

УДК 330

АЛГОРИТМ ИСПОЛЬЗОВАНИЯ КРАУДСОРСИНГОВЫХ ИНСТРУМЕНТОВ ДЛЯ СБОРА И АНАЛИЗА ДАННЫХ НАУЧНЫХ ИССЛЕДОВАНИЙ (НА ПРИМЕРЕ ПОДГОТОВКИ СИСТЕМАТИЗИРОВАННОГО ОБЗОРА ЛИТЕРАТУРЫ)

Таран Е.А., Маланина В.А., Касати Ф.

Томский политехнический университет

Аннотация.

Количество публикаций, посвященных проведению систематизированного обзора литературы, за последние годы имеет устойчивую тенденцию к росту и сопровождается высокой цитируемостью данного типа публикаций в реферативных базах данных. Опрос исследователей, имеющих опыт работы по составлению систематизированного обзора литературы, позволил определить, что выполнение данного трудоемкого процесса сопровождается наличием достаточно большого количества трудностей. Целью данной статьи является разработка алгоритма использования краудсорсинговых инструментов для сбора и анализа научных исследований (на примере подготовки систематизированного обзора литературы). На основе разработанного и валидированного алгоритма создан шаблон для краудсорсинговой платформы Яндекс.Толока. Он включает в себя следующие этапы: определение критериев отбора; определение ключевых слов для поиска; проверка правильности вопросов; отбор вопросов для краудсорсинга; создание инструкции для исполнителей; размещение инструкции на краудсорсинговой платформе; создание тестового/тренировочного пула статей; составление рейтинга вопросов для краудсорсинга; определение стоимости работ исполнителей; создание контрольных элементов; запуск краудсорсинга. Использование шаблона для осуществления систематизированного обзора литературы при помощи инструментов краудсорсинга на платформе Яндекс.Толока позволит существенно сократить работу исследователей, направленную на отбор статей для дальнейшего анализа.

Информация о статье

Принята 05 ноября 2020 г.

Ключевые слова: систематизированный обзор литературы, краудсорсинг, шаблон, алгоритм, Яндекс.Толока

DOI: 10.26730/2587-5574-2020-4-39-46

ALGORITHM FOR USING CROWDSOURCING TOOLS FOR COLLECTING AND ANALYSIS OF SCIENTIFIC RESEARCH DATA (BY THE EXAMPLE OF PREPARATION OF SYSTEMATIZED LITERATURE REVIEW)

Ekaterina A. Taran, Veronika A. Malanina, Fabio Casati

Tomsk Polytechnic University

Abstract.

The number of publications devoted to conducting a systematic literature review in recent years has a steady upward trend and is accompanied by a high citation rate of this type of publications in abstract databases. A survey of researchers with experience in compiling a systematic review of the literature made it possible to determine that the implementation of this laborious process is accompanied by the presence of a fairly large number of difficulties. The purpose of this article is to develop an algorithm for using crowdsourcing tools for collecting and analyzing scientific research (using the example of preparing a systematized literature review). Based on the developed and validated algorithm, a template was created for the crowdsourcing platform Yandex.Toloka. It includes the following stages: determination of selection criteria; definition of keywords for search; checking the correctness of the questions; selection of questions for crowd-sourcing; creation of instructions for performers; placement of instructions on a crowdsourcing platform; creation of a test / training pool of articles; rating of questions for crowdsourcing; determination

Article info

Received December 05, 2020

Keywords:

systematic literature review, crowdsourcing, template, algorithm, Yandex.Toloka.

of the cost of performers' work; creation of control elements; launch of crowdsourcing. Using a template to carry out a systematic review of literature using crowdsourcing tools on the platform Yandex.Toloka will significantly reduce the work of researchers aimed at selecting articles for further analysis.

1. Introduction / Введение

Систематизированный обзор литературы является одним из наиболее важных видов научных публикаций, который позволяет обобщить результаты, достигнутые различными исследовательскими группами [1- 3], и направлен на объективный анализ большого количества статей по заданной теме исследования, что в итоге позволяет предоставить исчерпывающий и объективный обзор исследований в данной области [4]. Кроме того, систематизированный обзор литературы является одним из наиболее цитируемых видов публикации и его популярность неуклонно растет как в медицинских, инженерных, так и в социально-гуманитарных исследованиях [5-9].

2. Materials and methods / Материалы и методы

Выполнение систематизированного обзора литературы – это долгий и трудоемкий процесс, в котором достаточно велик риск совершить ошибку. Данные организованного нами опроса исследователей, имеющих опыт работы по составлению систематизированного обзора литературы, подтверждают это. Проведенные интервью с исследователями позволили выявить основные трудности, с которыми сталкиваются ученые при составлении систематизированного обзора литературы и получить обратную связь от исследователей, столкнувшихся с этими трудностями. На платформе SurveyMonkey была создана форма опроса, которая включала закрытые вопросы, описывающие количественные параметры проведенных респондентами исследований по систематизированному обзору литературы (количество участников, составляющих обзор и вычитывавших статьи, объем времени, затраченного на проведение всех этапов обзора в целом и каждого в отдельности). Также для респондентов были сформулированы открытые вопросы, в ответах на которые они указывали трудности, возникшие в процессе проведения исследования, и возможные способы их преодоления. Были получены ответы на анкету от исследователей, имеющих опыт составления систематизированного обзора литературы в сфере медицины, социально-гуманитарных наук, психологии, инженерных и компьютерных наук, междисциплинарных исследований. Для целей данного исследования были также проведены Skype-интервью с исследователями, заполнившими анкеты. Данные интервью были организованы в форме неструктурированных бесед. Неструктурированность была продиктована различиями между областями наук, в которых респонденты проводят исследования. Это позволило получить расширенные комментарии респондентов относительно организованных ими систематизированных обзоров литературы. Ответы респондентов свидетельствуют, что в среднем от момента начала сбора информации до момента подачи ее в виде статьи для публикации в среднем прошло 12 месяцев (значения в пределах от 3 до 36 месяцев). На определение критериев включения статей в обзор, целей и задач исследования в среднем уходил месяц, на процесс отбора в соответствии с заявленными критериями - три, на вычитку прошедших отбор статей и их анализ - в среднем 7 месяцев. Количество исследователей, вычитывавших каждую статью в процессе обзора литературы, колеблется от 1 до 3 человек. Для целей данного исследования важным было выявить характер трудностей, с которыми они столкнулись:

- большое количество статей, проходящих первоначальный отбор по ключевым словам;
- необходимость корректировки критериев отбора;
- выработка успешной стратегии поиска;
- проблемы доступа к полнотекстовым версиям статей;
- получение исходных данных, на которых базируются написанные статьи;
- синтез полученных результатов по итогам прочтения всех прошедших отбор статей;
- разные мнения внутри команды исследователей по вопросам выбора критериев;
- слишком большое разнообразие методов и приемов, используемых авторами статей, необходимость их группировки и точной классификации;
- обеспечение поиска всех релевантных статей;

- низкое качество публикаций и данных;
- плохая автоматизация процесса отбора статей.

Таким образом иногда процесс обзора литературы становится настолько долгим, что работа не завершается и не становится общедоступной, тем самым она не выполняет свою полезную функцию для научного сообщества. С каждым годом все больше и больше научных работ публикуется, что усложняет задачу их анализа [10]. В итоге, уже опубликованные систематизированные обзоры литературы пропускают порядка 30 % статей, относящихся к теме заявленного исследования [11, 12].

Использование краудсорсинга для упрощения/сокращения процедуры проведения систематизированного обзора литературы является одним из способов преодоления сложившейся ситуации [13-15].

В статье мы опишем опыт работы с такой платформой как Яндекс.Толока [16]. Данный краудсорсинговый проект еще достаточно молодой, он запущен 6 лет назад. Основной целью проекта является осуществление быстрой разметки большого количества данных, которые затем используются для машинного обучения и совершенствования поисковых алгоритмов. Как правило, размещаемые задания достаточно простые, для их исполнения не требуется специальная подготовка. В основном они связаны с анализом и оценкой контента. Поэтому любой желающий может выполнять задания и получать за них вознаграждение. Однако в большинстве случаев использование данной платформы осуществляется для решения маркетинговых задач бизнеса. Поэтому все существующие шаблоны направлены на соответствующие запросы заказчиков. Мы же попытались использовать эту платформу для решения части научной задачи исследователя, работающего над обзором литературы. Цель нашего исследования – разработать шаблон систематизированного обзора литературы, описывающего алгоритм использования краудсорсинга для сбора и анализа данных научных исследований, и позволяющего сократить время подготовки обзора литературы и связанных с этим финансовых затрат.

3. Results and discussion / Результаты и обсуждение

На основе разработанного и валидированного алгоритма использования краудсорсинговых инструментов для сбора и анализа данных научных исследований (на примере подготовки систематизированного обзора литературы) был создан Шаблон. Данный Шаблон предназначен для конкретного этапа систематизированного обзора литературы, представляющего наиболее монотонный/утомительный/длительный процесс по отбору (скринингу) источников для последующего анализа (рисунок 1).



Рис. 1. Шаблон систематизированного обзора литературы (при использовании краудсорсинга на платформе Яндекс.Толока)

Первым этапом является определение критериев отбора, позволяющих осуществить выбор необходимых публикаций для последующего анализа. Например, для нашего систематизированного обзора литературы были определены следующие критерии включения/исключения:

C1: Does the paper describe an original study that employs quantitative methods?

C2: Does the study or analysis in the paper explicitly include measures of inequality?

C3: Does the study or analysis described in the paper include some measures of well-being?

C4: Does the paper present quantitative data regarding the association between inequality and well-being?

C5: Is the paper/the article published in a peer-reviewed journal?

C6: Does the study involve older adults, or enables the separate analysis of data for older adults?

Следующий этап «Определение ключевых слов для поиска» предполагает определение ключевых слов по теме исследования. На данном этапе составляется запрос для поиска статей по библиографическим и реферативным базам данных (Scopus, Web of Science, Elibrary и др.). В данном запросе отражаются ключевые слова-включения, слова-исключения, а также задаются дополнительные параметры (тип публикации: статья, обзор и т.д.). В результате получаем исходный набор статей-кандидатов для исследования. Например, для систематического обзора литературы, посвященного вопросам неравенства пожилых людей, был осуществлен поиск статей в базе данных Scopus, на основе данного запроса: слова-включения (wellbeing, well-being, life satisfaction, happiness, subjective well-being, quality of life, inequality, disparity, distribution, redistribution, older adults, elderly), слова-исключения (cancer, osteoporosis, diabetes, femoris, heart failure) и дополнительные параметры (article, review). В результате было получено 2619 статей.

Описанные выше этапы являются стандартными для осуществления систематизированного обзора литературы, и применяются как с использованием традиционных, так и краудсорсинговых технологий.

Третий этап «Проверка правильности вопросов» требует особого внимания со стороны исследователей, независимо от того какие инструменты (традиционные или краудсорсинговые) они планируют применять в дальнейшей организации научного поиска. Распространенная ошибка данного этапа - прямое разделение работы между исследователями для осуществления проверки аннотаций статей. По нашему опыту, это является плохой практикой, так как очень часто первоначальное определение критерия может быть неоднозначным и интерпретироваться по-разному. Поэтому от качества сформулированных критериев зависит успех дальнейшего исследования. Для проверки правильности сформулированных вопросов исследователям необходимо взять 10-20 статей и попытаться оценить предикаты, сравнить полученные результаты и внести необходимые корректировки. Это позволит прийти к более точному определению предиката, что в дальнейшем позволит доступно его описать для краудсорсинга. Так, в рамках нашего исследования первоначальные критерии включения/исключения были подвергнуты проверке на правильность их формулировки. В итоге мы получили следующий их обновленный вариант:

C1: Does the study use quantitative methods?

C2: Does the study include some measures of economic inequality?

C3: Does the study or analysis described in the paper include some measures of subjective well-being?

C4: Does the paper present quantitative data regarding the association between economic inequality and subjective well-being?

C5: Is the paper/the article published in a peer-reviewed journal?

C6: Does the paper describe a study involving older adults 55+ years?

На этапе «Отбор вопросов для краудсорсинга» необходимо правильно определить перечень вопросов, которые можно будет передать исполнителям на платформе Яндекс.Толока. Это должны быть вопросы простые для понимания и позволяющие дать максимально однозначный ответ. Например, при отборе вопросов для краудсорсинга в нашем исследовании были исключены C5 (так как ответ на данный вопрос был решен еще на 2 шаге при помощи инструментов, доступных в реферативной базе данных), C4 (так как данный вопрос решается получением ответов на C2 и C3). В результате были отобраны следующие вопросы:

C1: Does the study use quantitative methods?

C2: Does the study include some measures of economic inequality?

C3: Does the study or analysis described in the paper include some measures of subjective well-being?

C6: Does the paper describe a study involving older adults 55+ years?

Несколько небольших взаимосвязанных проектов, позволяют упростить задание для исполнителей на краудсорсинговой платформе, а значит снизить стоимость и повысить качество результатов.

После определения перечня вопросов, которые будут переданы на краудсорсинг, наступает этап создания инструкции для выполнения задания на краудсорсинговой платформе. Инструкция создается для каждого вопроса отдельно и должна включать в себя следующие основные элементы: инструкцию по выполнению задания, описание самого задания; вопрос с пояснениями (объяснение искомого параметра) и положительные и отрицательные примеры (с обязательным объяснением).

Разработанные материалы для инструкции размещаем на краудсорсинговой платформе Яндекс.Толока. На данном этапе важно проработать интерфейс задания, который определяет внешний вид задания для исполнителя и логику обработки ответов. Хороший интерфейс повысит качество результатов и позволит исполнителям выполнять задания быстрее, а это позволит вам снизить цену за задание. На рисунке 2 представлен пример инструкции, который видит исполнитель.

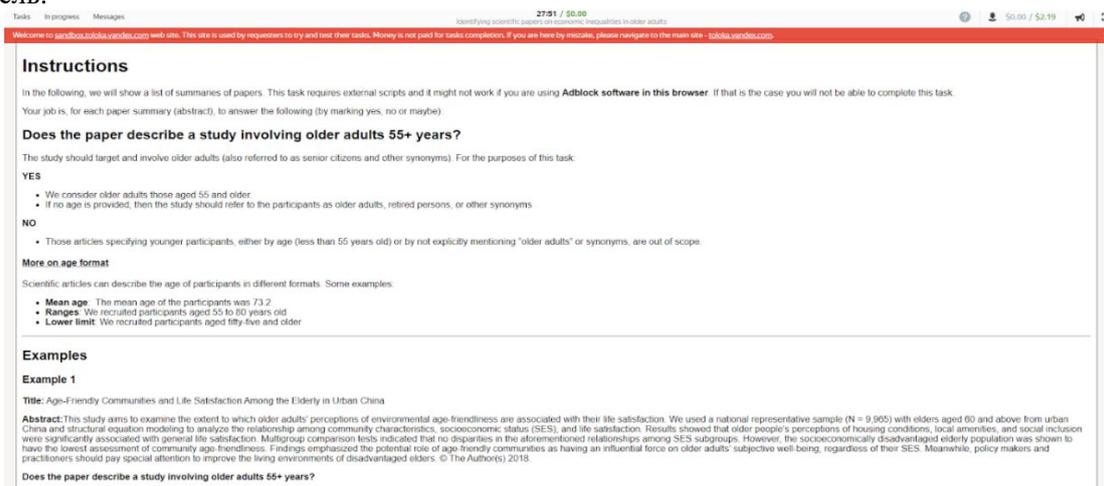


Рис. 2. Инструкция для исполнителя

Данная инструкция размещается для каждого критерия, выбранного на этапе 4. Фактическое задание, которое видит каждый исполнитель, будет представлять собой набор статей. При этом исполнитель сосредотачивается только на одном критерии отбора статей. Данный момент является важным, так как объяснение различных критериев одному и тому же исполнителю, и последующая проверка понимания этих вопросов, показала, что более эффективно сосредоточиться на одном вопросе. Это сокращает время обучения исполнителя и сосредотачивает внимание каждого исполнителя на ключевом показателе.

Затем для каждого вопроса создаем небольшой набор тестовых статей, который используется для проверки понимания исполнителем вопроса и способен ли он правильно выполнить задание. На данном этапе реализуются два важных момента. Во-первых, тестируются созданные инструкции, а полученные результаты определяют алгоритм дальнейших действий. При этом необходимо определить следующие параметры: вопросы, которые были легче всего для выполнения (наименьшее число разногласий среди исполнителей), затраченное время на выполнение задания. Во-вторых, происходит отбор потенциальных исполнителей, которые впоследствии будут допущены к работе с основной базой публикаций. В нашем исследовании исполнитель получал 3 статьи для обучения (они выпадали случайным образом из созданного пула обучающих статей). И, только при получении 3 правильных ответов исполнитель допускался к работе. Например, для С1 во время учебы было получено 468 ответов от 156 исполнителей и только 43 исполнителя были допущены до работы с основным пулом статей; для С2 - 393 ответа от 131 исполнителя во время обучения, а допущено к работе 35 исполнителей. Важность данного шага определяется еще и тем, что за выполнение заданий во время обучения оплата не производится. Таким образом, данный этап позволяет допускать к оплачиваемой работе только добросовестных исполнителей, правильно понявших поставленную задачу.

Составление рейтинга вопросов для краудсорсинга позволяет сократить как временные, так и финансовые затраты на осуществление отбора статей для анализа при помощи краудсорсинга.

Составляем рейтинг самых легких вопросов по результатам 7 этапа. Например, в нашем исследовании:

1. C6: Does the paper describe a study involving older adults 55+ years?
2. C2: Does the study include some measures of economic inequality?
3. C1: Does the study use quantitative methods?
4. C3: Does the study or analysis described in the paper include some measures of subjective well-being?

Для определения стоимости работ исполнителей необходимо сначала определить время принятия решения по каждому вопросу. Этот показатель рассматривается как мера рабочих усилий. Используем специальный код JavaScript для визуализации экспериментальных условий. Этот код отвечает за сбор данных каждый раз, когда исполнитель нажимает один из возможных ответов на предикат. Здесь также регистрируется время завершения загрузки интерфейса задачи, а затем полученные данные используются для определения временных меток для вычисления времени принятия решения для каждого документа, используя время загрузки страницы в качестве ориентира. Далее рассчитывается время принятия решения для исполнителей в каждом из условий эксперимента краудсорсинга. Для каждого документа используем предикат «более медленный» как время принятия решения. Эти результаты показывают, что запрос одного (более простого) предиката не всегда приводит к сокращению времени принятия решения по сравнению с запросом сложного. Однако более простые предикаты помогают более точно оценить усилия из-за их детализации, которая влияет на задержку выполнения задания. Так, например, для обучающего задания по критерию C1 время принятия решения составило 11,062 секунды, а для C2 - 16,908.

Данный показатель является важным для дальнейшего определения стоимости работы исполнителя. Цена указывается в долларах США из расчета оплаты за 1 страницу заданий (например, 3 аннотации на 1 странице). Для выполнения данного условия необходимо классифицировать все ваши аннотации по количеству слов на длинные (более 230 слов) и короткие (менее 230 слов). Минимальная цена на платформе составляет \$ 0.01. Комиссия – 20 % от цены, но не меньше \$ 0.005 [17]. В настройках шаблона использовали фиксированное ценообразование для каждого критерия, позволяющего установить единую стоимость одной страницы заданий. Таким образом, цена устанавливалась одинаковая для всех исполнителей, независимо от их навыков и опыта.

Следующий важный шаг – это создание контрольных элементов. Необходимо определить для каждого вопроса перечень аннотаций, позволяющих контролировать качество выполнения задания исполнителем. Контроль качества позволяет получать более точные ответы и ограничивать доступ к заданиям недобросовестным исполнителям. Чтобы следить за тем, как часто ошибаются исполнители в нашем шаблоне используются контрольные задания, позволяющие на основе полученных ответов исполнителя на контрольные задания, блокировать доступ исполнителям, которые дают неправильные ответы. Согласно рекомендациям платформы Яндекс.Толока в пул контрольных заданий следует добавлять не менее 1% от общего числа заданий [18].

Запускаем первый критерий из составленного рейтинга на краудсорсинг, это позволит сократить количество статей (и затраты), которые в дальнейшем поступят для анализа по второму критерию из рейтинга и т.д., так как несоответствие статьи хотя бы одному критерию приводит к исключению данного материала из базы для дальнейшего анализа. При осуществлении регистрации на Яндекс.Толока необходимо обратить внимание на выбор платформы, предназначенной именно для заказчиков.

4. Conclusion / Заключение

Таким образом, использование данного шаблона для осуществления систематизированного обзора литературы при помощи инструментов краудсорсинга на платформе Яндекс.Толока позволит существенно сократить работу исследователей, направленную на отбор статей для дальнейшего анализа.

5. Acknowledgement / Признательность

Исследование выполнено за счет гранта Российского научного фонда (проект №19-18-00282)

Список источников

1. Grant M., Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies // *Health Info Library Journal*. – 2009. – Vol. 26(2). – pp. 91-108.
2. Khan K., Kunz R., Kleijnen J., Antes G. Five steps to conducting a systematic review // *Journal of the Royal Society of Medicine*. – 2003. Vol. 96. – pp. 118-121.
3. Harris J., Quatman C., Manring M., Siston R., Flanigan D. How to write a systematic review // *The American journal of sports medicine*. – 2014. Vol. 42(11). – pp. 2761-2768.
4. Haidich A. Meta-analysis in medical research // *Hippokratia*. – 2010. – Vol. 14. – pp. 29-37.
5. Сенбеков М.Т., Токмурзиева Г.Ж. Совершенствование методов управления медицинским образованием. Обзор литературы // *Наука и здравоохранение*. – 2020. – № 1. – С. 27-38.
6. Соколов Д.В. Обзор новой литературы // *Управление наукой: теория и практика*. – 2020. – № 1. – С. 218-220.
7. Литвинова В.В. Финансовое поведение населения: статистика и обзор литературы // *Экономические исследования*. – 2019. – № 3. – С. 2.
8. Мокина М.И. Факторы роста семейного бизнеса в России: обзор литературы и направления будущих исследований // *Стратегические решения и риск-менеджмент*. – 2019. – № 1. – С. 88-97.
9. Krivosheev E., Benatallah B., Casati F. Crowd-based Multi-predicate Screening of Papers in Literature Reviews / *International World Wide Web Conferences Steering Committee*, 2018. URL: <https://arxiv.org/pdf/1803.09814v1.pdf> (дата обращения: 01.10.2020).
10. Krivosheev E., Caforio V., Benatallah B., Casati F. Crowdsourcing Paper Screening in Systematic Literature Reviews // *Proceedings of the fifth AAAI Conference on Human Computation and Crowdsourcing*, 2017. URL: <https://arxiv.org/pdf/1709.05168.pdf> (дата обращения: 01.10.2020).
11. Crequit P., Trinquart L., Yavchitz A., Ravaud P. Wasted research when systematic reviews fail to provide a complete and up-to-date evidence synthesis: the example of lung cancer // *BMC Medicine*, 2016. URL: <https://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-016-0555-0> (дата обращения: 01.10.2020).
12. Eickhoff C., Vries A. Increasing cheat robustness of crowdsourcing tasks // *Information retrieval* – 2012. – № 16(2) – pp. 121–137.
13. Русанова А.А., Таран Е.А., Гуменников И.В., Касати Ф. Современные методы анализа и применение краудсорсинговых инструментов в области социально-экономических исследований // *MODERN ECONOMY SUCCESS*. – 2019. – № 5. – С. 119-125.
14. Garneau N., Nuessle T., Sloan M., Santorico S., Coughlin B., Hayes J. Crowdsourcing taste research: genetic and phenotypic predictors of bitter taste perception as a model // *Frontiers in integrative neuroscience*. – 2014. URL: <file:///C:/Users/ektar/Downloads/fnint-08-00033.pdf> (дата обращения: 01.10.2020).
15. Edgar J., Murphy J., Keating M. Comparing Traditional and Crowdsourcing Methods for Pretesting Survey Questions // *First Publisher*. – 2016. – URL: <https://journals.sagepub.com/doi/pdf/10.1177/2158244016671770> (дата обращения: 01.10.2020).
16. Яндекс.Толока для заказчиков. URL: <https://toloka.ai/ru> (дата обращения: 20.11.2020).
17. Настройка ценообразования. Яндекс.Толока для заказчиков. URL: <https://yandex.ru/support/toloka-requester/concepts/dynamic-pricing.html> (дата обращения: 20.11.2020).
18. Контрольные задания. Яндекс.Толока для заказчиков. URL: <https://yandex.ru/support/toloka-requester/concepts/goldenset.html> (дата обращения: 20.11.2020).

References

1. Grant M., Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Library Journal*. 2009. Vol. 26(2). pp. 91 108.
2. Khan K., Kunz R., Kleijnen J., Antes G. Five steps to conducting a systematic review. *Journal of the Royal Society of Medicine*. 2003. Vol. 96. pp. 118 121.
3. Harris J., Quatman C., Manring M., Siston R., Flanigan D. How to write a systematic review. *The American journal of sports medicine*. 2014. Vol. 42(11). pp. 2761 2768.
4. Haidich A. Meta-analysis in medical research. *Hippokratia*. 2010. Vol. 14. pp. 29 37.
5. Senbekov M.T., Tokmurzieva G.ZH. Sovershenstvovanie metodov upravleniya me-dicinskim obrazovaniem. *Obzor literatury [Improvement of medical education management methods. Literature review]*. *Nauka i zdravoohranenie = Science and health*. 2020. Vol. 1. pp. 27 38.
6. Sokolov D.V. *Obzor novej literatury [Review of new literature]*. *Upravlenie naukoy: teoriya i praktika = Management of science: theory and practice*. 2020. Vol. 1. pp. 218-220.
7. Litvinova V.V. *Finansovoe povedenie naseleniya: statistika i obzor literatury [Financial behavior of the population: statistics and literature review]*. *Ekonomicheskie issledovaniya = Economic research*. 2019. Vol. 3. pp. 2.

8. Mokina M.I. Faktory rosta semejnogo biznesa v Rossii: obzor literatury i napravleniya budushchih issledovanij [Growth factors of family business in Russia: a review of the literature and directions for future research]. Strategicheskie resheniya i risk-menedzhment = Strategic decisions and risk management. 2019. Vol. 1. pp. 88-97.
9. Krivosheev E., Benatallah V., Casati F. Crowd-based Multi-predicate Screening of Pa-pers in Literature Reviews. International World Wide Web Conferences Steering Committee, 2018. URL: <https://arxiv.org/pdf/1803.09814v1.pdf> (last access: 01.10.2020).
10. Krivosheev E., Caforio V., Benatallah V., Casati F. Crowdsourcing Paper Screening in Systematic Literature Reviews. Proceedings of the fifth AAAI Conference on Human Computation and Crowdsourcing, 2017. URL: <https://arxiv.org/pdf/1709.05168.pdf> (last access: 01.10.2020).
11. Crequit P., Trinquart L., Yavchitz A., Ravaud P. Wasted research when systematic reviews fail to provide a complete and up-to-date evidence synthesis: the example of lung cancer. BMC Medicine. 2016. URL: <https://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-016-0555-0> (last access: 01.10.2020).
12. Eickhoff C., Vries A. Increasing cheat robustness of crowdsourcing tasks. Information retrieval. 2012. Vol. 16(2). pp. 121–137.
13. Rusanova A.A., Taran E.A., Gumennikov I.V., Kasati F. Sovremennye metody analiza i primeneniye kraudsorsingovyh instrumentov v oblasti social'no-ekonomicheskikh issledovanij [Modern methods of analysis and the use of crowdsourcing tools in the field of socio-economic research]. MODERN ECONOMY SUCCESS. 2019. Vol. 5. pp. 119-125.
14. Garneau N., Nuesse T., Sloan M., Santorico S., Coughlin B., Hayes J. Crowdsourcing taste research: genetic and phenotypic predictors of bitter taste perception as a model. Frontiers in integrative neuroscience. 2014. – URL: <file:///C:/Users/ektar/Downloads/fnint-08-00033.pdf> (last access: 01.10.2020).
15. Edgar J., Murphy J., Keating M. Comparing Traditional and Crowdsourcing Methods for Pretesting Survey Questions. First Publisher. 2016. URL: <https://journals.sagepub.com/doi/pdf/10.1177/2158244016671770> (last access: 01.10.2020).
16. Yandex.Toloka dlya zakazchikov [Yandex.Toloka for customers]. URL: <https://toloka.ai/ru> (last access: 20.11.2020).
17. Nastrojka cenoobrazovaniya. Yandex.Toloka dlya zakazchikov [Setting up pricing. Yandex.Toloka for customers]. URL: <https://yandex.ru/support/toloka-requester/concepts/dynamic-pricing.html> (last access: 20.11.2020).
18. Kontrol'nye zadaniya. YAndeks.Toloka dlya zakazchikov [Control tasks. Yandex.Toloka for customers]. URL: <https://yandex.ru/support/toloka-requester/concepts/goldenset.html> (last access: 20.11.2020).

Авторы

Таран Екатерина Александровна – канд. экон. наук, старший преподаватель
Томский политехнический университет
634050, г. Томск, пр. Ленина, 30
E-mail: ektaran@tpu.ru

Маланина Вероника Анатольевна – канд. экон. наук, доцент
Томский политехнический университет
634050, г. Томск, пр. Ленина, 30
E-mail: milanskaya@mail.ru

Касати Фабио – PhD, ведущий научный сотрудник
Томский политехнический университет
634050, г. Томск, пр. Ленина, 30
E-mail: kasati@tpu.ru

Authors

Ekaterina A. Taran – PhD, Senior Lecturer
Tomsk Polytechnic University
634050, Tomsk, Lenin Ave., 30
E-mail: ektaran@tpu.ru

Veronika A. Malanina – PhD, Assistant Professor
Tomsk Polytechnic University
634050, Tomsk, Lenin Ave., 30
E-mail: milanskaya@mail.ru

Fabio Casati – PhD, Leading Researcher
Tomsk Polytechnic University
634050, Tomsk, Lenin Ave., 30
E-mail: kasati@tpu.ru

Библиографическое описание статьи

Таран Е.А., Маланина В.А., Касати Ф. Алгоритм использования краудсорсинговых инструментов для сбора и анализа данных научных исследований (на примере подготовки систематизированного обзора литературы) // Экономика и управление инновациями — 2020. — № 4 (15). — С. 39-46.

Reference to article

Taran E.A., Malanina V.A., Casati F. Algorithm for using crowdsourcing tools for collecting and analysis of scientific research data (by the example of preparation of systematized literature review). Economics and Innovation Management, 2020, no. 4 (15), pp. 39-46.