

Практическая работа 2. Анализ частотности терминов в книжной продукции

Помимо поиска в интернет публикациях, на сегодняшний день возможно осуществление анализа частотности употребления отдельных слов и фраз, названий, имен и фамилий известных людей и т.п. в книжных коллекциях на основных языках мира. Для этого можно воспользоваться сервисом Ngrams (<https://books.google.com/ngrams>), работающим на материалах проекта Google Books, на сегодняшний день аккумулировавшего более четверти из когда-либо изданных человечеством книг.

Google Books Ngram Viewer

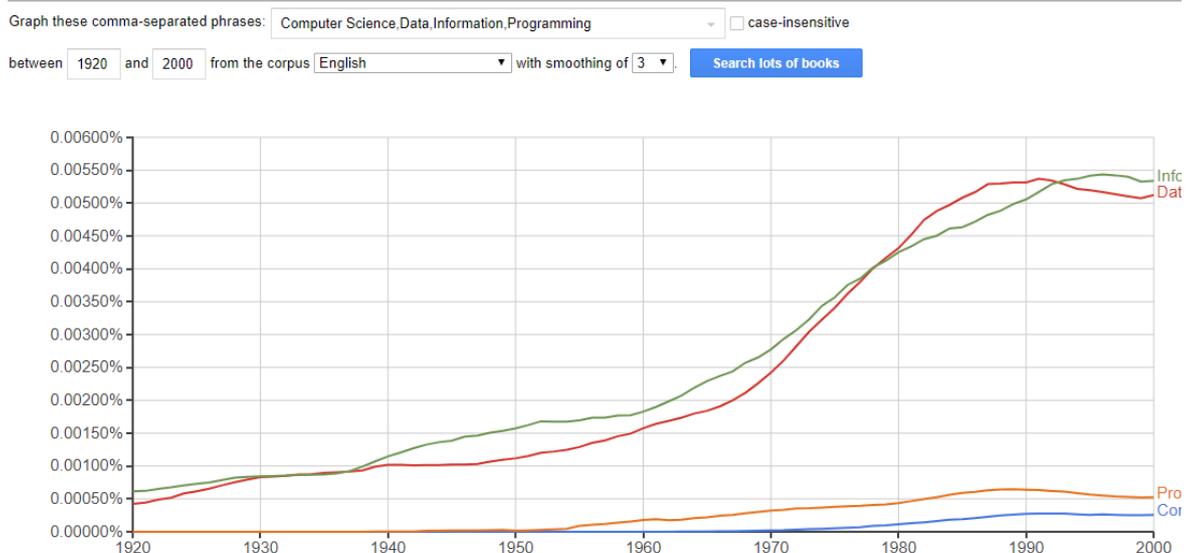


Рис. 1. Динамика популярности терминов Computer Science, Data, Information, Programming

Данные графика показывают, что за последние 100 лет трижды происходила смена приоритетов в использовании терминов Data и Information, два других термина - Computer Science и Programming, являются более узкоспециализированными, поэтому обладают меньшей относительной популярностью среди авторов книг и частоты их упоминания колеблется менее резко.

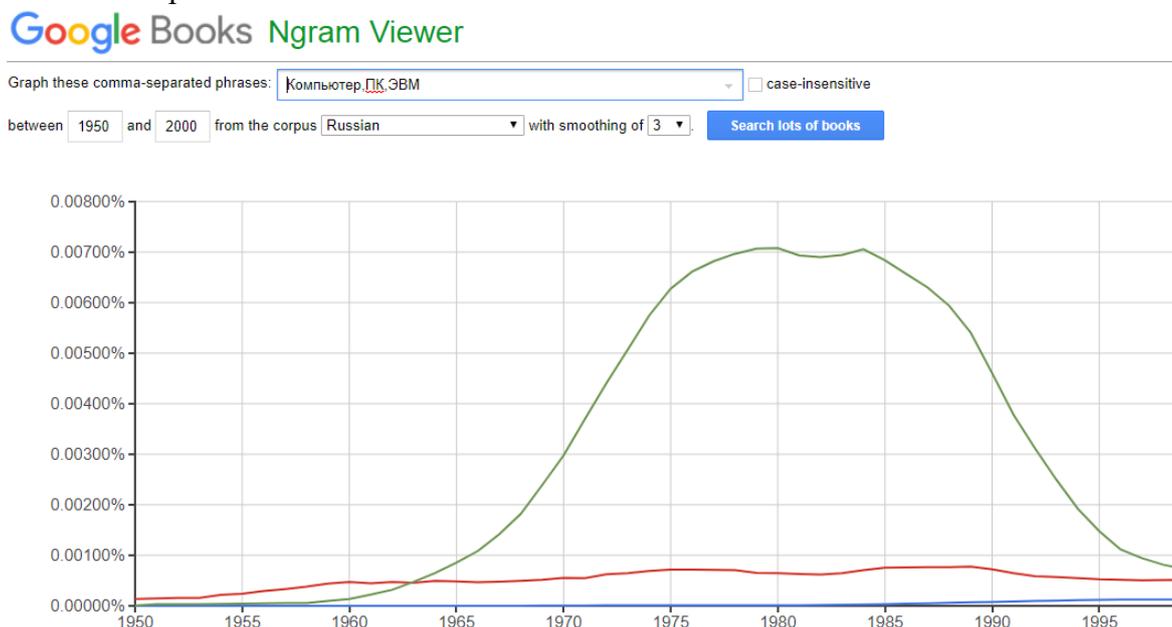


Рис.2. Динамика популярности названий вычислительной техники на русском языке.

Такого рода исследования можно выполнять не только на английском языке, но и, например, на русском. Для качественного выполнения данной работы следует поэкспериментировать с несколькими (минимум 5-7 группами) совокупностями синонимов или антонимов, обозначающих один и тот же объект, предмет, процесс и т.п. в различных областях деятельности. Также можно произвести поиск ФИО ученых, либо общественных деятелей, внесших значимый вклад в развитие какой-либо отрасли науки, культуры и т.д.

Решение:

Данные рисунка 1 показывают, что популярность слова- маркетплейс начала появляться после 2013 года, график резко пошел вверх и к 2019 году находится на самых высоких значениях за анализируемый период.

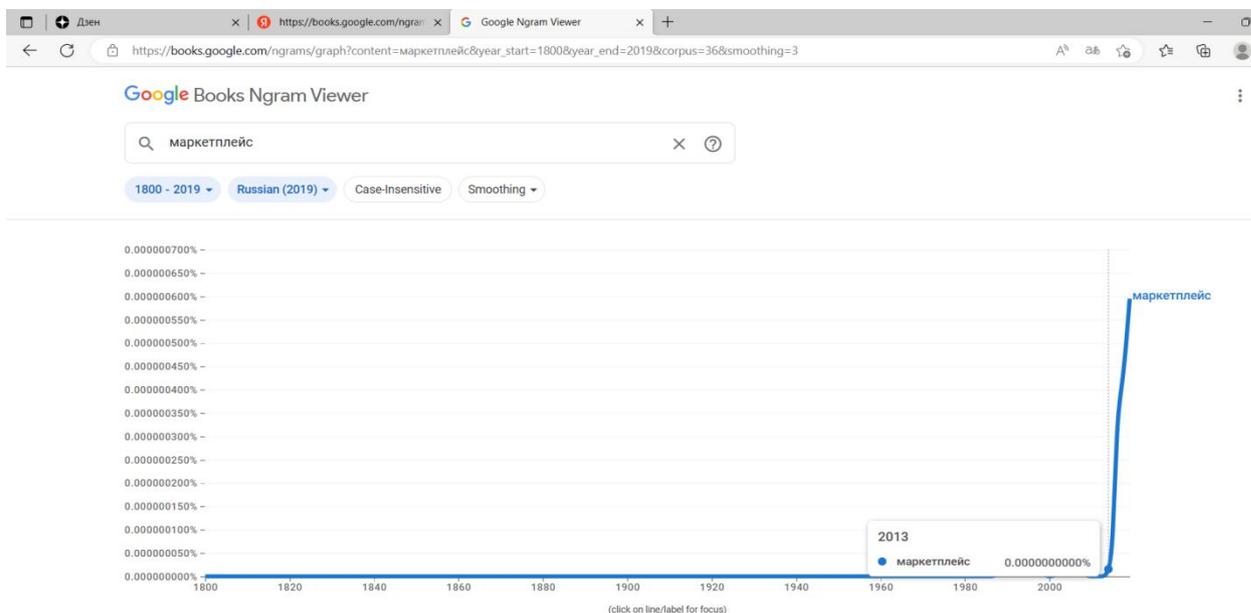
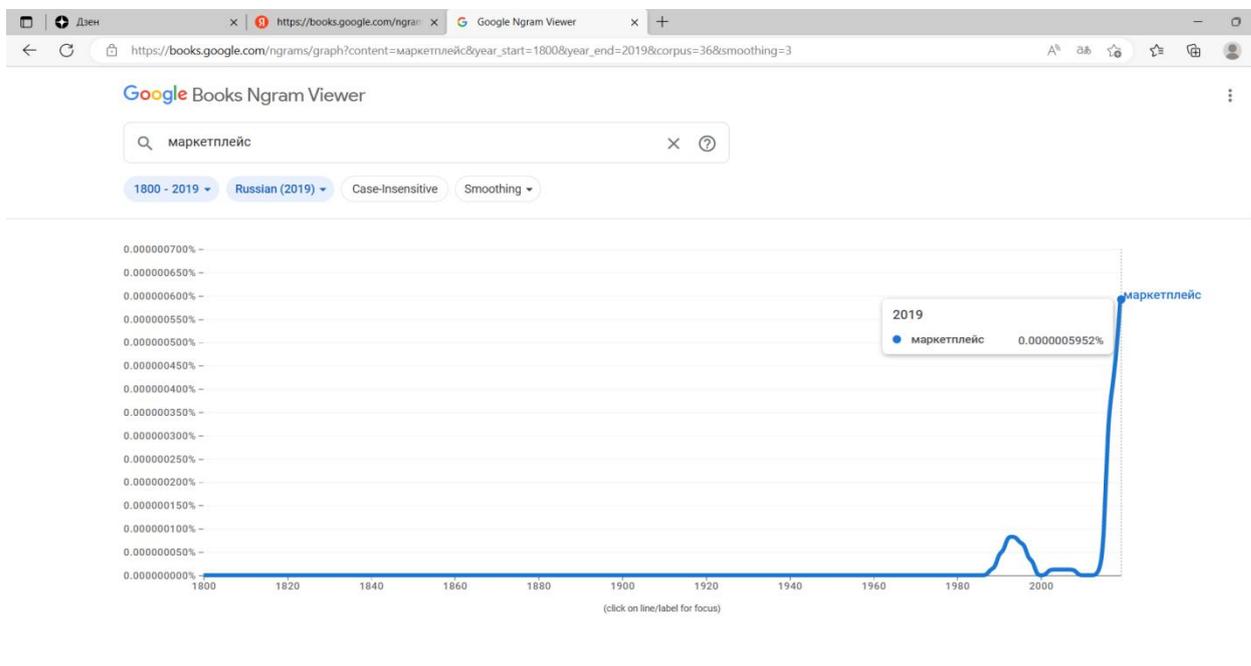


Рисунок 1-Динамика популярности термина «маркетплейс» с 1800-2019гг.

По запросу –интернет-магазин наблюдается начало роста после 1998года, далее резкий рост к 2012 году, был достигнут максимальный показатель, далее началось резкое снижение количества напоминаний по данному слову. К 2019 году количество напоминаний упало практически вдвое по сравнению с показателем 2012года. Очевидно, что после 2012 года на тему интернет-магазин стали меньше писать книг, рисунок 2.

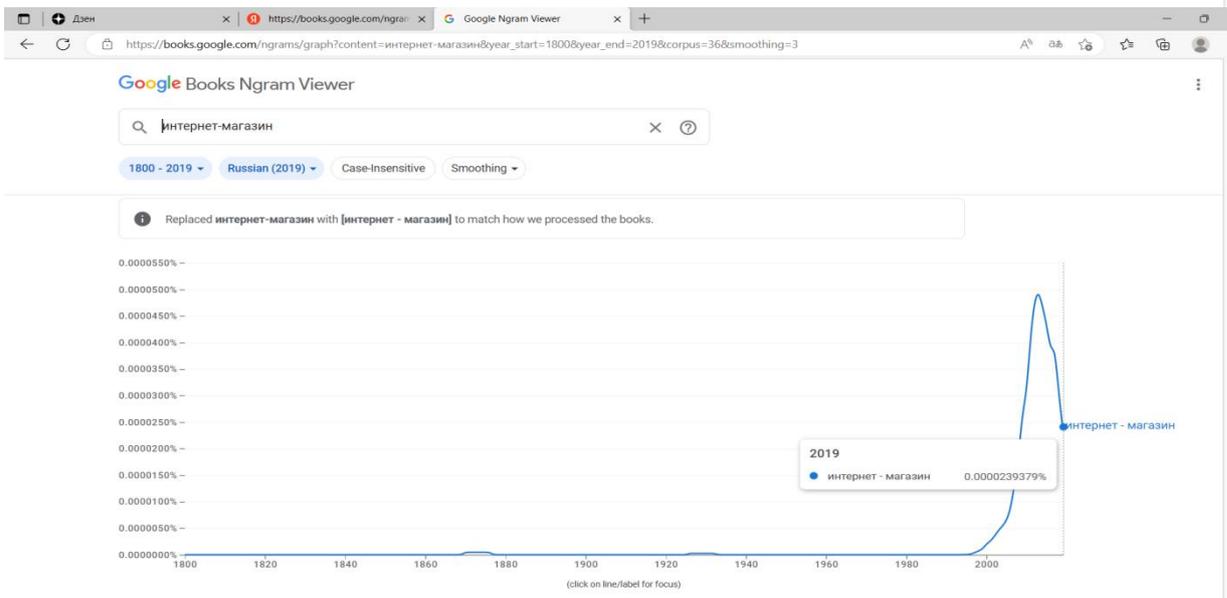
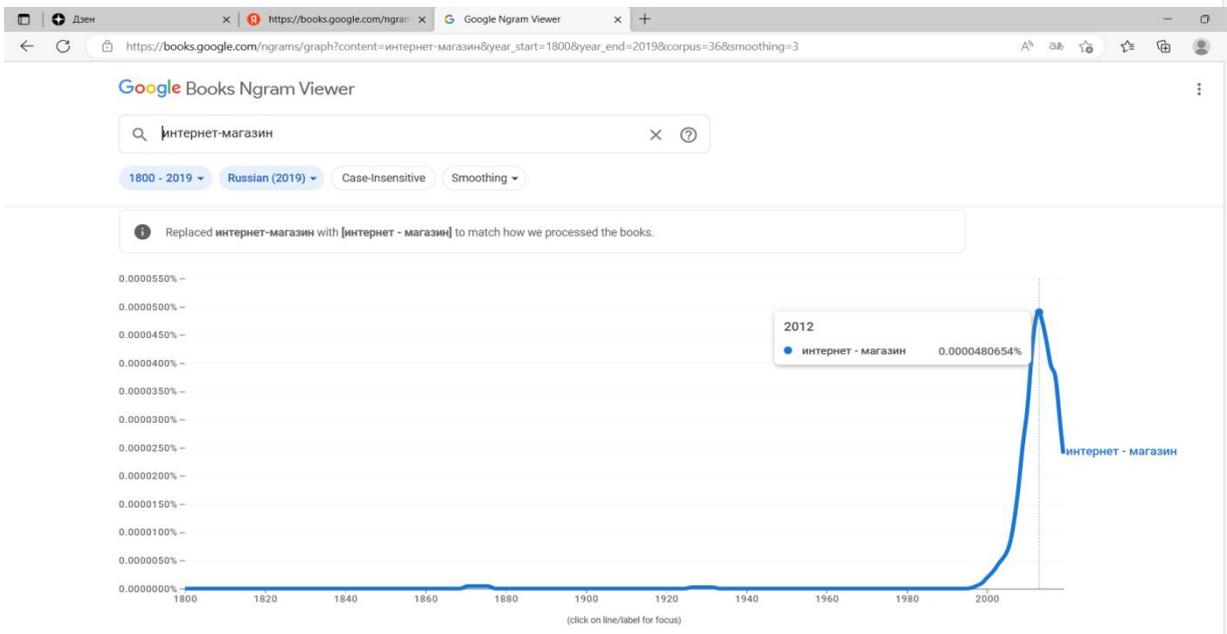
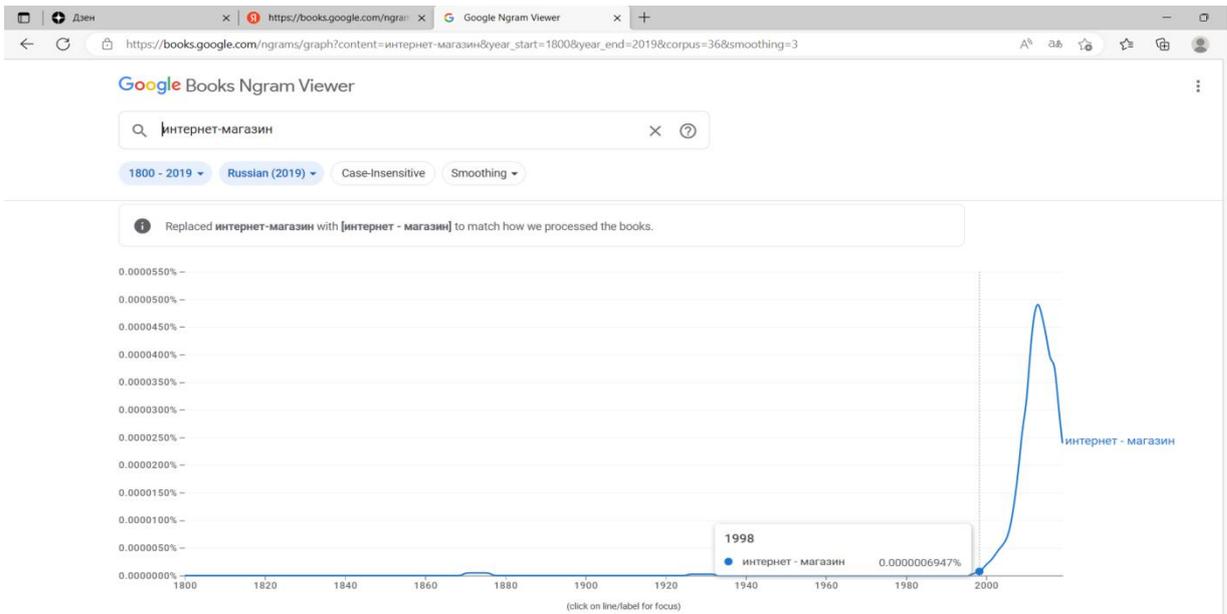


Рисунок 2-Динамика популярности термина «интернет-магазин» с 1800-2019гг.

По слову- интернет-сайт- популярность запроса начала расти в 1997году, далее график резко пошел вверх до 2007 года, в этот момент были достигнуты самые максимальные показатели по данному запросу. После 2007 года график стал снижаться и к 2019 году количество запросов снизилось практически в два раза по отношению к уровню 2007года. То есть после 2007года количество написанных книг, а следовательно и напоминаний на тему: интернет-сайт- стало снижаться , рисунок 3.

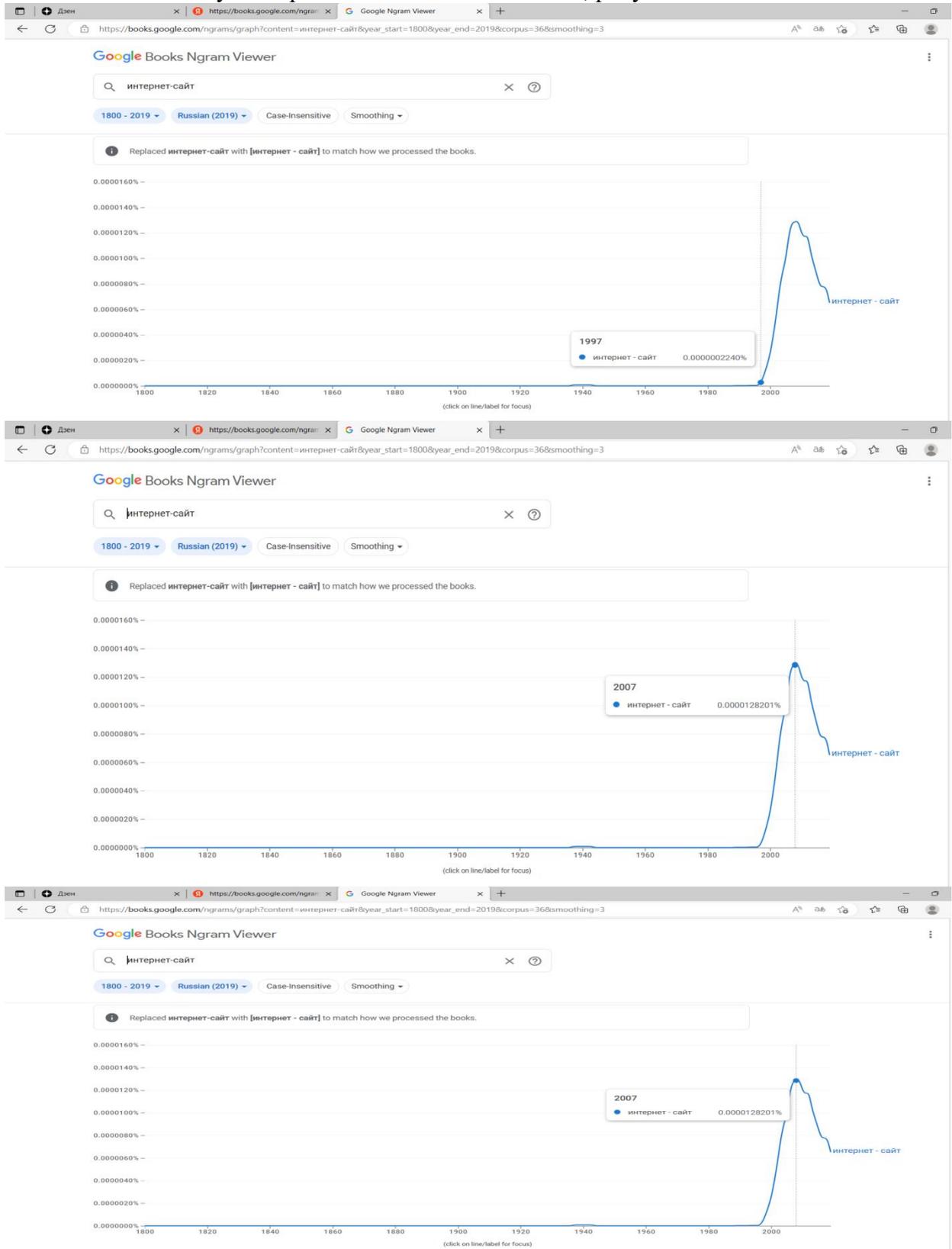


Рисунок 3-Динамика популярности термина «интернет-сайт» с 1800-2019гг.

Пик популярности термина «электронная коммерция» был достигнут в 2017 году, рост популярности начался с 1994 года, достиг максимального значения в 2003 году, потом популярность начала снижаться до 2009 года, после график развернулся и опять пошел вверх. Достигнув максимального значения в 2017 году, а к 2019 году замечено снижение популярности термина «электронная коммерция», рисунок 4.

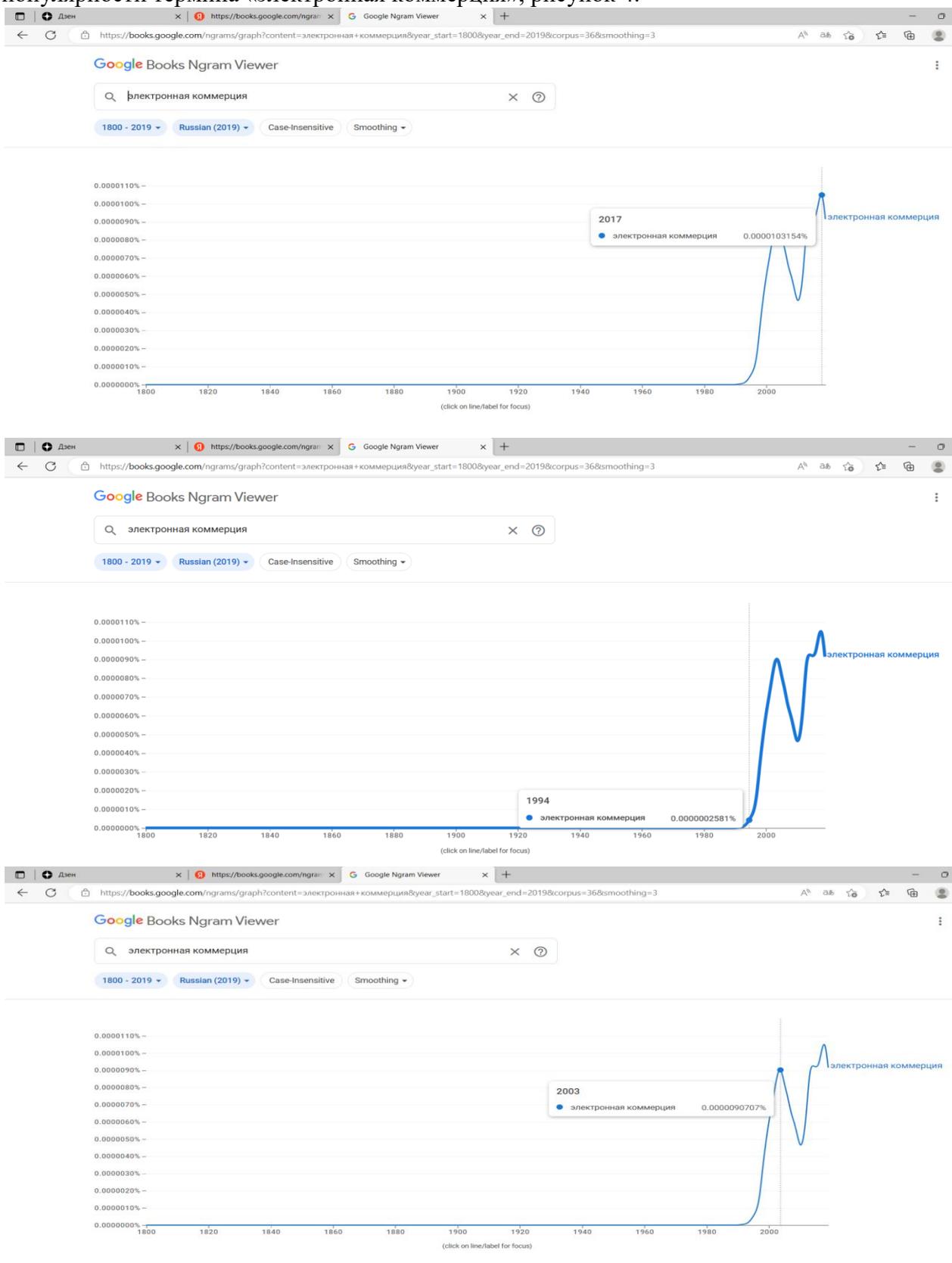


Рисунок 4- Динамика популярности термина «Электронная коммерция» с 1800-2019гг.

Термин- «Электронный бизнес» начал получать популярность после 1980 года, с 1994 года популярность термина резко пошла вверх и достигла пикового значения в 2003 году. Далее до 2009года популярность этого термина резко снижалась, потом начался плавный рост, но к 2019году термин так и не достиг максимальных значений. Очевидно, что количество книг на тему электронного бизнеса в настоящее время стало меньше, прошел пик популярности этого термина, рисунок 5.

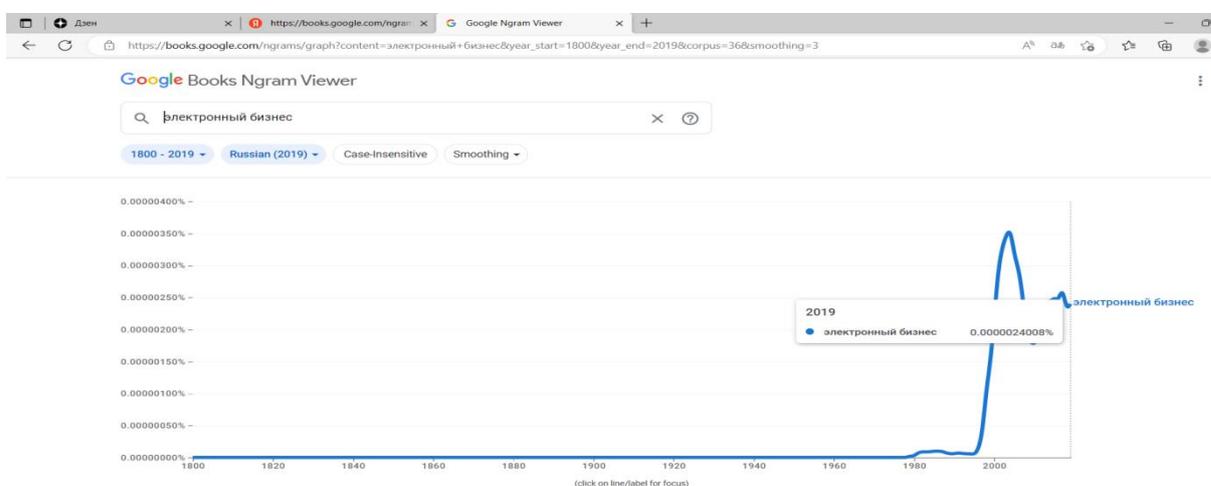
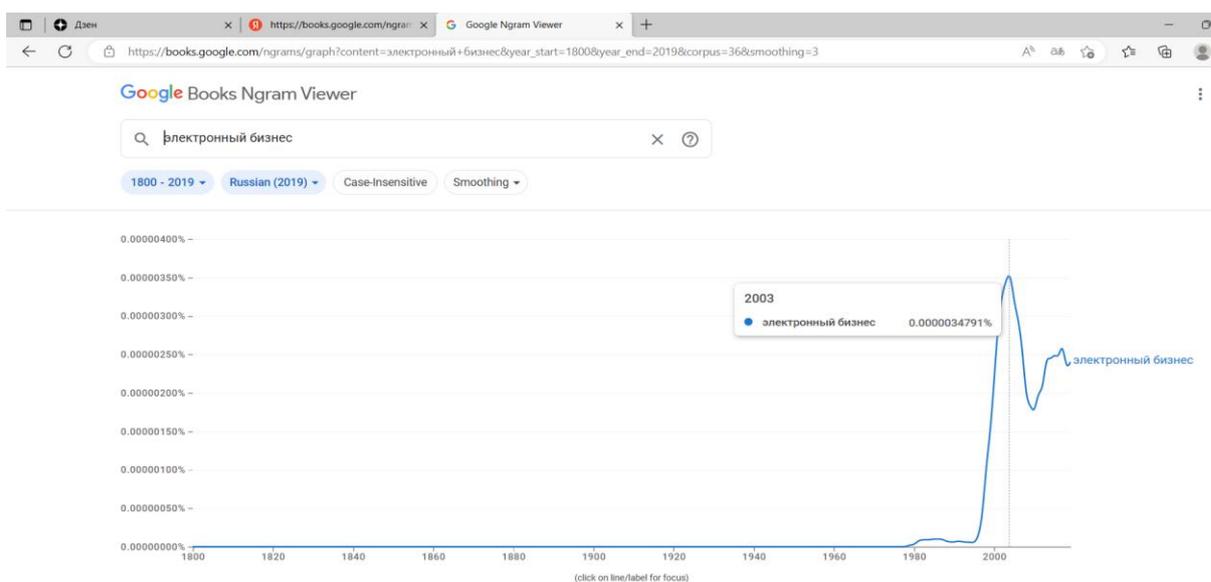
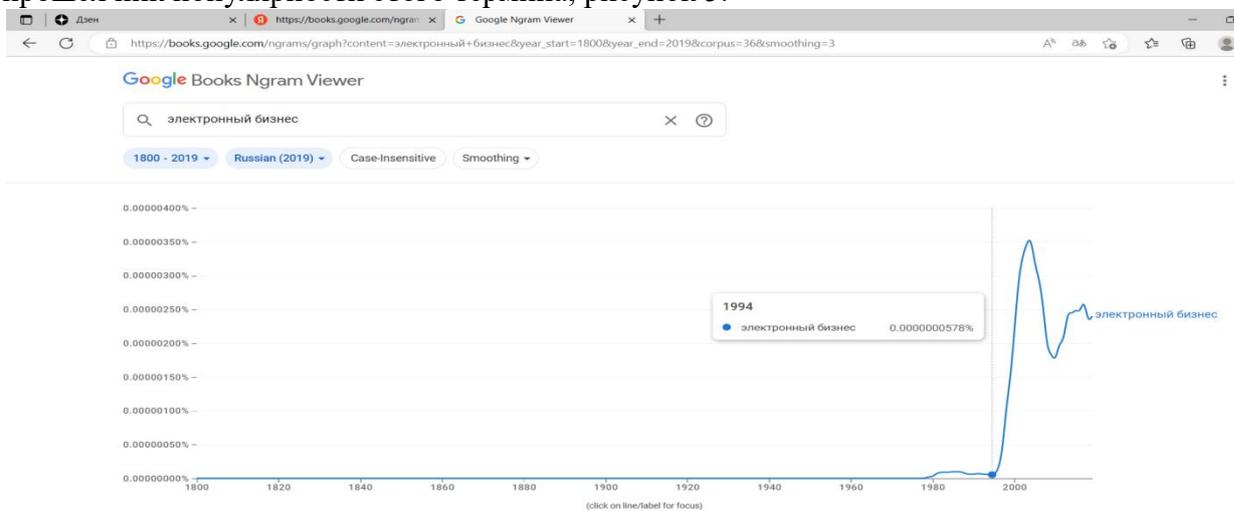


Рисунок 5- Динамика популярности термина «Электронный бизнес» с 1800-2019гг.

Таким образом, анализ частотности употребления терминов в книжной продукции на ресурсе (<https://books.google.com/ngrams>) показал, что термины связанные с понятием электронной коммерции, электронного бизнеса приобрели большую популярность в период с 2003-2017гг., к настоящему времени эти термины снизились относительно максимальных значений. То есть количество упоминаний этих слов в книжных коллекциях снизилось.