

УДК: 004.02

DOI: <https://doi.org/10.47813/2782-5280-2022-1-1-0229-0237>

Обзор инструментов кластеризации в SEO-проектировании

А.А. Яблокова

*Красноярский краевой Дом науки и техники Российского союза научных и инженерных общественных объединений, г. Красноярск, Россия
Красноярский государственный аграрный университет, г. Красноярск, Россия*

Аннотация. В статье приведена краткая характеристика методов кластеризации, рассмотрен ряд инструментов кластеризации ключевых слов в SEO – проектировании в соответствии с методами кластеризации: логической группировки, семантической группировки и группировки по топам поисковой выдачи. Отмечается, что для каждого метода кластеризации предусмотрен ряд специальных инструментов, даны рекомендации для успешного разделения ключевых слов на кластеры. Рекомендуется первоначально разбить ключи с помощью утилит на основе семантической схожести и определить интент, а затем соотнести результаты с результатами группировки по топам поисковой выдачи той системы, под которую будет оптимизирован ресурс. Далее на основе корреляции результатов вносятся изменения в кластеры.

Ключевые слова: семантическое ядро, кластеризация, логическая группировка, группировка по топам SERP, группировка по семантической схожести, инструменты кластеризации

Для цитирования: Яблокова А. (2022). Обзор инструментов кластеризации в SEO-проектировании. Информатика. Экономика. Управление - Informatics. Economics. Management, 1(1), 0229–0237. <https://doi.org/10.47813/2782-5280-2022-1-1-0229-0237>

Tools for keyword clustering in SEO

A.A. Yablokova

*Krasnoyarsk Regional Science and Technology City Hall of the Russian Union of Scientific and Engineering Public Associations, Krasnoyarsk, Russia
Krasnoyarsk State Agrarian University, Krasnoyarsk, Russia*

Abstract. The article provides a brief description of clustering methods, considers a number of keyword clustering tools in SEO - design in accordance with clustering methods: logical grouping, semantic grouping and grouping by search results tops. It is noted that for each clustering method a number of special tools are provided, recommendations are given for the successful division of keywords into

clusters. It is recommended to initially split the keys using utilities based on semantic similarity and determine the intent, and then correlate the results with the results of grouping by the tops of the search results of the system for which the resource will be optimized. Further, based on the correlation of the results, changes are made to the clusters.

Keywords: semantic core, clustering, logic-based keyword clustering, SERP-based keyword clustering, semantic-based keyword clustering, tools for keyword clustering

For citation: Yablokova A. (2022). Tools for keyword clustering in SEO. Informatics. Economics. Management, 1(1), 0229–0237. <https://doi.org/10.47813/2782-5280-2022-1-1-0229-0237>

ВВЕДЕНИЕ

Вопросам семантического поиска, классификации, структуризации и интеграции информации в контексте SEO – проектирования уделяется достаточно много внимания исследователями в области информационных технологий [1-3]. Одним из первых этапов разработки сайта является именно этап SEO – проектирования, в рамках которого закладывают основу для последующей индексации и продвижения ресурса через SERP. На этапе SEO – проектирования составляют семантическое ядро из ключевых слов, характеризующих тематику ресурса и распределяемых на кластеры в соответствии с общим свойством группируемых слов.

Выделяют три метода кластеризации: логическая группировка, семантическая группировка, группировка по топам поисковой выдачи. Метод кластеризации ключевых слов выбирают исходя из объема семантического ядра, тематики и типа сайта и полноты соответствия общему свойству, заданному в качестве тематического маркера [4].

Рассмотрим инструменты группировки ключевых слов на кластеры в соответствии с разбивкой по методам кластеризации.

ЛОГИЧЕСКАЯ ГРУППИРОВКА

Метод логической группировки используют для простых коммерческих сайтов, (сайт-визитка, лендинг, веб-портфолио, промо-сайт, сайт-витрина), семантические ядра которых содержат до нескольких сотен ключевых слов.

Группировка проводится вручную в одну или несколько итераций: выбирается тематический маркер, который, как правило, является одним из высокочастотных ключей из семантического ядра. При условии наличия смысловой близости с тематическим маркером ключевое слово добавляется в кластер [5]

Для группировки ключевых слов в ручном режиме используют облачные и локальные табличные процессоры [табличные процессоры – это комплекс прикладных программ, предназначенных для обработки информации, проведения расчетов и визуального представления результатов расчетов в виде графиков и диаграмм], наиболее популярными среди которых являются: Google Sheets, Microsoft Excel, LibreOffice Calc и OpenOffice Calc [6, 7].

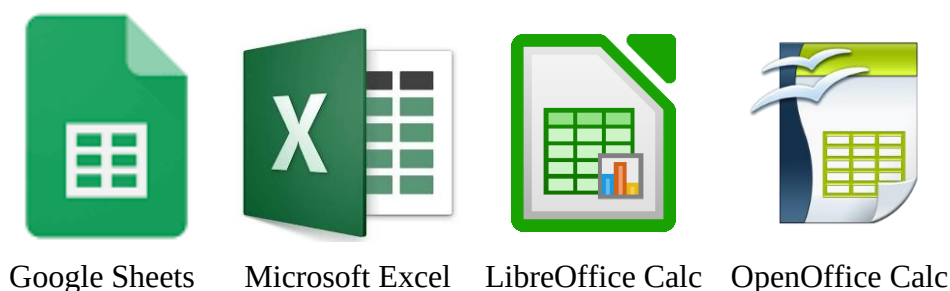


Рисунок 1. Облачные и локальные табличные процессоры.
Figure 1. Cloud and local spreadsheets.

В табличный процессор загружают список ключевых слов и сортируют по частотности, которую определяют с помощью разных сервисов, к примеру, Яндекс. Вордстата. Слова, которые обладают самой высокой частотностью, используют в качестве общих свойств, т.е. тематических маркеров, вокруг которых выстраивают средне – и низкочастотные ключевые слова с разделением по коммерческому и информационному интенту ключевых слов.

Существуют отдельные плагины и надстройки для табличных процессоров. К примеру, для Microsoft Excel был разработан плагин, который позволял обрабатывать разные по размеру семантические ядра. Кластеризация сводилась к следующему: в табличный процессор загружается семантическое ядро с указанием частотности ключевых слов, вычисляются доли весов, формируется справочная таблица с весами слов, выделяются леммы для слов, вычисляют вес лемм, формируют справочную таблицу с леммами, проводят первичную кластеризацию, укрупняют полученные группы.

На основе плагина создана надстройка Power Query, автоматизирующая механические действия плагина и сокращающая время обработки данных [8].

Иногда для логической кластеризации используют ментальные карты, к примеру, XMind или MindMeister.

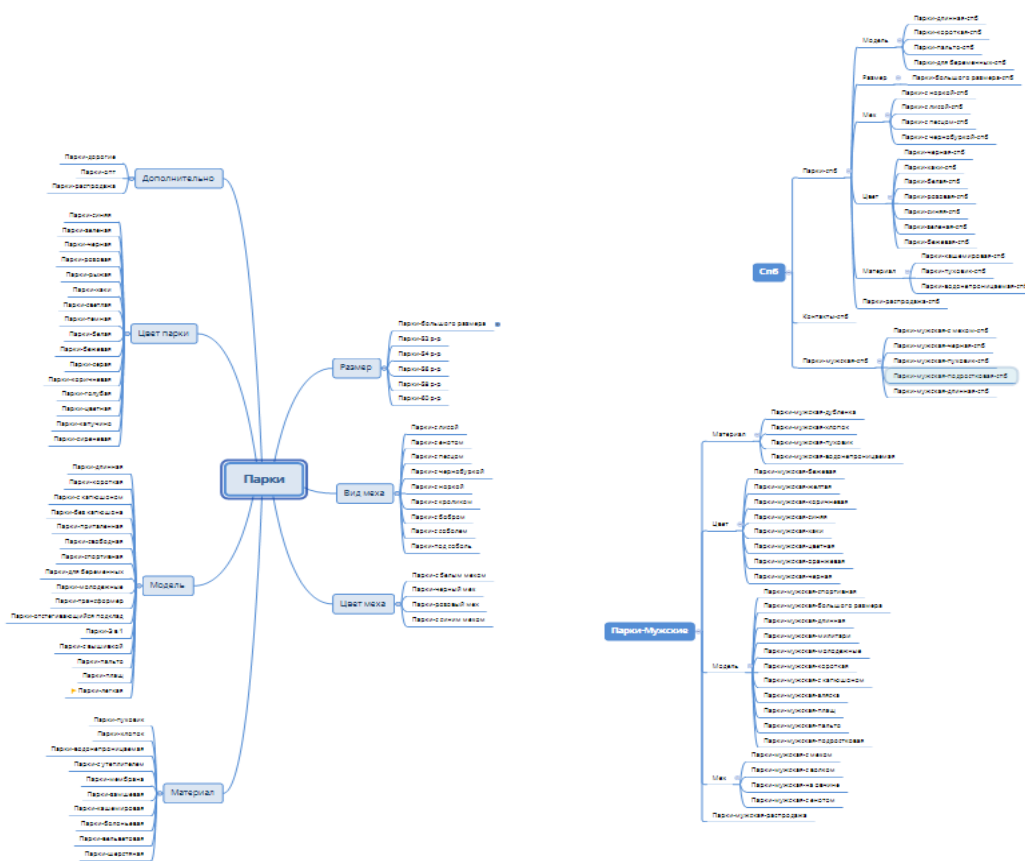


Рисунок 2. Пример кластеризации в виде ментальной карты.

Figure 2. An example of clustering in the form of a mental map.

Логическая группировка дает достаточно точный результат, но не рекомендуется использовать этот метод для кластеризации семантических ядер, которые содержат более несколько сотен слов, т.к. легко сбиться с принципа группировки, пропустить ключ или некорректно определить поисковой интент.

ГРУППИРОВКА ПО СЕМАНТИЧЕСКОЙ СХОЖЕСТИ

Метод группировки по семантической схожести применяют преимущественно для крупных коммерческих сайтов (агрегаторов, интернет-магазинов, сайтов услуг, маркетплейсов), семантические ядра которых включают в себя тысячи ключевых слов, т.к. метод основан на применении нейросетевых алгоритмов, что требует долгосрочной и дорогостоящей разработки формул и обучения нейросети для обработки и сегментации ключевых слов в автоматическом режиме.

Суть кластеризации заключается в разделении ключевых слов с учетом весовых коэффициентов на кластеры по смысловому совпадению. Выбирается тематический маркер с тяжелым весом, как и в случае с логической группировкой, маркер является одним из высокочастотных ключей из семантического ядра. При условии пересечения с маркером с тяжелым весом в кластер попадут семантические близкие ключи. В некоторых алгоритмах помимо требования семантической схожести может быть установлено дополнительное свойство для вхождения в кластер по морфологическому признаку.

Несмотря на то, что метод семантической схожести сложен в разработке, существует ряд упрощенных приложений, основанных на семантической кластеризации, к которым относятся Cluster Army, SpyFu, Simple SEO tool [9].

Следует учитывать, что автоматическая группировка дает достаточно точный результат, но кластеризация по семантической схожести не учитывает поисковый интент, поэтому определять интент ключей на коммерческую необходимость вручную или посредством сервиса определения процента коммерциализации.

ГРУППИРОВКА ПО ТОПАМ ПОИСКОВОЙ ВЫДАЧИ

Группировка по топам SERP поисковых систем может проводиться тремя способами: soft (мягкая), middle / moderate (умеренная) и hard (строгая). Мягкая кластеризация применяется в двух случаях: тематика сайта неконкурентная или для ресурса важно количество визитов, но не качество трафика. Строгая кластеризация используется для продвижения крупных коммерческих ресурсов и высококонкурентных тематик. Умеренную кластеризацию рекомендуют применять для продвижения коммерческих ресурсов в определенном регионе.

Суть кластеризации заключается в поиске пересечений ключевых слов на страницах сайтов, находящихся в топе поисковой выдачи. Поисковые системы формируют SERP на основании факторов ранжирования. Параметры, по которым оценивается сайт и определяется его позиция в органической выдаче, для каждой поисковой системы индивидуальны. Поэтому ключевые слова из семантического ядра поочередно проверяются и по каждому ключу составляется топ SERP в рамках конкретной поисковой системы. Результаты поиска по каждому из ключевых слов сравнивают с тематическим маркером на предмет одинаковых адресов страниц и в

зависимости от установленного порога пересечений ключевые слова объединяют в кластер.

При мягкой кластеризации ключевое слово попадет в кластер, если при сравнении с тематическим маркером количество общих адресов в топе SERP будет соответствовать установленному порогу пересечений. В кластере все ключевые слова связаны с маркером, но между собой по количеству общих адресов могут не пересекаться. При строгой кластеризации, как и при мягкой, ключевые слова сравниваются с маркером, но и дополнительно между собой. Для включения в группу ключи должны пересекаться не только с маркером по установленному порогу пересечений, но и между собой. При умеренной кластеризации ключевые слова связаны с маркером по установленному порогу пересечений, между собой обязательно связаны, но в разных парах количество общих адресов – разное [10].

К ряду основных инструментов кластеризации по топам SERP относятся онлайн – сервисы, среди которых наиболее популярными являются Serpstat, SERanking, Rush Analytics, SeoQuick, Topvisor и ПиксельТулс. Общий принцип работы данных сервисов заключается в загрузке в систему списка ключевых слов, выбора способа кластеризации, установке порога пересечений и выбора поисковой системы.

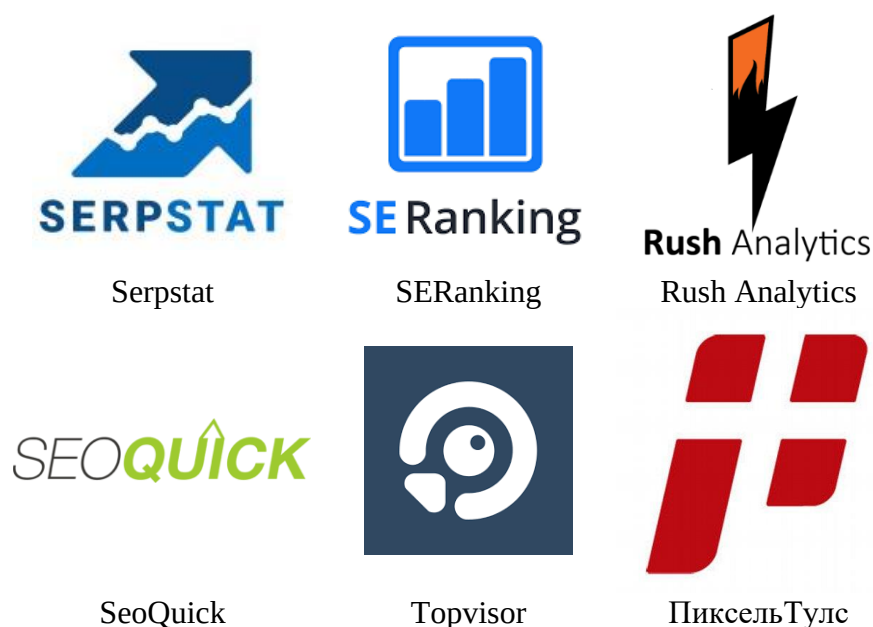


Рисунок 3. Инструменты кластеризации для SERP.

Figure 3. Clustering tools for SERP.

ЗАКЛЮЧЕНИЕ

Отметим, что для каждого метода кластеризации предусмотрены определенные инструменты. Но для успешного разделения ключевых слов на кластеры рекомендуется использовать следующую комбинацию: первоначально разбить ключи с помощью утилит на основе семантической схожести и определить интент, а затем соотнести результаты с результатами группировки по топам поисковой выдачи той системы, под которую будет оптимизирован ресурс. На основе корреляции результатов внести изменения в кластеры.

СПИСОК ЛИТЕРАТУРЫ

- [1] Прокимнов Н.Н. Технологии использования информационных ресурсов Интернета, Прикладная информатика, 5, 2008, 51-80.
- [2] Кравченко Ю.А. Задачи семантического поиска, классификации, структуризации и интеграции информации в контексте проблем управления знаниями, Известия Южного федерального университета. Технические науки, 7 (180), 2016, 5-18.
- [3] Чамина О.Г., Бексаева Е.А. SEO-аналитика веб-ресурсов электронной коммерции, Модели, системы, сети в экономике, технике, природе и обществе, 3 (15), 2015, 180-188.
- [4] Бохонок А., Синявский А. Кластеризация семантического ядра: что это, для чего нужно. URL: <https://elit-web.com/blog/klasterizaciya-semanticheskogo-yadra> (дата обращения: 01.08.2022).
- [5] Давиденко А. Полный гайд по кластеризации семантического ядра. URL: <https://kokoc.com/blog/klasterizaciya-zaprosov/> (дата обращения: 01.08.2022).
- [6] Миронов В. Инструменты для кластеризации запросов (семантического ядра). URL: <https://topodin.com/seo/post/instrumenty-dlya-klasterizatsii-zaprosov-semanticheskogo-yadra> (дата обращения: 01.08.2022).
- [7] Игошин Р. Excel-кластеризатор ключевых слов по весу. URL: <https://www.it-agency.ru/academy/excel-clusterer> (дата обращения: 01.08.2022).
- [8] Игошин Р. Кластеризатор ключевых слов на Power Query. URL: <https://www.it-agency.ru/academy/power-query-keywords-clusterizer/> (дата обращения: 01.08.2022).
- [9] Кластеризация семантического ядра: алгоритмы и подходы популярных SEO-инструментов. URL: <https://serpstat.com/ru/blog/kak-provesti-klasterizatciyu-zaprosov-s-promoshyu-serpstat/> (дата обращения: 01.08.2022).

[10] Ашманов И. С. Оптимизация и продвижение в поисковых системах. Санкт-Петербург: Питер, 2019. 512.

[11] Усачёв А.В., Яблокова А.А. Методы кластеризации в SEO проектировании: материалы II Всероссийской научной конференции «Наука, технологии, общество (НТО-II-2022)». 28-30 июля 2022, Красноярск; 2022: 233-238.

REFERENCES

[1] Prokimnov N.N. Tekhnologii ispol'zovaniya informacionnyh resursov Interneta, Prikladnaya informatika, 5, 2008, 51-80 (in Russian).

[2] Kravchenko YU.A. Zadachi semanticheskogo poiska, klassifikacii, strukturizacii i integracii informacii v kontekste problem upravleniya znaniyami, Izvestiya YUzhnogo federal'nogo universiteta. Tekhnicheskie nauki, 7 (180), 2016, 5-18 (in Russian).

[3] CHamina O.G., Beksaeva E.A. SEO-analitika veb-resursov elektronnoj kommercii, Modeli, sistemy, seti v ekonomike, tekhnike, prirode i obshchestve, 3 (15), 2015, 180-188 (in Russian).

[4] Bohonok A., Sinyavskij A. Klasterizaciya semanticheskogo yadra: chto eto, dlya chego nuzhno. URL: <https://elit-web.com/blog/klasterizaciya-semanticheskogo-yadra> (data obrashcheniya: 01.08.2022) (in Russian).

[5] Davidenko A. Polnyj gajd po klasterizacii semanticheskogo yadra. URL: <https://kokoc.com/blog/klasterizaciya-zaprosov/> (data obrashcheniya: 01.08.2022) (in Russian).

[6] Mironov V. Instrumenty dlya klasterizacii zaprosov (semanticheskogo yadra). URL: <https://topodin.com/seo/post/instrumenty-dlya-klasterizatsii-zaprosov-semanticheskogo-yadra> (data obrashcheniya: 01.08.2022) (in Russian).

[7] Igoshin R. Excel-klasterizator klyuchevyh slov po vesu. URL: <https://www.it-agency.ru/academy/excel-clusterer> (data obrashcheniya: 01.08.2022) (in Russian).

[8] Igoshin R. Klasterizator klyuchevyh slov na Power Query. URL: <https://www.it-agency.ru/academy/power-query-keywords-clusterizer/> (data obrashcheniya: 01.08.2022) (in Russian).

[9] Klasterizaciya semanticheskogo yadra: algoritmy i podhody populyarnyh SEO-instrumentov. URL: <https://serpstat.com/ru/blog/kak-provesti-klasterizatsiyu-zaprosov-s-pomoshyu-serpstat/> (data obrashcheniya: 01.08.2022) (in Russian).

[10] Ashmanov I. S. Optimizaciya i prodvizhenie v poiskovyh sistemah. Sankt-Peterburg:

Piter, 2019. 512 (in Russian).

[11] Usachyov A.V., Yablokova A.A. Metody klasterizatsii v SEO proektirovanii: materialy II Vserossiyskoj nauchnoj konferencii «Nauka, tekhnologii, obshchestvo (NTO-II-2022). 28-30 iyulya 2022, Krasnoyarsk; 2022: 233-238 (in Russian).

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Алена Александровна Яблокова,
аспирант, Красноярский краевой Дом науки
и техники Российского союза научных и
инженерных общественных объединений,
г. Красноярск, Россия
e-mail: alena.yabl@yandex.ru

Alena A. Yablokova
Post-graduate, Krasnoyarsk Regional Science
and Technology City Hall of the Russian Union
of Scientific and Engineering Public
Associations, Krasnoyarsk, Russia
e-mail: alena.yabl@yandex.ru

Статья поступила в редакцию 22.07.2022; одобрена после рецензирования 18.08.2022; принята к публикации 23.08.2022.

The article was submitted 22.07.2022; approved after reviewing 18.08.2022; accepted for publication 23.08.2022.