам» (*altmetrics*) оценки научного влияния, в числе которых указываются цитирования в Википедии, добавления в менеджер библиографии, загрузки pdf-файлов, лайки и репосты в социальных сетях, упоминания в популярных научных блогах, комментарии на страницах публикаций и др. [10, 11].

О расширении спектра продуктов научного исследования свидетельствует и ряд изменений в политике научных фондов, издателей и производителей инструментов:

- начиная с 2013 года Национальный научный фонд (*National Scientific Foundation, NSF*) запрашивает информацию от руководителей проектов не только о статьях, но и других продуктах исследований [12];
- Elsevier, один из крупнейших научных издателей, отмечает важность открытия доступа и организации научных данных [13];
- популярная социальная сеть для исследователей ResearchGate присваивает пользователям рейтинг, основанный, в числе прочего, на количестве положительно оценённых сообществом ответов на запросы других исследователей.

Продукты, описанные в примерах, отличаются ещё и по виду – в первых двух случаях речь идёт о данных и контенте, в третьем – о знаниях. Продукт исследования не гомогенен и отличается как на разных этапах исследования, так и внутри одного этапа. Контент понимается как слабоструктурированная информация, данные – отдельные упорядоченные факты, знания – результат интерпретации данных или метаданные (данные о данных). Результатом поиска информации на уровне контента может стать подборка релевантных статей; на уровне данных – сведённые воедино факты по интересующей тематике, выделенные из собранного контента; на уровне знаний – сформулированная гипотеза исследования (выражающая как знания исследователя об области, так и потребность в их увеличении) [14].

После формулировки гипотезы появляется возможность перейти к этапу сбора данных – продумать план эксперимента, его материальное выражение – протокол исследования, результатом проведения которого будут необходимые «сырые» данные, а также экспериментальные программы и другие составные части эксперимента.

Полученные данные обрабатываются с помощью программ анализа, и в результате получают численные выражения следствий выдвинутых гипотез – размеры эффектов и корре-

ляций, которые подлежат интерпретации. Полученные результаты оформляются в виде визуализаций (графиков, таблиц, схем) и вместе с описанием формируют подготовленный текст, который сопровождается техническими файлами, например, шаблонами оформления или изображениями.

На этапе публикации текст и другие материалы, например данные и код, отправляются в тематические журналы и репозитории и проходят процесс рецензирования, в процессе которого происходит коммуникация с редакторами и рецензентами. После публикации необходимо провести оценку: отследить отдельные упоминания – например, добавления в списки для чтения других исследователей, классические цитирования и т.п., - которые могут сформировать список цитирований и увеличить знания исследователя о восприятии проведённого исследования коллегами.

На этапе распространения и интеграции исследователь может выложить в тематические репозитории отдельные части своего исследования, или дополнительно описать их, или добавить полученные знания, например, в Википедию, в другие тематические проекты.

Обобщённо данный подход представлен на рисунке 4. Обратим внимание, что на большинстве этапов процесс получения итогового продукта идёт от контента к знаниям (получение знания – цель этих этапов); на этапах же сбора данных и оценки влияния – наоборот, отправной точкой становятся знания, а цель этих этапов – накопление данных.

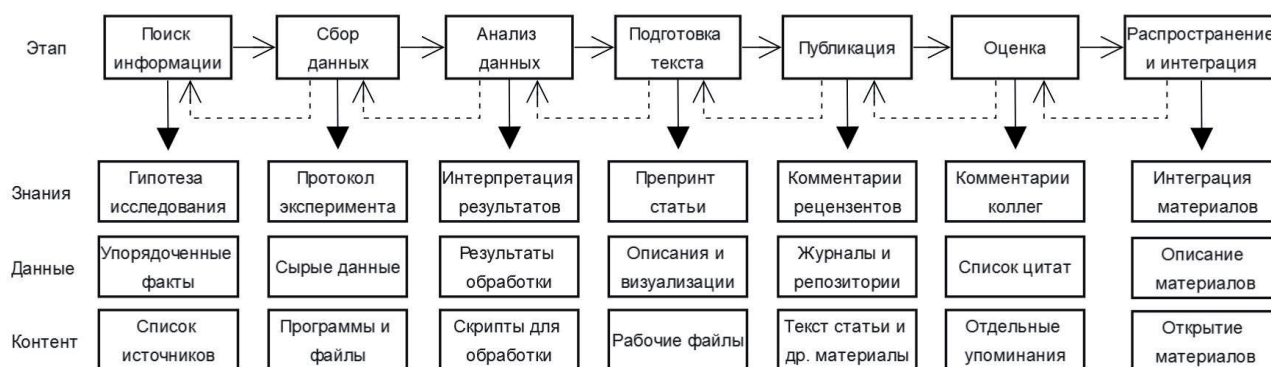


Рисунок 4 – Анализ продукта научного исследования с позиции его разделения на контент, данные и знания

Разделение на данные, знания и контент при этом достаточно условно и необходимо, в первую очередь, для охвата всего спектра продуктов исследования. Если рассматривать основной «физический» продукт этапа как данные, то контент – это те куски, из которых он получился, а знания – это метаданные о процессе и результате его получения. Ценность такого разделения - в экспликации всех продуктов исследования и включении их в явном виде в организационную часть исследования. Не все продукты обязательно должны сохраняться в процессе проведения исследования, но данный подход позволяет принять более обоснованные решения о такой необходимости.

Наиболее часто, когда говорят о результате исследования, речь идёт о данных и, в меньшей степени, о контенте. Так, например, специализированные хранилища данных созданы в первую очередь для контента и данных (материалов для подготовки экспериментов, массивов сырых данных и т.д.), однако добавленные в некоторые инструменты вики-системы и категории делают их пригодными и для работы со знаниями.

Пример. Кроме собственно текста статьи, являющегося основным продуктом на этапе «Подготовка текста», появляются также список источников и «выжимки» из них, схемы и таблицы, результаты обработки данных, информация о вкладе соавторов, идеи о дальнейшем направлении исследований. Первые три можно выявить, акцентируя внимание на то, из чего состоит текст (это контент, связанный с этим этапом), последние два отвечают на вопрос «что мы знаем об этом тексте?» – это знания данного этапа. Эти

продукты должны учитываться в процессе исследования – «выжимки» из источников сохраняться и фиксироваться, таблицы и схемы – существовать в оригинальных форматах (для возможности изменения и последующего использования) и т.д.

3 Доступность продукта исследования

Важная характеристика продукта исследования – уровень его доступности для других исследователей. Науку можно назвать «экономикой дара» (*gift economy*), в которой ценность исследования определяется во многом тем, насколько свободно доступны его результаты другим исследователям [15]. Свидетельством важности этого аспекта продукта исследования можно считать увеличение продуктов исследований, находящихся в открытом доступе. Например, количество таких статей возросло с менее чем двадцати тысяч в 2000-м году до почти двухсот тысяч к 2009-му [16], достигнув, таким образом, примерно двадцати процентов от общего количества статей [17] и продолжало расти, составляя к 2014 году уже около 24 процентов [18].

Доступность определяется двумя характеристиками: открытостью и интегрированностью. Открытым продуктом может воспользоваться каждый (обычно он распространяется под открытыми лицензиями). Приватный продукт – тот, который доступен только исследовательской группе и, возможно, организации (либо вообще не выложен в сеть и хранится, например, в частном хранилище, или выложен под частной лицензией). Открытость способствует распространению продукта исследования, но для полноценного распространения продукта недостаточно открыть к нему доступ – он должен находиться там, где потенциальный пользователь сможет его найти, то есть, быть не только открытым, но и интегрированным в информационное поле соответствующей области исследований. Например, массив данных исследования, доступный для скачивания только с сайта научной группы, будет хотя и открытым, но не интегрированным. Тогда как тот же массив, хранящийся под открытой лицензией в тематическом хранилище (например, figshare) – уже и открытый, и интегрированный. Основным критерий отнесения продукта к интегрированным – возможность доступа к нему посредством обращения к общепринятым поисковикам и хранилищам.

И открытость продукта, и его интеграция определяются возможностями инструментов работы с продуктом исследования, о которых речь пойдёт в следующей части. Поэтому при подборе инструментов нужно учитывать план по распространению продукта, фиксирующий точку зрения на аспекты открытости и интегрированности продуктов исследования.

Пример. Открытость текста статьи отчасти обеспечивается публикацией. Интегрированности способствует индексирование в *Web of Science* или *Scopus*. Интегрированность других продуктов, связанных с текстом, например, отдельных результатов обработки данных, зависит от исследовательской области – например, в области функциональной магнитно-резонансной томографии существует несколько проектов, предназначенных для интеграции полученных результатов сканирования головного мозга (*BrainMap*, *Neurosynth*, *NeuroVault*).

4 Инструменты для работы с продуктом исследования

Интернет-ресурсы для исследователей появились в начале 70-х годов [19], а уже через 30 лет сеть Интернет для научной коммуникации используют практически все исследователи [20]. В цифровое пространство переносятся практически все этапы исследовательского процесса: от общения с коллегами (через почтовые агенты и сервисы видеосвязи) до получения данных (например, проведение экспериментов или привлечение участников с помощью онлайн-сервисов) и доступа к опубликованным материалам [21]. К настоящему времени существует несколько сотен различных инструментов для обеспечения научного исследования, большинство из которых созданы специально для учёных, а не адаптированы ими [22].

Исследования инструментального обеспечения в основном имеют характер анализа средств научной коммуникации (*scholarly communications*). Термин «коммуникация» при этом понимается максимально широко – как любая передача информации, поэтому охватывает практически все возможные инструменты, используемые в исследованиях.

Можно выделить несколько десятков видов инструментов, которые используются в процессе работы над продуктом научного исследования (в скобках даны примеры):

- базы и поисковики статей (РИНЦ, Google Scholar, Scopus);
- сервисы рекомендаций (F1000Prime);
- инструменты доступа к литературе (Open Access Button);
- сервисы аннотирования и чтения (Hypothesis, ReadCube);
- инструменты для доступа к протоколу (плану) исследования (myExperiment, protocols.io);
- графические и видео редакторы (Photoshop, Inkscape);
- программы для демонстрации экспериментов (PsychoPy, Presentation);
- онлайн-опросы (Google Forms, SurveyMonkey);
- оффлайн-опросы экспериментальные методики;
- программы для обработки данных (R, SPSS);
- инструменты визуализации данных (Plotly);
- редакторы исходного кода и интегрированные среды разработки (Notepad++, RStudio, PyCharm);
- инструменты совместного авторства (Google Docs, Authorea);
- менеджеры библиографий (Mendeley, Papers);
- предпубликационная оценка (eGAP, OSF Preregistration);
- инструменты для подбора журнала (SCImago, Elsevier Journal Finder, Jane);
- инструменты публикации (doi, PeerJ, лицензии);
- постпубликационная оценка (PubPeer);
- инструменты отслеживания цитирований (Altmetrics, JCR, Scopus);
- исследовательские профили (ResearchGate, ORCID);
- блоги и микроблоги (Twitter);
- тематические каналы коммуникации (сообщества, списки рассылок);
- публичные хранилища (figshare, Open Science Framework, F1000Posters, arXiv.org);
- тематические базы знаний, например, описания методик и экспериментальных процедур;
- онтологии (An ontology of scientific experiments - EXPO, Mental Functioning Ontology - MFO, Cognitive Paradigm Ontology - CogPO).

Эти инструменты используются на разных этапах исследования, предназначены для определённой формы продукта исследования и могут обеспечивать разные уровни его доступности.

Наиболее часто инструменты классифицируются по *этапам исследования*, на которых они используются (рисунок 5), при этом этапы распространения и интеграции разведены. Такая классификация достаточно условна. С одной стороны, на одном этапе научного исследования могут использоваться инструменты из нескольких групп (например, при написании текста это инструменты совместного авторства, менеджеры библиографии и инструменты визуализации данных). С другой, возможна и обратная ситуация – один инструмент может обслуживать несколько этапов (так, некоторые менеджеры библиографии имеют встроенные модули для хранения статей и работы с текстом).

По *форме* продукта большинство инструментов предназначены для работы в первую очередь с данными и контентом (их поиска, хранения, организации и нахождения). Однако часть инструментов ориентирована на работу со знаниями, среди них:

- сервисы аннотирования статей;

- сервисы послепубликационной оценки статей;
- исследовательские профили;
- регистрации протоколов исследования;
- онтологии и вики-подобные системы.



Рисунок 5 – Вариант классификации инструментов работы с продуктом исследования по этапам

Большинство инструментов обеспечивают *доступность* продукта исследования: либо предусматривают размещение открытой информации, либо обеспечивают эту возможность опционно. Что касается непосредственно распространения информации, инструменты можно разделить на локализованные (предназначенные для работы отдельной научной группы или организации) и интегрирующие (предназначенные для распространения и интеграции информации). Зачастую один и тот же инструмент может обеспечивать и первый, и второй вариант работы. Инструменты, ориентированные на интеграцию, обычно располагают встроенными формами для метаданных, возможностью добавления тегов и системами категорий.

Пример. В процессе подготовки текста используются инструменты для организации списка источников (менеджеры библиографий), для предпубликационной оценки (для редактур текста), совместного авторства (для подготовки текста), визуализации (для графиков и таблиц). При этом часть инструментов используется на других этапах: менеджеры библиографий – на этапе поиска публикаций, инструменты визуализации – на этапе обработки данных. Эти инструменты предназначены для работы с информацией в разных формах – например, менеджер библиографии в первую очередь предназначен для сбора контента (текстов статей), но многие представители этого класса имеют модули для работы со знаниями (идеями исследователей), например, встроенную возможность добавления заметок к статьям. В частности, менеджеры библиографии могут открывать подготовленные коллекции статей, или обсуждать (в форме комментариев) статьи с другими исследователями.

5 Модель продуктов научного исследования

Исследование понимается как процесс, состоящий из последовательных итеративных этапов, каждый из которых связан с определённой деятельностью. Результаты исследования представляются в виде конкретных продуктов, которые могут быть соотнесены с этапами, на которых они произведены, и существуют в разных формах. Каждый из продуктов характеризуется доступностью, которая обеспечивается рядом инструментов. В общем виде модель научного исследования представлена на рисунке 6, которая может быть использована следующим образом.

- Для исследовательских коллективов модель может служить основанием для построения своей системы организации получаемых результатов. Зная какие, в каких формах и на ка-

ком этапе будут получены результаты, исследователи могут подобрать инструментарий, предлагающий как необходимое функциональное обеспечение, так и приемлемый уровень доступности.

- Модель может помочь разработчикам определить место разрабатываемого инструмента в общей структуре инструментального обеспечения и включить в него модули для работы с соответствующими продуктами.
- В зависимости от направленности создаваемых базы или портала знаний модель может быть дополнена онтологиями продуктов исследования и инструментарием специальных ПРО.

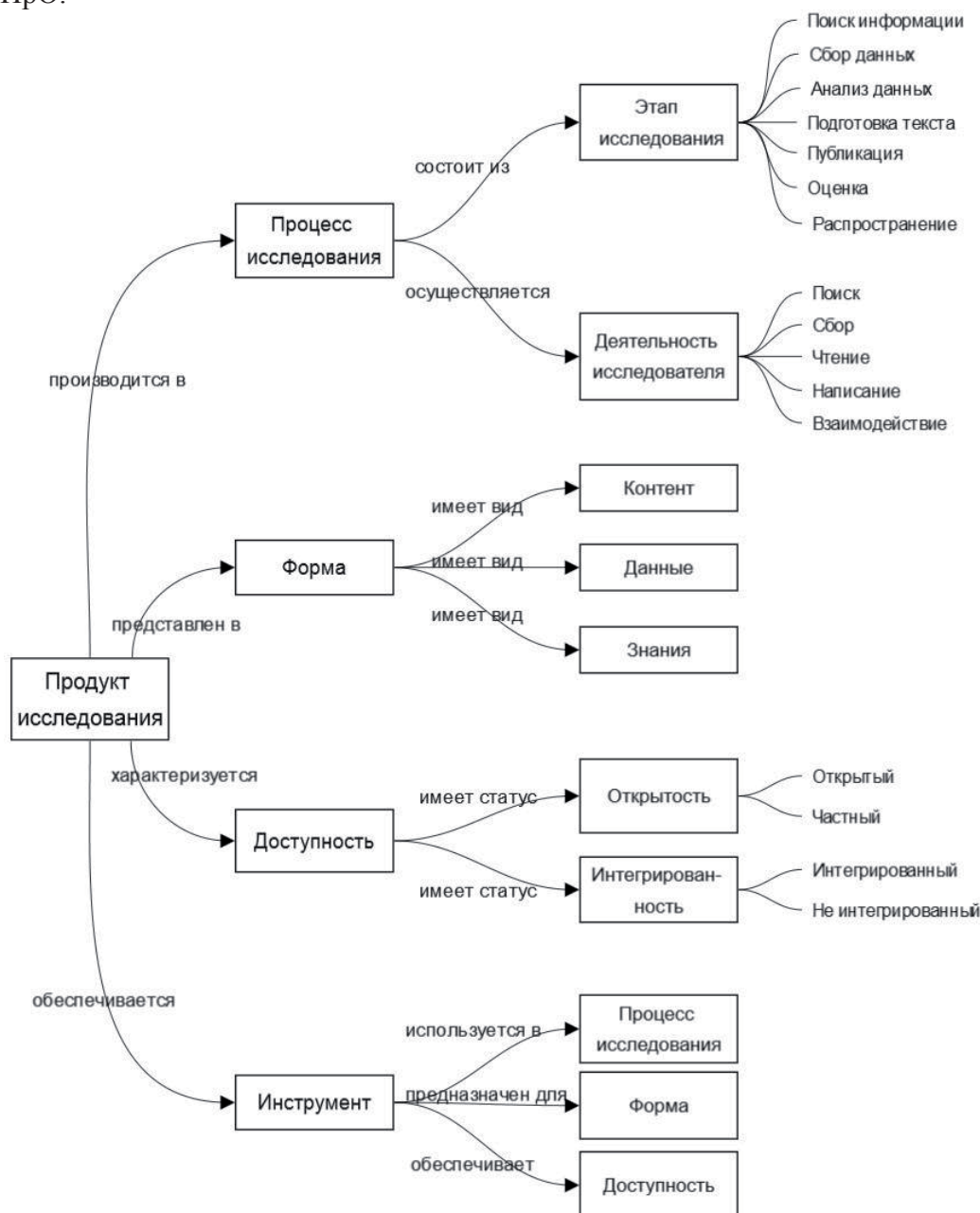


Рисунок 6 – Информационная модель продуктов научного исследования

Заключение

Достоинства предложенной модели состоят в том, что она объединяет ряд понятий, характеризующих продукт научного исследования, которые ранее не рассматривались в единой системе и представляет их в форме, которая может быть дополнена и расширена как предметными характеристиками, так и за счёт уточнения общих характеристик продукта исследования.

К недостаткам модели можно отнести слабую детализацию форм продуктов исследования, которые могут быть преодолены расширением модели и её конкретизацией. В статье не были также рассмотрены организационные аспекты исследовательских работ (финансирование, подготовка плана работ и т.д.). Они связаны с продуктами научного исследования и обладают собственным инструментарием. В данной области имеется ряд разработок (например, онтология проекта VIVO [23]), с которыми предложенная информационная модель продуктов научного исследования может быть объединена.

Благодарности

Автор выражает благодарность Татьяне Альбертовне Гавриловой за ценные комментарии и идеи по усовершенствованию статьи.

Список источников

- [1] **Bornmann, L.** Growth Rates of Modern Science: A Bibliometric Analysis Based on the Number of Publications and Cited References / L. Bornmann, R. Mutz // *J. Assoc. Inf. Sci. Technol. Technol.* 2015. Vol. 66, № 11. P. 2215–2222.
- [2] **Vines, T.H.** The availability of research data declines rapidly with article age / T.H. Vines, A.Y.K. Albert, R.L. Andrew, F. Débarre, D.G. Bock, et. al. // *Curr. Biol.* 2014. Vol. 24, № 1. P. 94–97
- [3] **Загорулько, Ю.А.** Подход к построению порталов научных знаний / Ю.А. Загорулько, О.И. Боровикова // *Автометрия*. 2008. Т. 44, № 1. С. 100–110.
- [4] **Загорулько, Г.Б.** Разработка онтологии для Интернет-ресурса поддержки принятия решений в слабоформализованных областях / Г.Б. Загорулько // *Онтология проектирования*. – 2016. – Т. 6, №4(22). – С. 485-500. – DOI: 10.18287/2223-9537-2016-6-4-485-500.
- [5] **Kramer, B.** Innovations in scholarly communication – global survey on research tool usage [version 1; referees: 2 approved] / B. Kramer, J. Bosman // *F1000Research*. 2016. Vol. 5:692. P. 11.
- [6] **Kubilius, J.** A framework for streamlining research workflow in neuroscience and psychology. / J. Kubilius // *Front. Neuroinform.* 2014. Vol. 7, № January. P. 52.
- [7] **Nicholas, D.** Social media use in the research workflow / D. Nicholas, I. Rowlands // *Inf. Serv. Use.* 2011. Vol. 31, № November. P. 61–83.
- [8] **Unsworth, J.** Scholarly Primitives: what methods do humanities researchers have in common, and how might our tools reflect this? / J. Unsworth // *Symposium on Humanities Computing: Formal Methods, Experimental Practice*. King's College, London. 2000. Vol. 13. P. 0–5.
- [9] **Palmer, C.L.** Scholarly Information Practices in the Online Environment: Themes from the literature and implications for library service development / C.L. Palmer, L.C. Teffeu, C.M. Pirmann // *Scholarly Information Practices in the Online Environment*. 2009. 57 p.
- [10] **Priem, J.** Scientometrics 2.0: New metrics of scholarly impact on the social Web / J. Priem, H.A. Piwowar, B.M. Hemminger // *First Monday*. 2010. Vol. 15, № 7. P. 1–14.
- [11] **Priem, J.** Altmetrics in the wild: Using social media to explore scholarly impact / J. Priem, B.M. Hemminger // *arXiv Prepr. arXiv1203.4745*. 2012. P. 1–23.
- [12] **Piwowar, H.A.** Value all research products / H.A. Piwowar // *Nature*. 2013. Vol. 493, № 7431. P. 159.
- [13] **de Waard, A.** Research data management at Elsevier: Supporting networks of data and workflows / A. de Waard // *Inf. Serv. Use.* 2016. Vol. 36, № 1–2. P. 49–55.
- [14] **Гаврилова, Т.А.** Инженерия знаний. Модели и методы. / Т.А. Гаврилова, Д.В. Кудрявцев, Д.И. Муромцев // Учебник. СПб: Лань, 2016. 324 с.

- [15] **Bollen, J.** A principal component analysis of 39 scientific impact measures / J. Bollen, H. Van de Sompel, A. Hagberg, R. Chute // PLoS One. 2009. Vol. 4, № 6. P. 1–11.
- [16] **Laakso, M.** The development of open access journal publishing from 1993 to 2009 / M. Laakso, P. Welling, H. Bukvova, L. Nyman, B.-C. Björk, T. Hedlund // PLoS One. Public Library of Science, 2011. Vol. 6, № 6. P. e20961.
- [17] **Björk, B.-C.** Open access to the scientific journal literature: situation 2009 / B.-C. Björk, P. Welling, M. Laakso, P. Majlender, T. Hedlund, G. Guðnason // PLoS One. Public Library of Science, 2010. Vol. 5, № 6. P. e11273.
- [18] **Khabsa, M.** The number of scholarly documents on the public web /M. Khabsa, C.L. Giles// PLoS One. 2014. Vol. 9, № 5.
- [19] **Kling, R.** Not just a matter of time: field differences and the shaping of electronic media in supporting scientific communication / R. Kling, G McKim // J. Am. Soc. Inf. Sci. Technol. 2000. Vol. 51, № 14. P. 1306–1320.
- [20] **Walsh, J.P.** Connecting minds: computer-mediated communication and scientific work / J.P. Walsh, S. Kucker, N.G. Maloney, S. Gabbay // J. Am. Soc. Inf. Sci. Technol. 2000. Vol. 51, № 14. P. 1295–1305.
- [21] **Pearce, N.** Digital scholarship considered: how new technologies could transform academic work / N. Pearce, M. Weller, E. Scanlon, S. Kinsley // Educ. 2010. Vol. 16, № 1. P. 33–44.
- [22] **Pearce, N.** A Study of Technology Adoption By Researchers / N. Pearce // Information, Commun. Soc. 2010. Vol. 13, № 8. P. 1191–1206.
- [23] **Krafft, D.B.** VIVO: Enabling National Networking of Scientists / D.B. Kraft, N.A. Cappadona, B. Caruso, J. Corson-Rikert, M. Devare, B.J. Lowe, // Proceedings of the WebSci10: Extending the Frontiers of Society On-Line. 2010.
-

INFORMATION MODEL OF SCHOLARLY PRODUCTS

A.M. Begler

Saint Petersburg State University, Saint Petersburg, Russia
alena.begler@gmail.com

Abstract

Results of the scientific research – scholarly products – are not homogenous. They appear in different forms during different research activities and might be described with several related concepts. First one is the research workflow, which consists of several phases: from searching of the relevant information through data collecting to dissemination of the results. Every phase characterized by the activities, which described by scholarly primitives and the research products are the results of this activities. Second, the research products itself are represented by information of different types: content, data and knowledge. For example, during searching of the relevant information scientist produce several types of information: collection of papers/datasets/etc, several facts, derived from them, and research hypothesis, which express researcher’s personal ideas. For producing them modern scientists use research tools (the third concept), which could be distinguished by the information type too, but not always one-by-one: for example, most of the reference management programs are not only storing the papers, but also maintaining the work on the text reading. One more related concept – information availability, reflects openness and discoverability of the scholarly products and might be to the different degree supported by the instruments. We put these four concepts into the information model of the scholarly products, which will be helpful for managing research results.

Key words: *scholarly information, scholarly products, research lifecycle, research workflow, research tools, research management.*

Citation: *Begler AM.* Information model of the scholarly products. *Ontology of designing.* 2017; 7(2): 160-171. DOI: 10.18287/2223-9537-2017-7-2-160-171.

References

- [1] **Bornmann L., Mutz R.** Growth Rates of Modern Science: A Bibliometric Analysis Based on the Number of Publications and Cited References. *J. Assoc. Inf. Sci. Technol. Technol.* 2015. Vol. 66, № 11. P. 2215–2222.

- [2] *Vines TH*. The availability of research data declines rapidly with article age. *Curr. Biol.* 2014. Vol. 24, 1. P. 94–97.
- [3] *Zagorulko YuA., Borovikova OI*. An approach to constructing knowledge portals. *Optoelectronics, Instrumentation and Data Processing*. Vol. 44. № 1. 2008. P. 75–82.
- [4] *Zagorulko GB*. Development of ontology for intelligent scientific internet resource decision-making support in weakly formalized domains. *Ontology of designing*. 2016; 6(4): 485–500. DOI: 10.18287/2223-9537-2016-6-4-485-500.
- [5] *Kramer B, Bosman J*. Innovations in scholarly communication – global survey on research tool usage [version 1; referees: 2 approved]. *F1000Research*. 2016. Vol. 5:692. P. 11.
- [6] *Kubilius J*. A framework for streamlining research workflow in neuroscience and psychology. *Front. Neuroinform.* 2014. Vol. 7, № January. P. 52.
- [7] *Nicholas D, Rowlands I*. Social media use in the research workflow. *Inf. Serv. Use*. 2011. Vol. 31, № November. P. 61–83.
- [8] *Unsworth J*. Scholarly Primitives: what methods do humanities researchers have in common, and how might our tools reflect this? *Symposium on Humanities Computing: Formal Methods, Experimental Practice*. King's College, London. 2000. Vol. 13. P. 0–5.
- [9] *Palmer CL, Teffeau LC, Pirmann CM*. Scholarly Information Practices in the Online Environment: Themes from the literature and implications for library service development. *Scholarly Information Practices in the Online Environment*. 2009. 57 p. - <http://www.oclc.org/content/dam/research/publications/library/2009/2009-02.pdf>.
- [10] *Priem J, Piwowar HA, Hemminger BM*. Scientometrics 2.0: New metrics of scholarly impact on the social Web. *First Monday*. 2010. Vol. 15, № 7. P. 1–14.
- [11] *Priem J, Hemminger BM*. Altmetrics in the wild: Using social media to explore scholarly impact. *arXiv Prepr. arXiv1203.4745*. 2012. P. 1–23.
- [12] *Piwowar HA*. Value all research products. *Nature*. 2013. Vol. 493, № 7431. P. 159.
- [13] *de Waard A*. Research data management at Elsevier: Supporting networks of data and workflows. *Inf. Serv. Use*. 2016. Vol. 36, № 1–2. P. 49–55.
- [14] *Gavrilova TA, Kudryavtsev DV, Muromtsev DI*. Knowledge engineering. Models and methods [In Russian]. SPb: Lan', 2016. 324 p.
- [15] *Bollen J, Van de Sompel H, Hagberg A, Chute R*. A principal component analysis of 39 scientific impact measures. *PLoS One*. 2009. Vol. 4, № 6. P. 1–11.
- [16] *Laakso M, Welling P, Bukvova H, Nyman L, Björk B-C, Hedlund T* (2011) The Development of Open Access Journal Publishing from 1993 to 2009. *PLoS ONE* 6(6): e20961. <https://doi.org/10.1371/journal.pone.0020961>
- [17] *Björk B-C, Welling P, Laakso M, Majlender P, Hedlund T, Guónason G* (2010) Open Access to the Scientific Journal Literature: Situation 2009. *PLoS ONE* 5(6): e11273. <https://doi.org/10.1371/journal.pone.0011273>
- [18] *Khabsa M., Giles CL*. The number of scholarly documents on the public web. *PLoS One*. 2014. Vol. 9, № 5.
- [19] *Kling R, McKim G*. Not just a matter of time: field differences and the shaping of electronic media in supporting scientific communication. *J. Am. Soc. Inf. Sci. Technol.* 2000. Vol. 51, № 14. P. 1306–1320.
- [20] *Walsh JP, Kucker S, Maloney NG, Gabbay S*. Connecting minds: computer-mediated communication and scientific work. *J. Am. Soc. Inf. Sci. Technol.* 2000. Vol. 51, № 14. P. 1295–1305.
- [21] *Pearce N, Weller M, Scanlon E, Kinsley S*. Digital scholarship considered: how new technologies could transform academic work. *Educ.* 2010. Vol. 16, № 1. P. 33–44.
- [22] *Pearce N*. A Study of Technology Adoption By Researchers. *Information, Commun. Soc.* 2010. Vol. 13, № 8. P. 1191–1206.
- [23] *Krafft D.B., Cappadona NA., Caruso B., Corson-Rikert J., Devare M., Lowe B.J.*, VIVO: Enabling National Networking of Scientists. *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line*. 2010.

Сведения об авторах



Беглер Алёна Маратовна, 1988 г. рождения. Окончила Санкт-Петербургский государственный университет в 2011 г. С 2014 года – менеджер научной группы, работающей в области когнитивной психологии (факультет психологии СПбГУ). Область научных интересов: организация исследовательского процесса, в особенности – получаемой информации, проектный менеджмент, инженерия знаний.

Alena M. Begler (b. 1988) graduated from the Saint Petersburg State University in 2011. Since 2014 working as a scientific group manager in cognitive psychology field (SPbU Department of Psychology SPbU). Main interests are: research workflow organizing (especially – management of the information products), project management, knowledge engineering.