

# Отбор тем в тематических моделях для разведочного информационного поиска.\*

Хоменко Р.Д.<sup>1</sup>, Воронцов К.В.<sup>1</sup>, А. О. Янина<sup>1</sup>

r.d.khomenko@yandex.ru; vokov@forecsys.ru; yanina.anastasia.mipt@gmail.com

<sup>1</sup> Московский физико-технический институт, Москва, Россия

Одним из приложений тематического моделирования является разведочный информационный поиск. Данный метод поиска позволяет пользователю получить набор релевантных документов, дорожную карту исследуемого раздела знаний, а также научиться ей пользоваться.

В данной работе исследуются методы отбора тем для повышения качества разведочного информационного поиска. Используются следующие методы отбора тем: метод рекурсивного исключения признаков, использование фоновых и предметных тем в тематических моделях, отбор тем по критерию чистоты и контрастности лексического ядра.

Тут должно быть краткое описание результатов работы.

**Ключевые слова:** *разведочный информационный поиск; тематическое моделирование; аддитивная регуляризация; методы отбора тем; рекурсивное извлечение признаков*

## 1 Введение

Тут будет введение и ссылки на литературу [5], [2], [1], [3], [4].

## 2 Постановка задачи

Тут будет постановка задачи.

### 2.1 Название параграфа

Разделы и параграфы, за исключением списков литературы, нумеруются.

## 3 Заключение

Желательно, чтобы этот раздел был, причём он не должен дословно повторять аннотацию. Обычно здесь отмечают, каких результатов удалось добиться, какие проблемы остались открытыми.

## Литература

- [1] Anastasia Ianina, Lev Golitsyn, and Konstantin Vorontsov. Multi-objective topic modeling for exploratory search in tech news. pages 181–193, 01 2018.
- [2] K. Vorontsov, O. Frei, M. Apishev, P. Romov, M. Suvorova, and A. Yanina. Non-bayesian additive regularization for multimodal topic modeling of large collections. 2015.
- [3] K. Vorontsov, A. Potapenko, and P. Alexander. Additive regularization of topic models for topic selection and sparse factorization. 2015.
- [4] K. V. Vorontsov. Additive regularization for topic models of text collections. 2013.
- [5] Воронцов К.В. Вероятностное тематическое моделирование: теория, модели и проект bigartm. 2020.

*Поступила в редакцию*

---

\*Задачу поставил: Воронцов К. В. Консультант: Янина А. О.