

ПРОДВИЖЕНИЕ НАУЧНЫХ ПУБЛИКАЦИЙ / PROMOTION OF SCIENTIFIC PUBLICATIONS

Оригинальная статья / Original paper

<https://doi.org/10.24069/SEP-23-15>

Академическая поисковая оптимизация: улучшение видимости и доступности научных публикаций

Е. В. Тихонова  

МГИМО Университет МИД РФ, г. Москва, Российская Федерация

 etihonova@gmail.com

Резюме. Адаптированный под инструменты академической поисковой оптимизации (academic search engine optimization, ASEO) текст научной публикации позволяет последней получить высокий рейтинг в контексте поисковых запросов в базах данных и других системах. Как результат, статья попадает в большее количество поисковых запросов, получает большее количество просмотров, скачиваний и цитат. ASEO способствует не только повышению принципиальной обнаруживаемости статьи, но и оптимизации ее метаданных. Тщательно сформулированные заглавия, ключевые слова и аннотация повышают интерес всех категорий потенциальных читателей к контенту статьи, способствуют более глубинному пониманию сути описанного исследования и мотивации авторов к его проведению. Цель представленной статьи – описать инструменты ASEO, позволяющие автору оптимизировать текст рукописи для ее постпубликационного продвижения. В статье представлена базовая информация о механизмах ранжирования, прокомментированы преимущества и ограничения оптимизации. Статья будет полезна не только авторам научных публикаций, но и редакторам научных журналов, заинтересованным в повышении видимости своих изданий.

Ключевые слова: академическая поисковая оптимизация, научная коммуникация, научные журналы, научные публикации, обнаруживаемость, продвижение, распространение

Благодарности: Автор выражает искреннюю благодарность Ольге Владимировне Кирилловой, президенту Ассоциации научных редакторов и издателей России (АНРИ) и главному редактору журнала «Научный редактор и издатель», за идейное сопровождение подготовки рукописи статьи и комментарии относительно ее содержания.

Для цитирования: Тихонова Е. В. Академическая поисковая оптимизация: улучшение видимости и доступности научных публикаций. *Научный редактор и издатель*. 2023;8(1):18–27. <https://doi.org/10.24069/SEP-23-15>

Academic Search Engine Optimization: Improving visibility and accessibility of scientific publications

E. V. Tikhonova  

MGIMO University, Moscow, Russian Federation

 etihonova@gmail.com

Abstract. A text of a scientific paper, adapted for academic search engine optimization (ASEO) tools, enables it to achieve a high ranking in the context of search queries in databases and other search systems. As a result, the article appears in more search queries, receiving more views, downloads, and citations. ASEO not only promotes the fundamental discoverability of the article but also optimizes its metadata. Carefully formulated titles, keywords, and abstracts increase the interest of all categories of potential readers in the paper content, facilitating a deeper understanding of the essence of the described research and the authors' motivation for conducting it. The purpose of the presented paper is to describe the ASEO tools,

enabling the author to optimize the manuscript text for its post-publication promotion. The paper presents basic information about ranking mechanisms, comments on the advantages and limitations of ASEO optimization. It will be useful not only to authors of scientific publications but also to editors of scientific journals who are interested in increasing their visibility.

Keywords: academic search engine optimization, ASEO, scholarly communication, academic journals, scholarly publications, findability, discoverability, dissemination

Acknowledgements: The author expresses sincere gratitude to Olga V. Kirillova, President of the Association of Science Editors and Publishers of Russia (ASEP/ANRI) and the Editor-in-Chief of the journal “Science Editor and Publisher”, for providing conceptual guidance throughout the preparation of the manuscript.

For citation: Tikhonova E. V. Academic Search Engine Optimization: Improving visibility and accessibility of scientific publications. *Science Editor and Publisher*. 2023;8(1):18–27. (In Russ.) <https://doi.org/10.24069/SEP-23-15>

Введение

Философия ASEO

Перенасыщенность интернет-пространства информацией обусловила необходимость поисковой оптимизации (*search engine optimization*, SEO), представляющей собой стратегию онлайн-маркетинга, которая нацелена на оптимизацию поиска веб-сайтов и документов в поисковых системах. SEO обуславливает необходимость создания или изменения веб-сайта таким образом, чтобы поисковым системам было проще сканировать и индексировать его контент. Прогрессирующий объем научной информации актуализировал необходимость развития инструментов и академической поисковой оптимизации (*academic search engine optimization*, ASEO). ASEO трактуется как создание, публикация и модификация научной литературы способом, который позволяет академическим поисковым системам эффективно ее сканировать и индексировать. Элементы рукописи, максимально поддерживающие поисковую оптимизацию: заглавие, аннотация, ключевые слова. Подобная оптимизация нацелена: (1) на предоставление исследователям/читателям максимально возможной поддержки в поиске релевантных результатов по их поисковым запросам; (2) на помощь авторам в достижении высокого рейтинга их публикаций в поисковых системах и базах данных.

Существуют четыре принципиальных отличия ASEO от SEO [1; 2].

1. В области веб-поиска Google является лидером для большинства стран, что поддерживает стремление разработчиков контента для сайтов оптимизировать его преимущественно под поисковые алгоритмы Google. В сфере научной информации столь же яркого лидерства нет, поэтому авторам научной литературы необходимо оптимизировать метаданные своих статей под

потребности нескольких поисковых систем. Поскольку разные поисковые системы основаны на разных методах сканирования и ранжирования, оптимизация контента усложняется.

2. Веб-мастерам, управляющим сайтами, в большинстве случаев не нужно беспокоиться об индексации сайта поисковой системой: если какая-либо веб-страница связана с уже проиндексированной страницей, она в какой-то момент будет просканирована и проиндексирована поисковыми системами. Иная ситуация в научных кругах, где лишь часть всех опубликованных материалов доступна академическим поисковым системам. Многие научные статьи хранятся в базах данных издателей, являясь частью академической «невидимой сети», поскольку академические поисковые системы обычно не могут получить доступ к этим статьям и индексировать их. Даже Google Scholar, сотрудничающий с издателями по всему миру, не охватывает все существующие статьи. Исследователям, редакторам и издателям журналов, заинтересованным в постпубликационном продвижении рукописей, необходимо продумывать возможные пути индексации своих статей академическими поисковыми системами.

3. Веб-мастера могут изменять страницы сайтов, добавляя или заменяя в режиме реального времени слова и ссылки, удаляя страницы, предлагая несколько версий с небольшими вариациями и т.д. Таким образом они могут тестировать новые методы представления контента и адаптироваться к изменениям в алгоритмах ранжирования. Авторы научных статей и редакторы журналов такой возможности не имеют: после публикации статьи изменить ее практически невозможно вне механизма ретракции или постпубликационного рецензирования с последующей доработкой рукописи и опубликования ее новой версии с новыми doi.

4. Поисковые интернет-системы обычно индексируют весь текст на веб-сайте. Напротив, большинство академических поисковых систем реферативных баз данных предоставляют для поиска только метаданные статей¹. Это означает, что авторам необходимо особенно вдумчиво подходить к подготовке метаданных в процессе подготовки рукописи статьи.

ASEO для исследователей и издателей

Востребованность каждой конкретной научной публикации измеряется посредством анализа целого ряда факторов, самыми очевидными из которых являются подсчет количества просмотров, загрузок и цитирований публикации. Иными словами, чем проще потенциальному читателю будет обнаружить статью конкретного автора, тем больше у нее шансов на успех. Отсюда и столько внимания, которое уделяется распространению и продвижению результатов исследований, в том числе с опорой на социальные сети. Научные журналы, стремясь к статусу высокорейтинговых, пристально анализируют потенциал каждой публикуемой статьи с позиций интереса читательской аудитории. Как результат, распространение стратегий ASEO становится все более повсеместным.

С прогрессирующим ростом объема научной литературы авторам становится все сложнее привлекать внимание к собственным публикациям, а читатели оказываются перед лицом потенциального кризиса доступности научного контента, поскольку идентификация релевантных научных источников оказывается все более сложной. И именно техники ASEO и обширные метаданные способствуют снижению информационной перегруженности и позволяют обеспечить эффективную маршрутизацию потоков научной коммуникации.

Этика научной коммуникации и ASEO

ASEO в первую очередь призвана способствовать прозрачности научной коммуникации, что акцентирует ее ориентацию на этические конвенции научно-исследовательской деятельности. Стремясь сформулировать заглавие собственной статьи таким образом, чтобы читатель мог обнаружить ее максимально оперативно, автор всегда должен поддерживать баланс между творческой научной свободой, публикационной этикой, качеством и содержанием исследования, а также

потенциальной наглядностью и доступностью метаданных. Оптимизация научной статьи с точки зрения наглядности представления результатов исследования не должна сопровождаться их преувеличением или искажением. Ложные ожидания читателей от контента статьи, базирующиеся на заглавии, аннотации и ключевых словах, приведут к обратному эффекту: статья скорее получит негативную критику, а не признание читательской аудитории. Любая «чрезмерная оптимизация» только усложнит поиск соответствующих исследований, параллельно нанося ущерб репутации не только отдельным авторам и научным изданиям, но и науке в целом. И даже максимально оптимизированный текст статьи, описывающий результаты некорректно проведенного исследования, не обусловит повышения ее рейтинга у читателей.

Цель данной статьи: представить базовую информацию о механизмах оптимизации доступности научных публикаций для потребителей их контента.

ASEO и рейтинг релевантности

Большинство поисковых систем научной литературы реализуют не только поиск элементов, соответствующих введенным критериям поиска, но и ранжирование результатов по релевантности, опирающееся на целый ряд факторов [3].

Успешность каждого поискового запроса зависит от качественного уровня метаданных: заглавий, подзаголовков, особенности написания фамилий и имен авторов, идентификаторов статей и ключевых слов, выходных данных. Поисковая система оценивает частоту встречаемости терминов поискового запроса в метаданных, а также местоположение терминов поискового запроса в тексте публикации. Например, документ, содержащий поисковый термин в заглавии, будет иметь более высокий рейтинг в списке источников, отображенных в результатах поиска, по сравнению с документом, в котором этот термин встречается только в аннотации. В определении рейтинга статьи внутри поискового запроса учитывается и частота употребления термина в метаданных или полном тексте статьи. Чем чаще поисковый термин появляется в метаданных, аннотации и полном тексте², тем больше баллов релевантности будет присвоено документу алгоритмом. Если полный текст статьи доступен в открытом доступе, можно осуществлять

¹ Только в тех случаях, когда в Интернете представлены полные тексты в открытом доступе, поисковые системы могут их индексировать.

² В случае наличия открытого доступа к тексту статьи.

поиск по более широкому кругу слов, что повышает рейтинг релевантности. Баллы суммируются, и документы ранжируются соответствующим образом [1; 4].

В зависимости от технологии поисковой системы могут также учитываться многие другие факторы, такие как год публикации (недавно опубликованные статьи и книги алгоритм трактует как более релевантные), количество цитирований и просмотров по отношению к общему количеству найденных документов, а также импакт-фактор журнала. В концепции ASEO статья должна иметь единую структуру и оформление, включая стандартные разделы внутри конкретного жанра научной коммуникации: введение, материалы и методы, результаты и т.д. Стандартизированная структура помогает академическим поисковым системам в Интернете идентифицировать статью как научную.

Понимание логики выстраивания поискового запроса внутри каждой поисковой системы позволит авторам и редакторам журнала расставить корректные акценты в рукописи, предопределяя тем самым ее потенциальный успех у потребителей контента. Например, поисковые алгоритмы Google Scholar и Microsoft Academics чувствительны в ранжировании к наличию цитирования на статью, весу цитирующего источника и h-индексу автора статьи [5–8]. То есть, чем более значим источник, в котором опубликована цитирующая статья, и чем выше h-индекс автора статьи, тем более высокое место в рейтинге поискового запроса она займет.

Помимо популярности источника, к числу наиболее значимых критериев выстраивания иерархии отобранных по запросу статей относятся релевантность, полезность, уникальность и индекс доверия. Релевантность измеряется отображением значимых ключевых слов, описывающих проблематику исследования, в заглавии и аннотации. Их плотность в указанных подразделах метаданных значительно повышает рейтинг статьи с точки зрения поисковых алгоритмов³. Полезность метаданных отражается в их способности ответить на все вопросы, связанные с исследованием (*Что? Кто? Где? Когда? Почему?*). Уникальность определяется не только отсутствием плагиатных заимствований, но и реальной способностью статьи заполнить пробел в предметном поле: в поисковой иерархии

побеждают статьи, тематика которых получила слабое освещение в публикациях, но сопровождается значительным количеством поисковых запросов. Индекс доверия к статье определяется действиями не только автора статьи, но также редакции и издателя журнала. Стремясь выставить барьер для деятельности журналов-хищников, поисковый алгоритм Google Scholar отслеживает сроки функционирования домена, адрес редакции, структурированность статей в контексте общепринятых предметных и жанровых конвенций. Например, использование векторной графики будет работать в пользу повышения индекса доверия, а растровая графика сыграет противоположный эффект [4].

В международной базе данных Scopus ранжирование в тематическом поисковом запросе значительным образом зависит от релевантности поискового термина (чем более общим является его содержание, тем ниже рейтинг), частотности и места его упоминания в тексте (вес упоминания ключевого слова в заглавии значительно выше, нежели в последней подсекции аннотации, например). Упоминание нескольких ключевых слов в тексте рядом также способствует повышению рейтинга статьи [5].

Оптимизация заглавия

Смысл vs лаконичность

Поисковые термины, встречающиеся в заглавии, имеют наивысшую релевантность не только для поисковых систем, но и для пользователей. Целый ряд статей посвящен исследованию влияния длины заглавия на количество просмотров и цитирований статьи. Отдельные исследователи рекомендуют стремиться к коротким заглавиям, поскольку более длинные «отпугивают авторов» [4; 9; 10]. Существуют и противоположные рекомендации: описательные заглавия в 10–13 слов дают полное представление о содержании статьи и повышают релевантность поискового запроса [11]. В целом, в контексте ASEO речь должна идти о взвешенном подходе в каждом конкретном случае: описательность или декларативность описания содержания статьи, заглавие в виде вопроса или утвердительного словосочетания и пр. диктуются не только контентом конкретной статьи, но и конвенциями предметной области.

Наиболее важные ключевые слова необходимо включать в первую часть заглавия, чтобы читателям было проще идентифицировать содержание публикации. Более того, поисковые

³ Elsevier. Scopus: Access and Use Support Center. What Does “Relevance” Mean in Scopus? Available at: https://service.elsevier.com/app/answers/detail/a_id/14182/supporthub/scopus/ (accessed 08.08.2023).

системы, как правило, выводят в списке поискового запроса первые 65 знаков заглавия. Некоторые поисковые системы, такие как Google, сокращают длинное заглавие (обычно демонстрируются только первые 10 слов) и, следовательно, отсекают последние (в данном случае самые важные) слова. Например, заглавие «О ключевых психосоциальных и управленческих аспектах вынужденного онлайн-обучения» начинается с менее незначимых слов («О ключевых психосоциальных и управленческих аспектах»), тогда как наиболее значимые слова («вынужденное онлайн-обучение») спрятаны в конце довольно длинного текста.

Потеря смысла заглавия вне контекста конкретного научного издания является значимой проблемой релевантности поисковых запросов. Например, в контексте специального выпуска журнала, посвященного вынужденному онлайн-обучению, заглавие статьи «Сильные и слабые стороны процесса онлайн-обучения» будет «работать» (пусть и не в должной мере, но информация будет считываться достаточно полно). Однако в контексте базы данных или каталога библиотеки, вне понимания фокуса специального выпуска, такое заглавие перестанет нести релевантную информацию о статье. О каком именно обучении идет речь? В каком контексте? Почему необходимо выделять сильные и слабые стороны? Даже попав в список статей в рамках поискового запроса, с большой долей вероятности такая статья не привлечет внимание потенциального реципиента.

Насыщенность заглавия ключевыми словами является механизмом ранжирования. В рамках поискового запроса алгоритм ранжирования будет оценивать заглавие каждой конкретной статьи, которое содержит только два из трех поисковых терминов, как более низкий в сравнении с вариантом, содержащим три ключевых слова. При этом даже если статья по содержанию идеально соответствует поисковому запросу, алгоритм реферативной базы данных, анализируя заглавие, не сможет узнать об этом. Поэтому, создавая заглавие, необходимо убедиться, что его толкование однозначно и не вводит в заблуждение, будучи представленным вне контекста номера журнала.

Существуют алгоритмы, призванные помочь авторам статей сконструировать эффективные заглавия (PCC/Population, Concept, Context), PICO/Population, Phenomenon of Interest, Context), SPICED (Setting, Population, Intervention/Phenomenon of Interest, Comparison, Evaluation, Design)

и др.⁴ [12; 13]. Например, алгоритм SPICED, популярный в медицинской среде, рекомендует в качестве обязательного шага упоминание в заглавии не менее двух из перечисленных элементов.

Крайне важным является упоминание дизайна исследования, поскольку одним из факторов поискового ранжирования является ориентация на заранее разработанный план исследования. Так, нарративный обзор в процессе ранжирования получит значительно меньшее количество баллов по сравнению с метаанализом. Например, типичное для гуманитарных российских статей заглавие типа «К вопросу о ... (высшем образовании/онлайн-обучении/развитии сферы здравоохранения)» нарушает все мыслимые механизмы построения заглавия, поскольку не несет вообще никакой информации о сути содержания и даже в самом отдаленном приближении не соответствует никаким алгоритмам его построения. Заглавия типа «Характеристики условий селекционной эволюции семян пшеницы», будучи чуть более конкретными, также, по сути, не дают читателю понимания специфики исследования. О каких именно характеристиках идет речь? В конкретных природных условиях какой страны реализовывалось исследование? Речь идет об обзоре предыдущих исследований в предметной области, или авторы статьи проводили собственное исследование? Характеризуется ли исследование хоть какой-то новизной? И это далеко не весь перечень вопросов, которые порождает непроработанное заглавие статьи.

Заглавие vs подзаголовок

Одна из стратегий оптимизации, которая связана не столько с технологией поисковых систем, сколько с поисковым поведением пользователей, заключается в использовании подзаголовков. Например, исследователи из сферы гуманитарных и социальных наук предпочитают креативные за-

⁴ Алгоритм построен на пошаговой структуре формулирования эффективного исследовательского вопроса (SPICE) для качественного исследования и подчеркивает чувствительность последнего к конкретному контексту и перспективам использования полученных результатов для конкретной целевой группы. Будучи дополненной компонентом D (Research Design), данный алгоритм создает подробную основу для конструирования эффективного заглавия. Исследователи из гуманитарной сферы получают, таким образом, эффективный инструмент, позволяющий отследить присутствие всех необходимых аспектов исследования в заглавии. Пример конструирования заглавия в рамках алгоритма: Setting (Где?), Perspective (Для кого?), Intervention/Phenomenon of Interest (Что?), Comparison (Сравнить с чем? – опционально), Evaluation (Оценить в сравнении с какими результатами?), Design (Каков дизайн исследования?). Сформулировав все эти положения, автору остается только отшлифовать заглавие, выбрав конкретную его оболочку и удалив пустые слова, чтобы в идеале оно не превышало 13 слов.

главия, даже если они не точно отражают содержание статьи. Спасением в такой ситуации выступает подзаголовок, поясняющий конкретный контекст исследования. Учитывая тот факт, что поисковые системы классифицируют подзаголовки как менее релевантные в сравнении с основным заглавием, с точки зрения ASEO, следует уделить больше внимания терминам, включаемым в заглавие. Ключевое слово, представленное в подзаголовке, привлечет более низкий рейтинг ранжирования статьи в списке источников, отобранных в рамках поискового запроса. Отдельные базы данных в поисковом запросе могут не отображать подзаголовки вообще или не учитывать их в связи с ограничением количества отображаемых в поисковой строке знаков. Особенно это касается заглавий, содержащих кавычки. Например, статья под названием «*В университетах нет цифры!*»: Анализ опыта вынужденного онлайн-обучения в период пандемии с использованием цифровых платформ» в поисковом запросе может отображаться только основным заглавием «*В университетах нет цифры!*», и читателю будет сложно связать его с темой отсутствия цифровой грамотности преподавателей и студентов. Следовательно, количество прочтений и потенциальных цитат изначально окажется под угрозой. Отсюда, привлекающие внимание фразы имеет смысл использовать только в подзаголовке, а самые важные ключевые слова должны быть размещены в начале основного заглавия.

Специальные символы

Специальные символы в заглавии осложняют поиск и ранжирование статей. Например, дефисы используются для сокращения списка сложных слов или фраз, например, *пре- и пост-релаксация*. Такие слова могут писаться одним словом или через дефис. Читатель может установить семантическую связь в подобных ситуациях, а поисковые системы зачастую не могут. Таким образом, в поиске по запросу «*пострелаксация*» заглавия, содержащие вариант слова «*пост-релаксация*», не будут рассматриваться как высокорелевантные, поскольку они не содержат поисковый термин в точной форме.

Боуман (Bowman) [14], описывая сложности поиска заглавий с дефисами, указал, что дефисы не только снижают возможность обнаружения и рейтинг заголовка в поисковой системе или базе данных, но также влияют на сам подсчет цитирований. Идею о том, что статьи с заглавиями, содержащими специальные символы, не всегда цитируются и обнаруживаются корректно, поддерживает и исследование Ли (Lee) с соавто-

рами [15]. Чем больше специальных символов содержится в заглавии, тем больше вероятность того, что авторы последующих рукописей процитируют его некорректно. Дальнейшая логика развития событий очевидна: инструменты анализа цитирования «потеряют» некорректно оформленную цитату. Поисковые системы и базы данных, которые включают подсчет цитирования в свои алгоритмы ранжирования, будут присваивать им более низкий ранг, что приводит к снижению видимости статьи и, следовательно, к меньшему количеству потенциальных цитирований. Поэтому авторам следует выбирать название, которое сведет к минимуму такие ошибки.

Аббревиатуры, диакритические знаки и формулы

ASEO не рекомендует использовать в заглавиях не принятые в тематической области аббревиатуры, диакритические знаки из других языков и математические формулы. Если платформа, на которой размещен источник, не может отображать аббревиатуры, диакритические знаки и формулы, функциональность поиска будет нарушена. Даже будучи распознанными, такие элементы заглавия могут отображаться некорректно, что нарушает прозрачность поискового запроса. Например, математические символы с верхними и нижними индексами часто не распознаются программным обеспечением библиотечных поисковых систем и баз данных. Значимость понимания этой поисковой специфики достаточно наглядно иллюстрируется тем фактом, что даже различия между американским и британским английским могут не учитываться поисковыми системами, выдавая различное число по термину, имеющему различные варианты написания в британском и американском вариантах английского языка.

Аббревиатуры, даже широко известные, выступают в качестве частотного провокатора ошибок. В различных предметных областях они имеют разные варианты прочтения и провоцируют нерелевантную подборку источников по поисковому запросу.

Оптимизация ключевых слов

Ключевые слова, точно описывающие содержание научной публикации, значительно улучшают ее доступность при поисковом запросе. В базах данных они включаются в отдельный раздел в поисковой системе, по которому предлагается уточнить поисковый запрос или результат поиска (*Refine results*). Проработанные ключевые слова, как правило, представляют собой отдельные слова или словосочетания, позволяющие наибо-

лее полно описать исследуемую проблематику для читательской аудитории [16; 17]. При этом отдельные слова должны представлять точные термины предметного поля или уточнять основной термин, стоящий рядом. Например, ключевое слово «экстренное дистанционное образование/emergency remote learning» окажется более весомым, нежели набор ключевых слов «экстренный», «дистанционный» и «образование».

Использование синонимов ключевых слов усиливает «видимость» статьи. Полезным для определения подходящих ключевых слов является использование тезаурусов, создающих устойчивую сеть синонимов, ассоциирующихся с каждым конкретным ключевым словом. В рамках поискового запроса тезаурус верифицирует подбор источников по всем синонимам ключевого слова. Многие базы данных основывают свои алгоритмы с опорой на конкретные тезаурусы и тематические словари. Существуют тезаурусы по конкретным дисциплинам, например, тезаурусы по медицине (MeSH, Embase), геологический тезаурус, словарь географических терминов (используются при индексировании публикаций в Scopus). Разработаны инструменты, которые помогают подобрать адекватные ключевые слова к тексту, такие как Google Trends, Google Insights, Google Adwords.

Эффективной практикой является и ориентация на термины, которые читатели могут использовать для поиска публикаций по своей теме: они далеко не всегда совпадают с выбором термина автором статьи. Важно различать «узкие» и «широкие» термины с точки зрения эффективности поиска. Если в заглавии упоминается «узкий» (то есть очень конкретный) термин, в ключевых словах лучше использовать «более широкий» термин, чтобы повысить обнаруживаемость статьи в поисковых базах данных и расширить читательскую аудиторию.

Ключевое слово для поисковых систем, например, Google, лучше формулировать в единственном числе, поскольку отдельные поисковые алгоритмы воспринимают термины в единственном и множественном числе как разные слова⁵. Следует избегать расплывчатых, нерепрезентативных,

«пустых» слов, таких как «сравнение», «характеристика», «критика», «описание» и пр. В описании содержания публикации полезными являются указания на временной охват, место проведения или страну исследования. Ключевые слова должны быть значимыми, описывать контекст исследования. Кроме того, количество используемых ключевых слов должно быть разумным, чтобы избежать «спама ключевыми словами»: перенасыщенный ими текст сложно читать, а вне библиографической базы данных полнотекстовый документ статьи может быть заблокирован как спам. В большинстве руководств для авторов на сайтах журналов рекомендуется использовать до семи ключевых слов в каждой статье.

Предпочтительно не выбирать только те ключевые слова, которые наиболее популярны в тематической области⁶. Необходимо понимать, что существуют поисковые запросы, состоящие из «long tail keywords» и «short tail keywords». Менее конкурентные, но более точные ключевые слова (long tail keywords) строятся по принципу точного понимания запрашиваемой информации при помощи поисковой фразы, состоящей из нескольких слов [18]: «уровень инсулина в крови детей», или словосочетания «кровь детей, уровень инсулина», а не «инсулин», «кровь», «дети» (short tail keywords). Последний тип ключевых слов более распространен, и количество источников, подобранных по таким словам, будет исчисляться сотнями тысяч. При этом большинство из них окажутся нерелевантными. Полезно выполнять запрос в академических поисковых системах, используя каждое отобранное ключевое слово; если поиск возвращает сотни документов, правильнее выбрать более корректное ключевое слово с меньшей поисковой конкуренцией.

Сконструированные ключевые слова необходимо упоминать в «правильных» частях рукописи статьи: в заголовке, аннотации и основном тексте (но в пропорции, не раздражающей читателей). Синонимы значимых ключевых слов также должны упоминаться в тексте несколько раз, чтобы статью мог найти человек, не знакомый с наиболее распространенной терминологией, используемой в исследовательской области или применяющий синонимичные термины.

⁵ Google интерпретирует ключевые слова в единственном и множественном числе как релевантные для различных поисковых запросов, ориентируясь на различные намерения пользователей. Потребители контента зачастую используют ключевые слова: (1) в единственном числе для информационных поисковых запросов, рекомендаций или общих тем по определенному предмету; (2) во множественном числе для поиска и сравнения продуктов, веб-сайтов или услуг (*Search Engine Journal*. Available at: <https://www.searchenginejournal.com/singular-plural-keywords-google/373297/> (accessed 08.08.2023)).

⁶ В рамках любой темы существуют несколько терминов/фраз, которые широко известны и по которым часто реализуют поисковые запросы. Параллельно существует и большое количество нишевых/более узких терминов/фраз: поисковые запросы, основанные на них, по отдельности менее частотны, но вместе составляют большую часть поискового объема. Поэтому нельзя игнорировать ни ту, ни другую группу ключевых слов [18].

Некорректное конструирование ключевых слов порождает становление феномена фантомных статей: такие статьи не обнаруживаются по ключевым словам из тезаурусов⁷, поскольку авторские ключевые слова либо содержат ошибку, либо используются исключительно в границах какой-то конкретной страны и не получили распространение в международных базах данных. Как правило, подобная ситуация характерна для статей начинающих исследователей, слабо владеющих английским языком и читавших мало зарубежной литературы по своему профилю на английском языке. Один из вариантов преодоления указанной проблемы – работа с тезаурусом в своей предметной области.

Оптимизация аннотаций

Аннотация научной статьи должна быть информативной, содержать описание дизайна и методов исследования, представлять основные результаты и соответствующие термины. Существуют различные стратегии оптимизации аннотаций для обеспечения их принципиальной «обнаруживаемости», например, использование синонимов: с одной стороны, синонимы увеличивают вероятность того, что аннотация окажется релевантной разным поисковым запросам; с другой стороны, их использование снижает риск того, что статья не будет ранжирована как релевантная только потому, что определенный поисковый термин не представлен в заглавии или подзаголовке.

Наиболее значимые термины лучше включать в начало аннотации, поскольку не все базы данных и поисковые системы отображают аннотацию целиком. Если в тексте статьи авторы оперируют сокращенной формой термина, в аннотации он сначала должен быть представлен целиком с последующим указанием его сокращенной формы. Предложения в аннотации должны быть короткими и максимально информативными: «пустые» слова воспринимаются подготовленными читателями как признак низкого уровня академической грамотности автора или низкой квалификации редакторов журнала, и их нужно избегать. Аннотация должна содержать самые значимые ключевые слова и давать полное представление о содержании статьи [19]. Повтор ключевых слов в тексте аннотации способен привести к увеличению рейтинга статьи в поисковом запросе: алгоритм воспринимает обоснованный повтор ключевых слов как сигнал их особой значимости и подтверждение их статуса в качестве ключевых.

Обширные метаданные

Каждый научный документ содержит значительное количество метаданных, в том числе заглавие, подзаголовки, дату публикации, информацию об авторах, жанр статьи, название издания и т.д. Эти метаданные используются поисковым алгоритмом для оценки релевантности документа в поисковом запросе. Рейтинг документа в списке результатов поискового запроса зависит, среди прочего, от объема информации, зашифрованной в метаданных документа. Поэтому при загрузке статьи, например, в репозиторий вуза с открытым доступом или в ResearchGate через форму загрузки правильным является включение как можно большего количества метаданных.

Оптимизация метаданных изображений

Отдельные поисковые системы проводят поиск информации в тексте средств визуализации научной информации и в подписях к ним. Текст на рисунках и таблицах должен быть машиночитаемым (вместо растровых изображений следует использовать векторную графику в формате, например, .svg и .eps, содержащую текст на основе шрифтов⁸), чтобы академические поисковые системы могли их индексировать [1]. Векторная графика выглядит более профессионально и более удобна для прочтения, нежели растровая. Графика, хранящаяся в файлах JPEG, BMP, GIF, TIFF или PNG, не является векторной графикой.

Идентификаторы

Одним из широко распространенных элементов метаданных является постоянный «идентификатор цифрового объекта» (DOI). Он гарантирует, что к публикации можно будет получить постоянный доступ и, следовательно, найти ее по стабильной ссылке, даже если ее местоположение (Uniform Resource Locator [URL]) изменилось. Становление CrossRef в роли ключевого реестра научных публикаций и репозитория метаданных, перевод в открытый доступ полных текстов и полей метаданных привели к появлению ряда открытых инструментов для поиска научных публикаций (Dimensions, Semantic Scholar, Lens, SciLit и OpenAlex). Помимо исследовательских информационных систем и баз данных публикаций, библиометрические инструменты, такие как Altmetrics или PlumX, также используют DOI в качестве источника данных. Использование идентификаторов способствует повышению информативности метаданных.

⁷ Словарь, отражающий систематизированную совокупность понятий какой-либо отрасли знаний.

⁸ Эти типы файлов можно создавать в программах векторной графики, таких как Windows PowerPoint, Adobe Illustrator, Inkscape или Corel Draw.

Заключение

ASEO складывается из двух взаимосвязанных блоков активностей: предпубликационной оптимизации текста рукописи и адекватного отражения контента в метаданных. Иными словами, речь идет не только о действиях и компетенциях авторов рукописи, но и о действиях редакции каждого конкретного журнала и его издателя. Автор, понимая алгоритмы поисковой оптимизации, может значительно повысить видимость своей статьи, но если он выберет в качестве публикации площадки журнал-хищник, все его усилия не возымеют должного эффекта. Даже оптимально сбалансированная с позиций ASEO рукопись вряд ли получит активное продвижение, если редакция журнала, куда она была направлена, не понимает значимость представления

корректных и обширных метаданных статьи и не занимается оптимизацией визуализации данных с целью продвижения своего издания.

Понимание механизмов ASEO предлагает авторам механизмы оптимизации ранжирования своих исследований в основных системах информационного поиска в академической среде. Вместе с тем, большая видимость статьи не равносильна ее успеху у читательской аудитории и росту числа ее цитирований. ASEO реализуется не с целью манипулирования результатами ранжирования в контексте поискового запроса, а для обеспечения прозрачности научной коммуникации и формирования у исследователей понимания необходимости оптимизации не только метаданных своей рукописи, но и «полировки» ее текста.

КОНФЛИКТ ИНТЕРЕСОВ

Автор заявляет об отсутствии конфликта интересов.

CONFLICT OF INTERESTS

The author declares no relevant conflict of interests.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Beel J., Gipp B., Wilde E. Academic Search Engine Optimization (ASEO): Optimizing Scholarly Literature for Google Scholar and Co. *Journal of Scholarly Publishing*. 2010;41(2):176–190. <https://doi.org/10.3138/jsp.41.2.176>
2. Enge E., Spencer S., Stricchiola J. *The Art of SEO: Mastering Search Engine Optimization*. 3rd ed. O'Reilly Media; 2015. 670 p. Available at: <https://archive.org/details/artofseomasterin0000enge/mode/2up> (accessed 08.08.2023).
3. Farhadi H., Salehi H., Yunus M. M., Aghaei Chadegani A., Farhadi M., Fooladi M., Ale Ebrahim N. Does it matter which citation tool is used to compare the h-index of a group of highly cited researchers? *Australian Journal of Basic and Applied Sciences*. 2013;7(4):198–202. Available at: <https://arxiv.org/ftp/arxiv/papers/1306/1306.0727.pdf> (accessed 11.08.2023).
4. Schilhan L., Kaier C., Lackner K. Increasing visibility and discoverability of scholarly publications with academic search engine optimization. *Insights*. 2021;34(6):1–16. <https://doi.org/10.1629/uksg.534>
5. Rovira C., Codina L., Guerrero-Solé F., Lopezosa C. Ranking by relevance and citation counts, a comparative study: Google Scholar, Microsoft Academic, WoS and Scopus. *Future Internet*. 2019;11(9):202. <https://doi.org/10.3390/fi11090202>
6. Martín-Martín A., Orduña-Malea E., Ayllón J. M., López-Cózar E. D. Does Google Scholar contain all highly cited documents (1950–2013) *EC3 Working Papers*. 2015;(19):1–96. <https://doi.org/10.6084/m9.figshare.1224314.v2>
7. Rovira C., Guerrero-Solé F., Codina L. Received citations as a main SEO factor of Google Scholar results ranking. *Profesional de la Información*. 2018;27(3):559–569. <https://doi.org/10.3145/epi.2018.may.09>
8. Orduña-Malea E., Martín-Martín A., Ayllón J. M., Delgado-Lopez-Cozar E. The silent fading of an academic search engine: The case of Microsoft Academic Search. *Online Information Review*. 2014;38(7):936–953. <https://doi.org/10.1108/OIR-07-2014-0169>
9. Jacques T. S., Sebire N. J. The impact of article titles on citation hits: An analysis of general and specialist medical journals. *JRSM Short Reports*. 2010;1(1):1–5. <https://doi.org/10.1258/shorts.2009.100020>
10. Jamali H. R., Nikzad M. Article title type and its relation with the number of downloads and citations. *Scientometrics*. 2011;88(2):653–661. <https://doi.org/10.1007/s11192-011-0412-z>

11. Тихонова Е.В., Косычева М.А. Заглавие научной статьи: лучшие практики конструирования. *Хранение и переработка сельхозсырья*. 2022;(3):11–25. (In Russ.). <https://doi.org/10.36107/spfp.2022.374>
Tikhonova E.V., Kosycheva M.A. Research paper title: Best design practices. *Storage and Processing of Farm Products*. 2022;(3):11–25. (In Russ.). <https://doi.org/10.36107/spfp.2022.374>
12. Booth A. Clear and present questions: Formulating questions for evidence based practice. *Library Hi Tech*. 2006;24(3):355–368. <https://doi.org/10.1108/07378830610692127>
13. Booth A., Noyes J., Flemming K., et al. Formulating questions to explore complex interventions within qualitative evidence synthesis. *BMJ Global Health*. 2019;4:e001107. <https://doi.org/10.1136/bmjgh-2018-001107>
14. Bowman J.H. The Catalog as Barrier to Retrieval—Part 1: Hyphens and Ampersands in Titles. *Cataloging & Classification Quarterly*. 2000;29(4):39–60. https://doi.org/10.1300/J104v29n04_04
15. Lee D.T.S., Zhou Z., Tse T.H. Metamorphic robustness testing of Google Translate. In: *Proceedings – 2020 IEEE/ACM 42nd International Conference on Software Engineering Workshops, ICSEW 2020*. New York, United States: Association for Computing Machinery; 2020, pp. 388–395. <https://doi.org/10.1145/3387940.3391484>
16. Тихонова Е.В., Косычева М.А. Эффективные ключевые слова: стратегии формулирования. *Health, Food & Biotechnology*. 2021;3(4):7–15. <https://doi.org/10.36107/hfb.2021.i4.s122>
Tikhonova E.V., Kosycheva M.A. Effective Keywords: Strategies for their formulation. *Health, Food & Biotechnology*. 2021;3(4):7–15. (In Russ.). <https://doi.org/10.36107/hfb.2021.i4.s122>
17. Chen Y.-N., Ke H.-R. A study on mental models of taggers and experts for article indexing based on analysis of keyword usage. *Journal of the Association for Information Science and Technology*. 2014;65(8):1675–1694. <https://doi.org/10.1002/asi.23077>
18. Johanson J. *Search Engine Optimization and the long tail of Web Search*. Uppsala: Uppsala University; 2016. 27 p. Available at: <https://cl.lingfil.uu.se/exarb/arch/johanssond2016.pdf> (accessed 08.08.2023).
19. Кириллова О. В. *Редакционная подготовка научных журналов по международным стандартам: рекомендации эксперта БД Scopus*. М.: Нобель пресс; 2013. 90 с. Режим доступа: <https://rassep.ru/academy/biblioteka/106592/> (дата обращения: 11.08.2023).

ИНФОРМАЦИЯ ОБ АВТОРЕ

Елена Викторовна Тихонова, кандидат исторических наук, доцент, МГИМО Университет МИД РФ, г. Москва, Российская Федерация; заведующий редакцией научного журнала «Journal of Language and Education», НИУ «Высшая школа экономики», г. Москва, Российская Федерация; <https://orcid.org/0000-0001-8252-6150>; e-mail: etihonova@gmail.com

INFORMATION ABOUT THE AUTHOR

Elena V. Tikhonova, Cand. Sci. (Hist.), Associated Professor, MGIMO University, Moscow, Russian Federation; Head of the Editorial Office of the *Journal of Language and Education*, HSE University, Moscow, Russian Federation; <https://orcid.org/0000-0001-8252-6150>; e-mail: etihonova@gmail.com

Поступила в редакцию / Received 05.09.2023

Поступила после рецензирования / Revised 09.10.2023

Принята к публикации / Accepted 20.10.2023