Министерство науки и высшего образования РФ ФГБОУ ВО «Удмуртский государственный университет» Институт математики, информационных технологий и физики

Отчет к исследовательской работе

«Методы исследования документальных информационных потоков»

Выполнил: студент 4 курса группы ОАБ-02.03.01-42 Лаврентьев

Введение

Основной причиной роста внимания к методам исследования документальных информационных потоков является их объективность в оценке динамично развивающегося научного потенциала страны, выбора научных приоритетов, исследования закономерностей функционирования научных систем, изучения вклада и продуктивности отдельного ученого/коллектива в развитие науки и прогнозирования дальнейших исследований.

Основные направления изучения информационных потоков

Документальные информационные потоки изучаются такими активно развивающимися направлениями, как библиометрия, информетрия, наукометрия, вебометрия, киберметрия, медиаметрия и др., позволяющими анализировать закономерности развития документопотоков.

Библиометрия - применение математических и статистических методов к изучению книг, периодических изданий и прочих публикаций.

Информетрия - это направление, связанное с исследованиями всех количественных аспектов информации, информационных процессов и явлений.

Наукометрия - область науковедения, проводящая исследова-ние науки количественными методами; научная дисциплина, изучающая эволюцию науки через многочисленные измерения и статистическую обработку научной информации (количество научных статей, опубликованных в данный период времени, цитируемость и т. д.)

Вебометрия - раздел информатики, в рамках которого исследуются количественные аспекты конструирования и использования информационных ресурсов, структур и технологий применительно к Всемирной паутине.

Киберметрия - научная дисциплина, занимающаяся количественным анализом новых возможностей по обработке информации, хранимой в электронном виде, и её визуализации

Медиаметрия - практическая деятельность по изучению и сбору численных и качественных данных медиаканала и аудитории, проводимое с целью исследования эффективности средств массовой информации.

Соответственно перечень методов носит название "библиометрических", "информетрических", "информационных". Различие этих методов, по мнению В.И. Горьковой, в несходстве предметов и сфер применения, объективных количественных закономерностей, а также целей практической деятельности. Термины являются схожими по значению. В то же время они используются для описания различных подходов и, как отмечают W.W. Hood и C.S. Wilson: "У каждого своя история возникновения и сфера использования. При этом они не ограничиваются рамками

только научной продукции и коммуникаций, а могут применяться в бизнесе, политике и других областях."

История изучения

количественные исследования Изначально документопотока называли 'статистической библиографией'. Термин ввел в 1923 г. Е. Хамл, применив его к ранжированию стран по числу журнальных статей в определенных областях. Затем был введен термин "библиометрия", который, по мнению Ю. Гарфилда, впервые встречается в работах П. Отле (1934) и А. Причарда (1969). Концептуальные положения о библиометрии, ее месте в системе наук и используемых методах остаются, несмотря на многочисленность исследований, весьма актуальными проблемами и сегодня. Зарубежные исследователи чаще всего указывают на работу Ф. Коула и Н. Ильса 1917 г., в которой был проведен статистический анализ литературы по сравнительной анатомии. О. Воверене, вслед за историком науки С.В. Альтшуллером отмечает, что несправедливо забыто имя русского ученого П. Вальдена, который еще в 1911 г. впервые применил метод анализа цитирования для изучения вклада ученых отдельных стран в развитие химии и, таким образом, должен быть признан родоначальником данного метода - наиболее часто используемого в настоящее время при оценке научной результативности. А первые попытки количественного изучения потока отечественной литературы были сделаны А. Шторхом и Ф. Аделунгом. Ими был проанализирован собственный указатель, отражающий литературу 1801- 1806 гг. по разным параметрам.

В 1969 г. В.В. Налимов и З.М. Мульченко ввели термин "наукометрия", в дальнейшем получивший распространение как "scientometrics". Он относится к области науковедения, изучающей закономерности функционирования и развития науки, структуру, динамику научной деятельности, взаимодействие науки с другими сферами материальной и духовной жизни общества, т.е. исследует процессы, явления научной деятельности в более широком контексте.

В 1979 г. в статьях немецких авторов L. Blackert, S. Siegel и О. Nacke появился термин "информетрия", который первоначально был определен как комплекс математических методов для исследования объектов информационной науки, описания и анализа их свойств, а также закономерностей в целях оптимизации этих объектов при принятии решений. Десятью годами позже под информетрией стали понимать использование разнообразного математического аппарата для анализа, выявления закономерностей, формулировки законов информационной деятельности и научной информации, а также для принятия решений в информационной практике. Специалист Института научной и технической документации и информации (Гавана, Куба) Мельвин Моралес выделяет следующие характерные аспекты информетрии:

- 1. Количественный рост литературы.
- 2. Старение и рассеяние информации.
- 3. Эффективность информационных продуктов и услуг в сфере производства, науки и техники.
- 4. Эффективность информационной системы и информационных органов в их совокупности.
 - 5. Роль различных видов документов как средств научной коммуникации.
 - 6. Роль неформальных каналов научной коммуникации.
 - 7. Релевантность и пертинентность информации.
 - 8. Ранжирование периодических и продолжающихся изданий по различным аспектам.
 - 9. Тематическая близость периодических и продолжающихся изданий.
 - 10. Значение практики цитирования между учеными и ее развитие.
- 11. Внутридисциплинарные и междисциплинарные связи на основе библиографических ссылок.

В основе вновь возникших родственных терминов, таких как "киберметрия", "интернетометрия" и "вебометрия", также находятся библиометрические приемы, только связаны они с исследованием новых форм представления информации. Так, под киберметрией понимают анализ потоков киберинформации (всех видов медиаинформации) с использованием наукометрических, библиометрических и информационных подходов. Издается журнал с одноименным названием — "Cybermetrics: International Journal of Scientometrics, Informetrics and Bibliometrics" (ISSN 1137-5019), отражающий развитие электронных коммуникаций и включающий анализ результатов научной деятельности, проведенных в сети World Wide Web (WWW).

Термин "вебометрия" впервые был применен в 1996 г. R.H. Abraham и R.R. Larson в целях построения когнитивных карт и математических моделей WWW. Первоначально предложенная технология подразумевала "bibliometrics on the Web", т.е. использование библиометрии при анализе Web. На сегодняшний день существует, как минимум, два его толкования. Первое - вебометрия - это статистика, цитирование веб-сайтов. Вторая позиция, которую поддерживают Lennart Bjorneborn и Peter Ingwersen, заключается в более углубленной теоретической проработке данной дефиниции. Авторы определяют вебометрию как область, изучающую с помощью библиометрических методов характер и свойства сети Web. Некоторые исследователи предлагают под вебометрией понимать методологию исследования World Wide Web.

Сущность формализованных методов анализа ДИП заключается в подсчете, сочетании, интерпретации, сравнении некоторых элементов документопотока, воздействующих друг на друга в процессе развития научного направления.

Методы количественного анализа, базирующиеся на исследовании вторичных источников информации

Основными параметрами количественного анализа вторичного документопотока, в частности метода анализа реферативных журналов, предложенного в 1981 г. А. Причардом и Г. Виттингом, являются:

- объем документопотока
- динамика и тенденции его формирования
- закономерности концентрации и рассеяния
- изменение во времени по отраслям науки и научным направлениям
- изменение во времени по странам и коллективам.

При этом анализ динамики развития документопотока, рост, старение, ранговое распределение основываются на классических законах и моделях, рассмотренных ниже.

Закон обратного квадрата А. Лотки

Важное эмпирическое исследование массива публикаций в целях анализа научной продуктивности, принадлежащее А. Лотке, появилось в 1926 г.. Автор подсчитал число ученых, написавших одну, две и т.д. статьи, на массиве публикаций, отраженных в издании "Chemical Abstracts" за 1907-1916 гг., а также проанализировал справочник открытий по физике за 1600-1900 гг. Ф. Ауэрбаха "Geschichtstafeln der Physik" (1910). В обоих случаях для числа n(x) ученых, написавших х статей, было получено следующее распределение научной продуктивности:

n(x)=AxI где A - число ученых, написавших всего одну статью; x = 1, 2... хмах(максимальная продуктивность ученого).

Данное распределение носит название Закона Лотки или Закона обратных квадратов: число ученых, написавших данное число статей, обратно пропорционально квадрату этого числа статей. Закон многократно проверялся на различных информационных массивах и оказался справедливым для всего диапазона научной продуктивности, кроме самых низких (n = 1) и самых больших значений ("краевые эффекты"). Таким образом, А. Лотка доказал, что ученые отличаются по числу опубликованных статей, т.е. распределение научной продуктивности носит асимметрический, неравномерный характер. Инвариантность, устойчивость этого закона во множестве научных областей позволяет говорить о нем как об одном из важнейших библиометрических законов.

Закон распределения Дж. Ципфа

Дж. Ципф, собрав огромный статистический материал, открыл закон распределения слов естественного языка. Он установил, что если к какому-либо достаточно большому тексту составить список всех встретившихся в нем слов, затем расположить их в порядке убывания частоты их встречаемости и пронумеровать от 1 до R, то для любого слова произведение его порядкового номера (R) на эту частоту будет величиной постоянной, имеющей примерно одинаковое численное значение для слов из рассматриваемого списка. Выражается закон Ципфа следующим образом: fr=c

где f - частота встречаемости слова в тексте; r - ранг (порядковый номер) слова в списке; с - эмпирически постоянная величина.

Полученная подчиненность графически выражается гиперболой. Исследовав таким образом самые разнообразные языки, Дж. Ципф для каждого построил указанные зависимости, при этом все кривые имели одинаковую форму - "гиперболической лестницы", т.е. при замене одного текста другим общий характер распределения не менялся. Ранговому гиперболическому закону Ципфа подчиняется распределение не только слов, но и различных явлений социального характера, в том числе распределение ученых (по числу опубликованных ими работ). Ранговый подход позволяет достаточно точно проранжировать распределение ученых по уменьшению продуктивности.

Закон рассеяния С.К. Брэдфорда

В дальнейшем С. Брэдфорд сформулировал закономерность распределения публикаций по изданиям. Согласно открытому им закону периодические издания можно разделить на группы, равные по общему числу статей в каждой (т.е. на совокупности журналов с одинаковым суммарным числом статей по данной тематике). Первая, наиболее "продуктивная", зона (ядро массива - P1) содержит малое число специальных журналов; вторая, более обширная, содержит статьи в научных журналах по смежным специальностям(P2) и, наконец, последняя зона ("периферия" - P3) включает наиболее обширное число журналов по разным областям, в которых рассеяны немногочисленные статьи по данной тематике.

P1:P2:P3=1:a:a^2

Если научное направление достаточно развито, то журналы по соответствующей проблеме обычно распадаются на три группы, каждая из которых содержит по одной трети публикаций. В каждой группе должно быть одинаковое количество публикаций, но число периодических изданий в них - различно. С. Брэдфордом была построена кривая рассеяния публикаций, в которой виден начальный участок, т.е. зона "ядерных" периодических изданий, а далее - прямолинейный участок. Выполнимость данного закона считается косвенным признаком "классичности" области исследования и используется для оценки развития и самостоятельности научного направления

Закон Б. Викери

Б. Викери доработал модель С. Брэдфорда. Он определил, что журналы, выстроенные в порядке уменьшения в них числа статей по конкретной проблеме, можно разбить не на три зоны, а на нужное количество зон. Закономерность имеет следующее выражение: Tx:T2x:T3x:T4x...=1:a:a^2:a^3...,

где x - количество статей в каждой группе; а - коэффициент рассеяния, постоянная величина для данного предмета и времени; Тx - количество журналов, содержащих x статей по данному предмету; Т2x - количество журналов, содержащих 2x статей, т.е. сумма наименований журналов первой и второй групп; Т3x - количество журналов, содержащих 3x статей; Т4x - количество журналов, содержащих 4x статей и т.д.

Закон Брукса

Формула Брукса, предложенная в 1969 г., также уточняет закон Брэдфорда и состоит из двух уравнений:

$$R_{(n)} = \operatorname{an^b}(1 \le n \le c)$$

$$R_{(n)} = k \log_b \frac{n}{s} (c \le n \le N),$$

где R(n) - суммарное число публикаций по данной отрасли, содержащееся в п периодических изданиях, проранжированных в порядке уменьшения "продуктивности"; n - ранг (порядковый номер) периодического издания в списке по убывающей продуктивности; а - количество релевантных публикаций в самом продуктивном журнале; с - число периодических изданий в "ядре"; N - общее количество периодических изданий в ранговом списке; b, k, s - эмпирические коэффициенты.

Первое уравнение описывает начальный участок кривой, который имеет экспоненциальный характер, а второе - прямолинейную часть кривой Брэдфорда. В зависимости от соотношения коэффициентов делают заключение о состоянии научного направления. Установлено, что для узкоспециализированных областей науки $s \le 1$, а величина $c \ge 3$ и связана $c \le 3$ прямой зависимостью.

Б. Брукс объясняет механизм функционирования Закона рассеяния С. Брэдфорд, предлагая следующую гипотезу. Первые статьи по новому предмету исследования публикуют в нескольких наиболее соответствующих тематике периодических изданиях. В результате данные журналы помещают, по мере развития рассматриваемого предмета, все больше подобных статей. В то же время статьи по этому предмету начинают публиковаться и в других периодических изданиях. Если предмет продолжает развиваться, то со временем складывается брэдфордовское "ядро" периодических изданий, в которых помещается наибольшее число публикаций по данной теме.

Истинность вышеперечисленных законов подтверждается огромным количеством последующих исследований. Вместе с тем обобщение практического опыта и эксперименты показали, что, например, Закон рассеяния С. Брэдфорда строго выполняется лишь при соблюдении следующих условий:

- 1. Отрасль, тема или предмет должны быть четко определены.
- 2. Рассматриваемый список периодических изданий по такой отрасли, теме или предмету, а также учет содержащихся в этих периодических изданиях релевантных публикаций должны быть полными.
- 3. Время, за которое рассматриваются периодические издания, должно быть четко определено и учтены все помещенные в эти издания релевантные публикации.

Большая заслуга Дж. Ципфа, С. Брэдфорда и А. Лотки состоит в том, что они положили начало строгому исследованию документальных потоков, которые представляют собой совокупности научных документов-публикаций неопубликованных материалов. Дальнейшие исследования, среди которых видное занимают работы В.И. Горьковой, показали, что можно определять количественные параметры не только совокупностей научных документов, но и совокупностей элементов признаков научных документов: авторов, терминов, индексов классификационных систем, наименований изданий, т.е. наименований элементов, характеризующих содержание научных документов. Например, можно расположить журналы в порядке убывания числа печатающихся в них авторов, в порядке убывания средней величины публикующихся в них статей или упорядочить совокупность документов по любому ее элементу.

Закон В.И. Горьковой

В.И. Горькова доказала, что одним из критериев, позволяющих оценить развитие отдельных научных дисциплин и их составляющих, может являться

средняя скорость роста количества публикаций за интервал времени Δt. Автором предложена следующая формула:

$$V = \frac{\Delta N}{N \Delta t'}$$

где N - число публикаций в базисном году; ΔN - прирост за период Δt ; V - скорость роста числа публикаций, определяемая их приростом (ΔN) во времени (Δt) по отношению к базисному году (N) по данным экспериментальных исследований за ряд лет (более 5).

Данные законы и установленные закономерности являются наиболее распространенными в исследованиях документопотоков, связаны с важными явлениями в научных коммуникациях и имеют большое практическое значение.

Значение законов распределения состоит в том, что их основатели положили начало системному исследованию документопотоков. Можно утверждать, что эти законы являются методологической основой при изучении документопотоков. Их применение в управлении массивами журналов или для оценки научного влияния журналов, авторов, статей, исследований - важное условие информационной поддержки научной деятельности.

Метод моделирования лексического анализа документальных БД

Метод оценки тенденций развития научного направления, основанный на анализе лексики проблемно-ориентированных БД, предложен Е.Ю. Павловска. Автор предлагает выбрать определенное научное направление, для которого в БД можно проследить моменты зарождения научных направлений исследований по какой-либо проблеме, их развитие, распад или трансформацию в новые направления, попытаться выявить закономерности в динамических характеристиках информационного потока в различные периоды его "жизни" и, когда такие закономерности будут найдены, проверить высказанную гипотезу либо на других массивах, либо в другом временном интервале.

Контент-анализ

Контент-анализ при исследовании развития научных направлений используется группа методов, основанных на анализе текстов документов, и берущих начало от социологического метода контент-анализа. Различные модификации контент-анализа направлены на получение объективной информации о некоторой совокупности однородных документов (информационного потока) путем фиксации существенных характеристик содержания и их количественного описания. Контент-анализ плодотворен при изучении больших объемов текста, в частности интернет-ресурсов и полнотекстовых изданий и коллекций. Единицей анализа может являться слово или некоторая совокупность слов, выражающая проблему (предмет, отрасль, направление и др.), автор, географическая рубрика и др.

Метод совместной встречаемости ключевых слов

К середине 50-х гг. XX в. исследователи стали обращать внимание не только на простое наличие или отсутствие единиц анализа в тексте, но и на связи между ними, т.е. совместную встречаемость (сооссигенсе) слов различных категорий. Для этой цели подсчитывают обычный коэффициент корреляции, который указывает на связи между категориями и знак этой связи. В результате может оказаться, что для некоторых категорий наблюдается тенденция их совместного употребления, а для других - наоборот. На этой основе создан так на зываемый метод "co-word analysis", предложенный группой исследователей в 1984 г. (J.P. Courtial, M. Callon, M. Sigogneau) и применимый к различным информационным ресурсам. Каждое поле в

полнотекстовом источнике или БД может быть использовано для анализа, чаще это дескрипторы или ключевые слова. Метод является полезным инструментом для картирования науки, так как способствует выявлению ассоциаций среди дескрипторов или терминов и на их базе построению сетей, отражающих картину эволюции любой дисциплины. Для этой цели создаются указатель включения (inclusion index) и указатель близости (proximity index), используемые для измерения силы связей между отраженными понятиями. С помощью этих указателей понятия объединяются в кластеры и представляются на сетевых картах. Результаты исследований позволяют делать выводы о постоянстве или "смене акцентов" в определенных темах исследовательских областей, проследить "созревание" или утрату значения научных проблем. Данный метод, по мнению ученых, применим ко многим тематическим областям при наличии достаточного "корпуса" текстов и поддерживающего его программного обеспечения.

Метод семантического спектра

Метод семантического спектра позволяет в графической форме исследовать динамику ключевых слов во времени и рассматривать разнообразные аспекты структурных изменений в различных областях. Метод также основан на анализе частот встречаемости ключевых слов. По мере развития научного направления их частота растет, прекращение работ по какому-либо научному направлению приводит к уменьшению частоты встречаемости определенных терминов. Таким образом, в совокупности выделяют три составляющих: низкочастотная - соответствующая развивающемуся научному направлению; высокочастотная присущая установившейся терминологии, низкочастотная характерная ДЛЯ научных направлений, работы по которым прекращаются или переходят в стадию производства. Анализ динамики "семантического спектра" позволяет выявлять эти составляющие и, соответственно, делать выводы об интенсивности развития научных направлений или их возникновении; другими словами, динамика "семантического спектра" отражает структурно-частотные изменения в документальных массивах, которые в свою очередь являются индикатором возникновения и развития научных направлений.

Метод логико-смыслового моделирования

Еще один разработанный метод, основанный на анализе текстов документов, - метод логико-смыслового моделирования, предложенный М. Субботиным. Он основан на использовании в качестве исходных элементов любых высказываний, которые могут быть выражены отдельным словом, словосочетанием или целым предложением. Для каждого высказывания выделяются все его непосредственные логические связи с другими высказываниями в данной предметной области, образуя неориентированный граф, вершинами которого являются высказывания, а ребрами - связи между ними, - так называемый ЛС-граф, рассматриваемый как логико-смысловая модель данной области. Метод позволяет следить за развитием предметной области: многосвязные и центральные высказывания в ЛС-графах свидетельствуют о наиболее важных проблемах на данном этапе развития, смещения этих центров на ЛС-графах следующих этапов свидетельствуют об определенных тенденциях развития.

При формировании логико-смысловых графов исследователь должен пользоваться определенными критериями и процедурами, чтобы отличать прямую связь от косвенной. Смежными по смыслу считаются лишь те понятия и утверждения, которые можно объединить при помощи логических связок ("есть", "является причиной", "поэтому", "в этих целях" и т.п.). Сегодня логико-смысловые графы приблизились к гипертекстам, их можно читать, вставляя при переходе к смежному узлу соответствующую логическую связку.

Анализ цитирования

Система методов изучения документопотока на основе анализа цитирования, применяемых для получения структурной картины состояния научных исследований и оценки результативности ученых, разрабатывается интенсивно. В данном случае объектом анализа является поток цитируемых и цитирующих публикаций, а предметом - количественные его характеристики. Анализ цитирования чаще всего проводят на базе специальных вторичных источников информации - указателей научных ссылок, например в изданиях Института научной информации США ("Science Citation Index" и др.). Ссылки, являясь средством научной коммуникации, позволяют проследить развитие научного направления, дают общие сведения о проблеме, отражают литературу, создающую контекст данной работы. Исследование внутренней структуры областей знания, выявление исторических особенностей и тенденций развития науки и техники, основанных на анализе сетей цитирования публикаций, подробно освещены в ряде работ Ю. Гарфилда, В.А. Маркусовой, И.В. Маршаковой.

Вместе с тем исследователи отмечают ряд значимых недостатков анализа цитирования. Напри мер, существует такой феномен, получивший название "Спящая красавица в науке", когда статья после выхода в свет остается незамеченной длительное время, но потом неожиданно привлекает к себе большое внимание. Примером такой работы, сильно опередившей свое время, является исследование Г. Менделя по генетике растений, опубликованное в 1865 г. и остававшееся в тени 34 года. Спорности и "опасности" индекса цитируемости как главного критерия оценки качества научной деятельности посвящена работа О.В. Михайлова. Одним из подходов к более достоверному анализу цитирования является дифференцированный подход. О его необходимости впервые стал говорить А.И. Уемов. В дальнейшем было предложено выделять пять типов цитирований: авторитетное, конструктивное, информирующее, критическое и негативное цитирование. Эта классификация была применена Э. Ваттером для анализа структуры цитирования в области информатики. Еще более детальная классификация цитирований была предложена Е.Д. Гражданниковым и Т.В. Сорокиной в 1976 г. Сегодня в науковедении выделяют положительное (справочное, обзорное, конструктивное и др.) и отрицательное (критическое, отвергающее, обвиняющее и др.) цитирования, позволяющие более объективно проанализировать цитируемый документопоток.

На основе принципов цитирования предложены дополнительные методы. Наиболее употребляемые среди них следующие: метод ко-цитирования и кластерный анализ, метод библиографического сочетания.

Метод ко-цитирования и кластерный анализ

Используя метод совместного цитирования двух публикаций третьей, так называемого ко-цитирования (от англ. "co-citation"), определяют тенденции развития коммуникационных процессов в науке. Данный метод, разработанный одновременно Г. Смоллом и И. Маршаковой, позволяет сгруппировать статьи в зависимости от принадлежности к той или иной тематике; вносить коррективы в полученную структуру в процессе выявления новых библиографических ссылок, отражающих изменения во времени исследуемого направления, и тем самым может служить инструментом оперативной коррекции научных проектов.

В основе метода ко-цитирования лежит принцип выделения взаимосвязи между двумя публикациями на основе цитирования их одними и теми же документами. Родство публикаций или авторов публикаций определяется числом работ, цитирующих одновременно обе статьи, т.е. появляются невидимые связи, которые при наглядном представлении образуют смысловые сгустки (кластеры).

Начиная с 1981 г. метод ко-цитирования используется в Институте научной информации (США) при построении кластеров публикаций, отражающих активные исследовательские фронты в различных областях знаний. Кластерный анализ является действенным инструментом картирования различных свойств, выявления активных исследовательских фронтов или динамики исторического развития определенных научных направлений или дисциплин. Как отмечает С.А. Рожков, "кластер - это своего рода "образ" исследовательской области, созданный всем стоящим за ним мировым научным сообществом ученых, работающих в данном направлении".

Метод библиографического сочетания

Выявлять СВЯЗИ В документопотоке предлагается методом также библиографического сочетания ("bibliographic coupling"), который был предложен М. Кесслером в 1963 г. В основе этого метода лежит принцип выделения взаимосвязи между двумя публикациями на том основании, что цитируется один и тот же документ, причем интенсивность их взаимосвязи определяется числом библиографических ссылок, общих для обеих публикаций. В соответствии с этим методом сила связи двух публикаций определяется числом приведенных в этих публикациях общих (одинаковых) ссылок, и на этой основе устанавливается связь между публикациями (на основании общих в них ссылок). Еще один вид взаимосвязи между двумя публикациями посредством ссылок - анализ предметного направления, который проводится путем изучения количества ссылок, содержащихся в первой публикации, на работы, содержащиеся во второй. По методу Кесслера две публикации прочно связаны и эта связь не меняется при появлении новых публикаций, т.е. не зависит от изменений, происходящих в науке. Исходя из этого, такую связь называют ретроспективной, в отличие от другого метода -ко-цитирования, где эта связь является проспективной и позволяет исследовать совместно цитируемые публикации в новых работах.

Заключение

Все перечисленные методы анализа ДИП имеют свои особенности, достоинства и недостатки. Каждый метод имеет более или менее определенную область применения, в границах которой он эффективен.

Методы анализа количественных характеристик первичных документов можно назвать перспективными, учитывая увеличение числа полнотекстовых информационных ресурсов. Методы анализа вторичных источников информации, основанные на классических библиометрических законах, анализе видовой структуры и количественных характеристик документопотоков (мобильность, стабильность, скачкообразность), позволяют оперативно следить и выявлять развивающиеся и затухающие направления. Вопрос о целесообразности применения того или иного метода решается в зависимости от цели исследования, изучаемой проблемы, степени ее разработанности в литературе, от тех информационных ресурсов, которыми располагает исследовательский коллектив.

Применение закона Ципфа

Проверим закон Ципфа для детской сказки о золотой рыбке (на англ). Напишем программу на языке питон, которая берет все слова из текстового файла "Story.txt", считает сколько раз они встречаются,строит таблицу на основании полученных данных и построит график зависимости частоты встречаемого слова от ранга

```
import re
import plotly.graph_objects as go
import matplotlib.pyplot as plt

if __name__ == '__main__':
    dictionary = {}
    items = {
        'words': [],
        'values': [],
        'rates': [],
```

```
'multiplication rate': []
        with open("Story.txt", 'r') as book:
            line = book.readline()
             punctuations = ['(', ')', ';', ':', '[', ']', ',',
111, 121, 11]
            word count = 0
            while line:
                tokens = re.split('[ ()\';:\[\]!?,\n\t.]', line)
                line = book.readline()
                 words = [word.lower() for word in tokens if word
not in punctuations and word != '']
                word count += len(words)
                for word in words:
                    if word in dictionary:
                        dictionary[word] += 1
                    else:
                        dictionary[word] = 1
            i = 0
                   for key, value in sorted(dictionary.items(),
key=lambda item: item[1], reverse=True):
                i += 1
                items['words'].append(key)
                items['values'].append(value)
                items['rates'].append(i)
            for index in range(0, len(items['words']) - 1):
                coef = float(items['values'][index] / word count)
                items['rate'].append(coef)
```

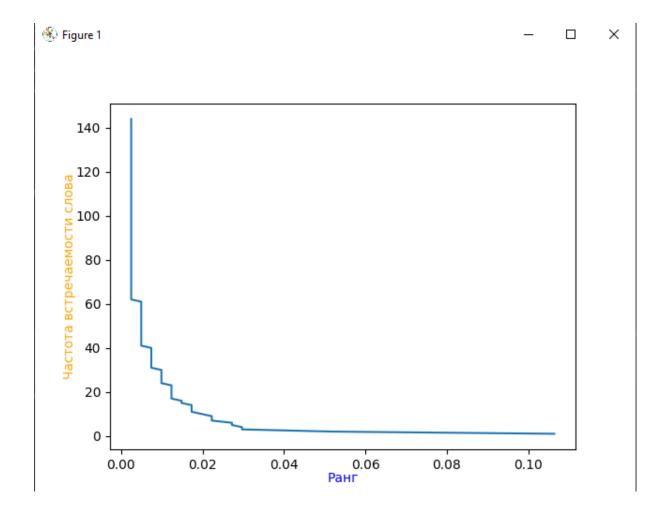
```
items['multiplication_rate'].append(float(items['rates'][index] *
coef))
                                             table
go.Figure (data=[go.Table(header=dict(values=['Слово', 'Сколько раз
встречается', 'Место', 'Частота встречаемости', 'Место *
Частота']), cells=dict(values=[items['words'], items['values'],
items['rates'], items['rate'], items['multiplication rate']]))])
     plt.plot(items['rate'],
items['rates'][0:len(items['rates'])-1])
    ax = plt.gca()
    ax.set_xlabel("Ранг", fontsize=10, color='blue', labelpad=0)
# +
    ax.set ylabel("Частота встречаемости слова", fontsize=10,
color='orange', labelpad=0) # +
     plt.show()
     table.show()
```

_	-			
Слово	Сколько раз встречается	Место	Частота встречаемости	Место * Частота
the	43	1	0.10643564356435643	0.10643564356435643
to	21	2	0.05198019801980198	0.10396039603960396
a	12	3	0.0297029702970297	0.0891089108910891
and	12	4	0.0297029702970297	0.1188118811881188
of	11	5	0.027227722772277228	0.13613861386138615
fisherman	11	6	0.027227722772277228	0.16336633663366337
was	9	7	0.022277227722772276	0.15594059405940594
fish	9	8	0.022277227722772276	0.1782178217821782
he	9	9	0.022277227722772276	0.20049504950495048
sea	8	10	0.019801980198019802	0.19801980198019803
in	7	11	0.017326732673267328	0.1905940594059406
his	7	12	0.017326732673267328	0.20792079207920794
wish	7	13	0.017326732673267328	0.22524752475247528
wife	7	14	0.017326732673267328	0.2425742574257426
for	6	15	0.01485148514851485	0.22277227722772275
she	6	16	0.01485148514851485	0.2376237623762376
but	5	17	0.012376237623762377	0.21039603960396042
golden	5	18	0.012376237623762377	0.22277227722772278
returned	5	19	0.012376237623762377	0.23514851485148516
her	5	20	0.012376237623762377	0.24752475247524752
sent	5	21	0.012376237623762377	0.2599009900990099
loaf	5	22	0.012376237623762377	0.2722772277227723
bread	5	23	0.012376237623762377	0.2846534653465347

ned 4 24 28 0.0000000000000000000000000000000000					
Casaphi	not	4	24	0.009900990099009901	0.2376237623762376
Casaphi	41-4	4	25	0.000000000000000001	0.24752475247524752
this	tnat	4	25	0.00990099009901	0.24/524/524/524/52
this 4 28 0.000900000000000001 0.2772777277777777777777777777777777777	old	4	26	0.009900990099009901	0.25742574257425743
this 4 28 0.000900000000000001 0.2772777277777777777777777777777777777	cought	4	27	0.000000000000000001	0.26722672267226724
there	caugnt	4	2/	0.00990099009901	0.20/320/320/320/34
Home	this	4	28	0.009900990099009901	0.27722772277227725
Home	da	4	20	0.000000000000000001	0.2071207120712071
Index	day	4	29	0.00990099009901	0.28/128/128/128/1
time	home	4	30	0.009900990099009901	0.297029702970297
Indit				0.00740574057405	
Far 3 3 34 0.007/127/1277/227732 0.2260/20090090090090090	time	3	31	0.00/425/425/425/425	0.23019801980198018
Far 3 3 34 0.007/127/1277/227732 0.2260/20090090090090090	land	3	32	0.007425742574257425	0.2376237623762376
on 3 3 3 4 0.007437427427427 50 0.25909000000000000000000000000000000000					
would 3 3 3 5 0.007437427427427 0.02790000000000000000000000000000000000	far	3	33	0.00/425/425/425/425	0.24504950495049505
would 3 3 3 5 0.007437427427427 0.02790000000000000000000000000000000000	on	3	34	0.007425742574257425	0.2524752475247525
when 3 3 37 .0.007427427427425 0.257227227227274 new 3 3 8 .0.007427427427425 0.2572472742745 0.257247427437437437437437437437437437437437437437					
husband	would	3	35	0.00/425/425/425/425	0.2599009900990099
husband	when	3	36	0.007425742574257425	0.26732673267326734
new					
all 3 9 0.00742742742742742 0.02990909090909090909090909090909090909	husband	3	37	0.007425742574257425	0.2747524752475247
all 3 9 0.00742742742742742 0.02990909090909090909090909090909090909	new	3	38	0.007425742574257425	0.28217821782178215
all 3 40 0.0074274274274275 0.0297070270270270270270270270270270270270270					
shack 2	find	3	39	0.007425742574257425	0.2896039603960396
shack 2	all	3	40	0.007425742574257425	0.297029702970297
### STATE 1.00.099904990499049049049049049060 0.22792722722722728 ### 0.00499049049049049049049049049049060 0.227927272727272727272727272727272727272					
### ### #############################	once	2	41	0.0049504950495049506	0.20297029702970298
### ### #############################	shack	2	42	0.0040504050405040506	0.2070207020702070
Bilk	SildCK				
Bike	morning	2	43	0.0049504950495049506	0.21287128712871287
Bike		2	44	0.0040504050405040506	0.21702170217021702
Mindhearted 2	as	2	44	0.0049304930493049306	0.21/821/821/821/82
Mindhearted 2	like	2	45	0.0049504950495049506	0.22277227722772278
Netro	Librard Comment of	2	45	0.0040504050405040506	0.0077007700770
however	Kindnearted	2	40	0.0049304930495049506	0.22//22//22//3
however					
however 2 47 0.0049504950495049506 0.2326732673267326732673267326732673267326	0	Curani un	M	Heren	Mc * II-
s 2 48 0.0049049049049050 0.23762372372372 back 2 2 49 0.0049049049049050 0.2475247523762272 back 2 2 50 0.0049049049049050 0.24752475247527457 back 2 2 51 0.0049049049049049050 0.247524752475275 dath 2 2 51 0.0049049049049049050 0.25247524752475275 dath 2 2 52 0.004904904904909050 0.25247524752475276 dath 2 2 53 0.004904904904909050 0.2524752475247524 with 2 2 54 0.004904904904909090 0.2524752475247624 with 2 2 55 0.0049049049049090 0.2527627367326734 with 2 2 55 0.0049049049049090 0.252762736736736736734 hashbir 2 2 55 0.0049049049049090 0.2727727272727373 bit 2 2 57 0.0049049049049090 0.2727727272727373 bit 2 2 57 0.0049049049049090 0.2727727272727373 bit 2 2 59 0.0049049049049090 0.27277272772727373 bit 2 2 59 0.0049049049049090 0.2727727272727373 dressed 2 2 59 0.0049049049049090 0.28271827128712871287 dressed 2 2 59 0.0049049049049090 0.28271827128712871287 dressed 2 2 50 0.0049049049049090 0.28271827128712871287 dressed 2 2 50 0.0049049049049090 0.2827182712871287 dressed 3 2 0.0049049049049090 0.2827182772972972972972972 dreen 2 0.0049049049049090 0.2827182772972972972972972972972972972972972972	Слово	сколько раз встречается	Место	частота встречаемости	место - частота
s 2 48 0.0049049049049050 0.23762372372372 back 2 2 49 0.0049049049049050 0.2475247523762272 back 2 2 50 0.0049049049049050 0.24752475247527457 back 2 2 51 0.0049049049049049050 0.247524752475275 dath 2 2 51 0.0049049049049049050 0.25247524752475275 dath 2 2 52 0.004904904904909050 0.25247524752475276 dath 2 2 53 0.004904904904909050 0.2524752475247524 with 2 2 54 0.004904904904909090 0.2524752475247624 with 2 2 55 0.0049049049049090 0.2527627367326734 with 2 2 55 0.0049049049049090 0.252762736736736736734 hashbir 2 2 55 0.0049049049049090 0.2727727272727373 bit 2 2 57 0.0049049049049090 0.2727727272727373 bit 2 2 57 0.0049049049049090 0.2727727272727373 bit 2 2 59 0.0049049049049090 0.27277272772727373 bit 2 2 59 0.0049049049049090 0.2727727272727373 dressed 2 2 59 0.0049049049049090 0.28271827128712871287 dressed 2 2 59 0.0049049049049090 0.28271827128712871287 dressed 2 2 50 0.0049049049049090 0.28271827128712871287 dressed 2 2 50 0.0049049049049090 0.2827182712871287 dressed 3 2 0.0049049049049090 0.2827182772972972972972972 dreen 2 0.0049049049049090 0.2827182772972972972972972972972972972972972972	b and	2		0.004050405040504050	0.00067006700670067
So	however	2	4/	0.0049504950495049506	0.23267326732673269
So	S	2	48	0.0049504950495049506	0.2376237623762376
back					
did 2 2 52 0.004950495049506 0.2247524752475274751 next 2 2 53 0.00495049506 0.0247524752475247524 next 2 2 53 0.00495049506 0.0247524752475247524 next 2 2 53 0.00495049506 0.022776277627762767326734 with 2 2 54 0.004950495049506 0.02673267326734 next 2 2 55 0.004950495049506 0.02673267326734 next 2 2 55 0.004950495049506 0.0277227727727723 still 2 2 55 0.004950495049506 0.02772277227727723 next 2 2 57 0.004950495049506 0.027722772277277273 next 2 2 57 0.0049504950495060 0.0277227722772277227722772277227722 next 2 2 59 0.0049504950495060 0.027722772277227722772277227722 next 2 2 59 0.0049504950495060 0.02772277227722772277227722 next 2 2 59 0.0049504950495060 0.0277227722772277227722 next 2 2 50 0.0049504950495060 0.0277227722772277227722 next 2 2 50 0.0049504950495060 0.02772277227722772277227722 next 2 2 50 0.0049504950495060 0.02772277227722772277227722772277227722	S0	2	49	0.0049504950495049506	0.2425/425/42574257
did 2 2 52 0.004950495049506 0.2247524752475274751 next 2 53 0.004950495049506 0.22475247524752745 next 2 53 0.004950495049506 0.26732673267326 with 2 2 54 0.004950495049506 0.26732673267326 with 2 2 55 0.004950495049506 0.26732673267326734 washtub 2 5 55 0.004950495049506 0.26732673267326734 washtub 2 7 55 0.004950495049506 0.277227727727723 still 2 7 50 0.0049504950495049506 0.2772277227727723 still 2 9 56 0.0049504950495049506 0.27722772277277273 still 2 9 57 0.0049504950495049506 0.2772277227722772727 become 2 5 58 0.0049504950495049506 0.287128712871282 become 2 6 0.0049504950495049506 0.287128712871287 queen 2 6 0.0049504950495049506 0.287128712871287 queen 2 6 0.0049504950495049506 0.287029702970297050 queen 2 6 0.0049504950495049506 0.287029702970297050 queen 3 7 6 0.0049504950495049506 0.287029702970297029702970297029702970297029	hack	2	50	0.0049504950495049506	0.24752475247524752
next					
next	catch	2	51	0.0049504950495049506	0.2524/524/524/525
next	did	2	52	0.0049504950495049506	0.25742574257425743
with 2 54 0.0049504950495049506 0.26732673267326734 washtub 2 55 0.004950495049506 0.227272277227723 still 2 56 0.004950495049506 0.227272277227723 it 2 57 0.0049504950495049506 0.227272277227723 it 2 5 50 0.0049504950495049506 0.22727227722772277 it 2 5 50 0.00495049504950495049506 0.2821782178217822 become 2 5 58 0.0049504950495049506 0.2821782178217822 dressed 2 5 59 0.0049504950495049506 0.2821782178217822 queen 2 6 60 0.0049504950495049506 0.2920792079209720 even 2 6 61 0.0049504950495049506 0.29207920792079209 upon 1 1 62 0.002475247524753475 0.15394654653465346 away 1 1 63 0.002475247524753475 0.15394654653465346 away 1 1 64 0.002475247524753475 0.15394059405904 lived 1 64 0.002475247524753475 0.15394059405904 very 1 65 0.002475247524753475 0.158940594059109 poor 1 66 0.002475247524753475 0.1639603366336337 couple 1 67 0.002475247524753475 0.1639603366336337 couple 1 67 0.002475247524753475 0.16391841584158415841584158415841584158415841					
washtub 2 55 0.049504950495049506 0.2722772277227723 still 2 56 0.049504950495069 0.27727727722772277237 it 2 57 0.049504950495049506 0.28217821782178217821782178217821 decome 2 58 0.0049504950495049506 0.28217821782178217 dessed 2 59 0.0049504950495049506 0.28207920720792079 even 2 61 0.0049504950495049506 0.297029702970297 even 2 61 0.0049504950495049506 0.0297029702970297 upon 1 62 0.00247524752475247534733 0.153465346534653463 away 1 63 0.00247524752475247534733 0.15841584158415842 very 1 65 0.00247524752475247534733 0.15891089108910 poor 1 67 0.00247524752475247534733 0.16891089108910 poor 1 67 0.00247524752475247534733 0.168916891689108910 poor 1 67 0.00247524752475247534733 0.16891689168918	next	2	53	0.0049504950495049506	0.2623762376237624
washbub 2 55 0.049504950495049506 0.2722772277227723 still 2 56 0.049504950495069 0.27727727722772277237 it 2 57 0.049504950495049506 0.28217821782178217821782178217821 dessed 2 59 0.049504950495049506 0.28217821782178217 queen 2 60 0.049504950495949506 0.297029702970297 even 2 61 0.049504950499596996906 0.297029702970297 upon 1 62 0.0024752475247524753473 0.15346534653465346 away 1 63 0.00247524752475247534733 0.15581594059405940994099 lived 1 64 0.00247524752475247534733 0.15841594158415842 very 1 65 0.00247524752475247534733 0.165816981098109 poor 1 67 0.00247524752475247534733 0.165816981681858 from 1 68 0.00247524752475247534733 0.165841584158611865 from 1 69 0.00247524752475247534733 0.17072	with	2	54	0.0049504950495049506	0.26732673267326734
### 1811 2 56 0.0049504950495049506 0.27722772772775 ### 2 57 0.004950495049506 0.28217821782178217821 ### 2 57 0.004950495049506 0.28217821782178217821 ### 2 58 0.0049504950495049506 0.28207920792079206 ### 2 60 0.0049504950495049506 0.29207920792079206 ### 2 60 0.0049504950495049506 0.29207920792079206 ### 2 61 0.0049504950495049506 0.29207920792079206 ### 2 61 0.0049504950495049506 0.292079207920792079206 ### 2 61 0.0049504950495049506 0.292079207920792079206 ### 2 62 0.0024752475247524753 0.1534634653465346 ### 3 0.0024752475247524753 0.1534634653465346 ### 3 0.0024752475247524753 0.1534634653465346 ### 2 1 64 0.002475247524753473 0.1534618415841584 ### 2 1 65 0.002475247524753473 0.153461841584 ### 2 1 66 0.002475247524753473 0.163816336336337 ### 2 0.0024752475247524753 0.163816336336336337 ### 2 0.002475247524753473 0.1638163316833 ### 2 0.002475247524753473 0.1638163316833 ### 2 0.002475247524753473 0.1638163316833 ### 3 0.002475247524753473 0.177920792079208 ### 3 1 71 0.002475247524753473 0.177920792079208 ### 3 1 71 0.002475247524753473 0.17874257425742574 ### 3 0.002475247524753473 0.17874257425742574 ### 3 0.002475247524753473 0.1881818181818181 ### 1 74 0.0024752475247524753473 0.1881881881881 ### 1 74 0.0024752475247524753473 0.1881881188118811 ### 1 76 0.0024752475247524753473 0.1881881188118811 ### 1 77 0.0024752475247524753473 0.1881881188118811 ### 1 78 0.0024752475247524753473 0.1881881188118811 ### 1 76 0.0024752475247524753473 0.1881881188118811 ### 1 78 0.0024752475247524753473 0.1881881188118811 ### 1 79 0.0024752475247524753473 0.19809309309307 ### 1 81 0.0024752475247524753473 0.1980930930930930930930930930930930930930930					
it 2 57 0.0049504950495049506 0.2821782178217822 become 2 58 0.0049504950495049506 0.2871287128712871 dressed 2 59 0.0049504950495049506 0.2871287128712871 dressed 2 59 0.0049504950495049506 0.287028702907907906 queen 2 60 0.0049504950495049506 0.2970297029702907906 even 2 61 0.0049504950495049506 0.0198019801990197 upon 1 62 0.00247524752475247524753 0.15346534653465346 away 1 0.03 0.00247524752475247524753 0.15346534653465346 away 1 0.00247524752475247524753 0.15346534653465346 away 1 0.00247524752475247524753 0.15346534653465346 away 1 0.00247524752475247524753 0.15346534653465346 away 1 0.00247524752475247524753 0.153465346534653465346 away 1 0.00247524752475247524753 0.163693109108910 poor 1 0.00247524752475247524753 0.163963369336337 couple 1 0.00247524752475247524753 0.163863366336337 couple 1 0.00247524752475247524753 0.16381831631833 away 1 0.00247524752475247524753 0.16381831631833 away 1 0.00247524752475247524753 0.16381831631833 away 1 0.00247524752475247524753 0.16381831631833 away 1 0.00247524752475247524753 0.16381831631831833 away 1 0.00247524752475247524753 0.16381831631831833 away 1 0.00247524752475247524753 0.16381831631831833 away 1 0.0024752475247524753 0.17326732673267328 away 1 0.0024752475247524753 0.183683163316 away 1 0.0024752475247524753 0.193693069307 away 1 0.0024752475247524753 0.1936930693097 away 1 0.0024752475247524753 0.002475247524753 0.002475247524753 0.002475247524753 0.002	washtub	2	55	0.0049504950495049506	0.2722772277227723
it 2 57 0.0049504950495049506 0.2821782178217822 become 2 58 0.0049504950495049506 0.2871287128712871 dressed 2 59 0.0049504950495049506 0.2871287128712871 dressed 2 59 0.0049504950495049506 0.287028702907907906 queen 2 60 0.0049504950495049506 0.2970297029702907906 even 2 61 0.0049504950495049506 0.0198019801990197 upon 1 62 0.00247524752475247524753 0.15346534653465346 away 1 0.03 0.00247524752475247524753 0.15346534653465346 away 1 0.00247524752475247524753 0.15346534653465346 away 1 0.00247524752475247524753 0.15346534653465346 away 1 0.00247524752475247524753 0.15346534653465346 away 1 0.00247524752475247524753 0.153465346534653465346 away 1 0.00247524752475247524753 0.163693109108910 poor 1 0.00247524752475247524753 0.163963369336337 couple 1 0.00247524752475247524753 0.163863366336337 couple 1 0.00247524752475247524753 0.16381831631833 away 1 0.00247524752475247524753 0.16381831631833 away 1 0.00247524752475247524753 0.16381831631833 away 1 0.00247524752475247524753 0.16381831631833 away 1 0.00247524752475247524753 0.16381831631831833 away 1 0.00247524752475247524753 0.16381831631831833 away 1 0.00247524752475247524753 0.16381831631831833 away 1 0.0024752475247524753 0.17326732673267328 away 1 0.0024752475247524753 0.183683163316 away 1 0.0024752475247524753 0.193693069307 away 1 0.0024752475247524753 0.1936930693097 away 1 0.0024752475247524753 0.002475247524753 0.002475247524753 0.002475247524753 0.002	etill	2	56	0.0040504050405040506	0.27722772277227725
become 2 58 0,0049504950495066 0.2871287128712871 dressed 2 59 0.0049504950495049506 0.29207920792079207 queen 2 60 0.0049504950495049506 0.29702770297029702970 even 2 61 0.004950495049504950690506 0.29702770297029702970 upon 1 62 0.00475247524752475247524753 0.1584059405940594 lived 1 63 0.00247524752475247524753 0.1584059405940594 lived 1 64 0.00247524752475247524753 0.158415815815415842 very 1 65 0.00247524752475247524753 0.1584158158158415842 very 1 66 0.00247524752475247524753 0.163863366336337 couple 1 67 0.00247524752475247524753 0.1638168316833 edge 1 67 0.00247524752475247524753 0.16381683168316833 couple 1 70 0.00247524752475247534753 0.16381683168316833 chege 1 70 0.00247524752475247524753					
dressed 2 59	it	2	57	0.0049504950495049506	0.2821782178217822
dressed 2 59	hecome	2	50	0.0040504050405040506	0.2071207120712071
queen 2 60 0.00495049504950495049506 0.2970297029702970297029702970297029702970					
even 2 61 0.0049504950495049506 0.30198019801970 ироп 1 62 0.0024752475247524753 0.15346534653465346534653465346534653465346	dressed	2	59	0.0049504950495049506	0.29207920792079206
even 2 61 0.0049504950495049506 0.3019801980197 upon 1 62 0.0024752475247524753 0.15346534653465346534653465346534653465346	gueen	2	60	0.0040504050405040506	0.207020702070207
upon 1 62 0.0024752475247524753 0.153465346534653465346 away 1 63 0.00247524752475247524753 0.1594159405940594 lived 1 64 0.00247554752475247524753 0.15941584158415841 very 1 65 0.0024755475247524753 0.160891089109 poor 1 66 0.00247554752475247524753 0.16383636336336336336336336336336336336336	queen	2	00	0.0049304930493049306	0.29/029/029/029/
away 1 1 63 0.002475247524753 0.15594059405940594 fived 1 1 64 0.002475247524753 0.15694059405940594 fived 1 1 64 0.0024752475247534753 0.15698189189199 рог 1 1 65 0.0024752475247534753 0.1669819819199 рог 1 66 0.0024752475247534753 0.1669819819199 рог 1 67 0.0024752475247534753 0.1633663366336633637 соцре 1 67 0.0024752475247534753 0.1633633663366336533 edge 1 1 68 0.0024752475247534753 0.16584158415841585 from 1 68 0.002475247524753 0.16584158415841583 edge 1 1 69 0.002475247524753753 0.16584158415841583 edge 1 1 70 0.002475247524753 0.1707920792079208	even	2	61	0.0049504950495049506	0.30198019801980197
ilved 1 1 63 0.002475247524753 0.15594059405940594 ilved 1 1 64 0.002475247524753 0.1569415941594159415941594159415941594159415			60	0.0004750475047504750	0.15346534653465346
lived 1 1 64 0.002475247524753 0.1584158415842 very 1 1 65 0.002475247524753 0.160891089109 рог 1 1 66 0.0024752475247534753 0.160891089109 рог 1 1 66 0.0024752475247534753 0.1636633663366337 соирle 1 1 67 0.0024752475247534753 0.1638168316833 edge 1 1 69 0.0024752475247534753 0.16381683168318831 edge 1 1 69 0.0024752475247534753 0.16381683168316831 edge 1 1 69 0.0024752475247534753 0.1707920792079209 cdge 1 1 70 0.0024752475247534753 0.1707920792079209 cdge 1 1 70 0.0024752475247534753 0.1707920792079209 cdge 1 1 70 0.002475247524753 0.170920792079209 cdge 1 1 70 0.002475247524753 0.180930693069307 cdge 1 1 70 0.002475247524753 0.1801881188118812 cdge 1 1 70 0.002475247524753 0.1801881188118812 cdge 1 1 70 0.002475247524753 0.190590059406 cdge 1 1 70 0.002475247524753 0.190590059406 cdge 1 1 80 0.0024752475247524753 0.190590059406 cdge 1 1 80 0.0024752475247524753 0.190590059406 cdge 1 80 0.0024752475247524753 0.2029702970297029702970297029702970297029	upon	1	02	0.0024/524/524/524/53	0.15340534053405340
lived	away	1	63	0.0024752475247524753	0.15594059405940594
very 1 1 65 0.0024752475247524753 0.160810891089109 рог 1 66 0.0024752475247524753 0.16336336336337 соирle 1 67 0.0024752475247524753 0.1633633663365336			64	0.0004750475047504750	0.15041504150415040
роог 1 66 0.0024752475247524753 0.16336633663366337 couple 1 67 0.0024752475247524753 0.16584158415841581 from 1 68 0.0024752475247524753 0.163816831683168316831 edge 1 1 69 0.0024752475247524753 0.163831683168316831 edge 1 1 69 0.0024752475247524753 0.1707920792079208 **Cлово **Cколько раз встречается	lived	1	64	0.0024/524/524/524/53	0.15841584158415842
роог 1 66 0.0024752475247524753 0.16336633663366337 couple 1 67 0.0024752475247524753 0.16336633663366337 strength of the str	very	1	65	0.0024752475247524753	0.1608910891089109
Couple 1 67 0.0024752475247524753 0.1658415841585168516851 from 1 68 0.0024752475247533 0.168316831683168316831863318633186331863			66	0.0004750475047504750	0.1600660066007
from 1 68 0.002475247524753 0.16831683168316831 edge 1 1 69 0.002475247524753 0.170792079208 Cлово	poor	1	00	0.0024/524/524/524/53	0.16336633663366337
from 1 68 0.002475247524753 0.16831683168316831 edge 1 1 69 0.002475247524753 0.170792079208 Cлово	couple	1	67	0.0024752475247524753	0.16584158415841585
edge 1 69 0.0024752475247524753 0.1707920792079208 Cлово Сколько раз встречается Mecto Частота встречаености Mecto ** Частота their 1 70 0.0024752475247524753 0.173267326732673267328 only 1 71 0.0024752475247524753 0.17574257425745273 means 1 72 0.0024752475247524753 0.178217821782 food 1 73 0.0024752475247534753 0.1805906930693069307 man 1 74 0.0024752475247534753 0.18316831683168316 one 1 75 0.0024752475247524753 0.185643564356435643564 usual 1 76 0.0024752475247524753 0.18811881188118812 routine 1 77 0.0024752475247524753 0.18811881881188118811881188118811881188					
Cnoso Сколько раз встречается Место Частота встречаемости Место " Частота their 1 70 0.0024752475247524753 0.17326732673267328 only 1 71 0.0024752475247524753 0.175742574257425742573 means 1 72 0.0024752475247524753 0.1806930693069307 man 1 74 0.002475247524753 0.1806930693069307 man 1 74 0.002475247524753 0.1831683168316 one 1 75 0.002475247524753 0.18811881188118812 oroutine 1 76 0.002475247524753 0.19059405940594059 took 1 78 0.002475247524753 0.19059405940594059 took 1 78 0.002475247524753 0.19306930693069307 fishing 1 79 0.002475247524753 0.19306930693069307 fishing 1 80 0.0024752475247534753 0.19306930693069307 down 1 81 0.0024752475247534753 0.1980198019801980198019801980198019801980	from	1	68	0.0024752475247524753	0.16831683168316833
Cnoso Сколько раз встречается Место Частота встречаемости Место " Частота their 1 70 0.0024752475247524753 0.17326732673267328 only 1 71 0.0024752475247524753 0.175742574257425742573 means 1 72 0.0024752475247524753 0.1806930693069307 man 1 74 0.002475247524753 0.1806930693069307 man 1 74 0.002475247524753 0.1831683168316 one 1 75 0.002475247524753 0.18811881188118812 oroutine 1 76 0.002475247524753 0.19059405940594059 took 1 78 0.002475247524753 0.19059405940594059 took 1 78 0.002475247524753 0.19306930693069307 fishing 1 79 0.002475247524753 0.19306930693069307 fishing 1 80 0.0024752475247534753 0.19306930693069307 down 1 81 0.0024752475247534753 0.1980198019801980198019801980198019801980	edae	1	69	0.0024752475247524753	0.1707920792079208
their 1 70 0.002475247524753 0.17326732673267328 only 1 71 0.002475247524753 0.173267326732673267326 only 1 71 0.0024752475247534753 0.17574257425742573 means 1 72 0.002475247524753 0.17821782178217821 food 1 73 0.0024752475247534753 0.1806930693069307 man 1 74 0.0024752475247534753 0.1806930693069307 man 1 75 0.0024752475247534753 0.18168316831683168316 one 1 75 0.0024752475247534753 0.18516831683168316 one 1 75 0.0024752475247534753 0.18516831683168316 usual 1 76 0.0024752475247534753 0.18516831683168316 took 1 77 0.0024752475247534753 0.19059405940594059405940594059405940594059					
their 1 70 0.002475247524753 0.17326732673267328 only 1 71 0.002475247524753 0.173267326732673267326 only 1 71 0.0024752475247534753 0.17574257425742573 means 1 72 0.002475247524753 0.17821782178217821 food 1 73 0.0024752475247534753 0.1806930693069307 man 1 74 0.0024752475247534753 0.1806930693069307 man 1 75 0.0024752475247534753 0.18168316831683168316 one 1 75 0.0024752475247534753 0.18516831683168316 one 1 75 0.0024752475247534753 0.18516831683168316 usual 1 76 0.0024752475247534753 0.18516831683168316 took 1 77 0.0024752475247534753 0.19059405940594059405940594059405940594059					
their 1 70 0.002475247524753 0.17326732673267328 only 1 71 0.002475247524753 0.173267326732673267328 only 1 71 0.002475247524753 0.17574257425742573 means 1 72 0.002475247524753 0.17821782178217821782 food 1 73 0.0024752475247534753 0.1806930693069307 man 1 74 0.0024752475247534753 0.1806930693069307 man 1 75 0.0024752475247534753 0.18168316831683168 fone 1 75 0.0024752475247534753 0.18168316831683168 one 1 75 0.002475247524753 0.188118811881188118 fone 1 77 0.002475247524753 0.188118811881188118 for 1 77 0.002475247524753 0.1816831683168316 fook 1 78 0.002475247524753 0.19059405940594059405940594059405940594059	Слово	Сколько раз встречается	Место	Частота встречаемости	Место * Частота
only 1 71 0.0024752475247524753 0.17574257425742573 means 1 72 0.0024752475247524753 0.17821782178217821782 food 1 73 0.0024752475247524753 0.1806930693069307 man 1 74 0.0024752475247524753 0.183168316831683168 one 1 75 0.0024752475247524753 0.185643564356435643 usual 1 76 0.0024752475247524753 0.18811881188118812 routine 1 77 0.0024752475247524753 0.19069306930693069306 took 1 78 0.0024752475247524753 0.193069306930693069306 took 1 78 0.0024752475247524753 0.193069306930693069307 fishing 1 79 0.0024752475247524753 0.1980198019803 down 1 81 0.0024752475247524753 0.1980198019803 down 1 81 0.0024752475247524753 0.2004950495049505 something 1 82 0.0024752475247524753 0.2004950495049504					
only 1 71 0.0024752475247524753 0.17574257425742573 means 1 72 0.0024752475247524753 0.17821782178217821782 food 1 73 0.0024752475247524753 0.1806930693069307 man 1 74 0.0024752475247524753 0.183168316831683168 one 1 75 0.0024752475247524753 0.185643564356435643 usual 1 76 0.0024752475247524753 0.18811881188118812 routine 1 77 0.0024752475247524753 0.19069306930693069306 took 1 78 0.0024752475247524753 0.193069306930693069306 took 1 78 0.0024752475247524753 0.193069306930693069307 fishing 1 79 0.0024752475247524753 0.1980198019803 down 1 81 0.0024752475247524753 0.1980198019803 down 1 81 0.0024752475247524753 0.2004950495049505 something 1 82 0.0024752475247524753 0.2004950495049504	their	1	70	0.0024752475247524753	0.17326732673267328
means 1 72 0.0024752475247524753 0.17821782178217821782 food 1 73 0.0024752475247524753 0.1806930693067307 man 1 74 0.0024752475247524753 0.18316831683168316 one 1 75 0.0024752475247533 0.18564356435644 usual 1 76 0.0024752475247533 0.1881188118811881188118 routine 1 77 0.0024752475247524753 0.1905940594059405 took 1 78 0.0024752475247524753 0.19309930693067 fishing 1 79 0.0024752475247524753 0.1995945544554555 net 1 80 0.0024752475247533 0.19901980198019801 down 1 81 0.0024752475247534753 0.19901980198019801 something 1 82 0.0024752475247524753 0.2004950495049505 something 1 83 0.0024752475247524753 0.2004950495049505 unusual 1 83 0.0024752475247524753 0.20549554455445546		1		0.0034753475347534753	
food 1 73 0.0024752475247524753 0.1806930693069307 man 1 74 0.0024752475247524753 0.1831683168316 one 1 75 0.0024752475247524753 0.1856435643564564 usual 1 76 0.0024752475247524753 0.188118811881188118812 routine 1 77 0.0024752475247524753 0.1905940594069406 took 1 78 0.0024752475247524753 0.19306930693069307 fishing 1 79 0.0024752475247524753 0.19554455445544554 net 1 80 0.0024752475247524753 0.1980198019803 down 1 81 0.0024752475247524753 0.200495049505 something 1 82 0.0024752475247524753 0.200495049505 unusual 1 83 0.0024752475247524753 0.20247029702970297029 unusual 1 84 0.0024752475247524753 0.205445544554654654654654654654654654654654	Offic	1			
food 1 73 0.0024752475247524753 0.1806930693069307 man 1 74 0.0024752475247524753 0.1831683168316 one 1 75 0.0024752475247524753 0.1856435643564564 usual 1 76 0.0024752475247524753 0.188118811881188118812 routine 1 77 0.0024752475247524753 0.1905940594069406 took 1 78 0.0024752475247524753 0.19306930693069307 fishing 1 79 0.0024752475247524753 0.19554455445544554 net 1 80 0.0024752475247524753 0.1980198019803 down 1 81 0.0024752475247524753 0.200495049505 something 1 82 0.0024752475247524753 0.200495049505 unusual 1 83 0.0024752475247524753 0.20247029702970297029 unusual 1 84 0.0024752475247524753 0.205445544554654654654654654654654654654654	means	1	72	0.0024752475247524753	0.1782178217821782
man 1 74 0.0024752475247524753 0.1831683168316 one 1 75 0.0024752475247524753 0.18564356435643564 usual 1 76 0.0024752475247524753 0.1881188118811881182 routine 1 77 0.0024752475247534753 0.1905940594059406 took 1 78 0.0024752475247524753 0.19306930693069307 fishing 1 79 0.0024752475247524753 0.19504554455445544555 net 1 80 0.0024752475247524753 0.1980198019803 down 1 81 0.0024752475247524753 0.2004950495049505 something 1 82 0.0024752475247524753 0.2004950495049505 something 1 83 0.0024752475247524753 0.20297029702970297029 unusual 1 83 0.0024752475247524753 0.20479029702970297029 begged 1 84 0.0024752475247524753 0.2103960396039604 spare 1 86 0.0024752475247524753 0.21287128712871287 <td></td> <td></td> <td></td> <td></td> <td></td>					
one 1 75 0.0024752475247524753 0.185643564356435644 usual 1 76 0.0024752475247524753 0.18811881188118812 routine 1 77 0.0024752475247524753 0.1905940594059406 took 1 78 0.0024752475247524753 0.19306930693069307 fishing 1 79 0.002475247524753 0.1955445544555544555 net 1 80 0.0024752475247524753 0.1980198019803 down 1 81 0.002475247524753 0.204950495049505 something 1 82 0.0024752475247524753 0.20297029702970297 unusual 1 83 0.0024752475247524753 0.20297029702970297 begged 1 84 0.0024752475247524753 0.2103960396039604 spare 1 86 0.0024752475247524753 0.21287128712871287 life 1 87 0.0024752475247524753 0.21782178217821782 return 1 89 0.0024752475247524753 0.2202970297029703	1000	1	/3		
one 1 75 0.0024752475247524753 0.185643564356435644 usual 1 76 0.0024752475247524753 0.18811881188118812 routine 1 77 0.0024752475247524753 0.1905940594059406 took 1 78 0.0024752475247524753 0.19306930693069307 fishing 1 79 0.002475247524753 0.1955445544555544555 net 1 80 0.0024752475247524753 0.1980198019803 down 1 81 0.002475247524753 0.204950495049505 something 1 82 0.0024752475247524753 0.20297029702970297 unusual 1 83 0.0024752475247524753 0.20297029702970297 begged 1 84 0.0024752475247524753 0.2103960396039604 spare 1 86 0.0024752475247524753 0.21287128712871287 life 1 87 0.0024752475247524753 0.21782178217821782 return 1 89 0.0024752475247524753 0.2202970297029703	man	1	74	0.0024752475247524753	0.18316831683168316
usual 1 76 0.0024752475247524753 0.18811881188118812 routine 1 77 0.0024752475247524753 0.190594059406 took 1 78 0.002475247524753 0.19306930693067 fishing 1 79 0.002475247524753 0.19554455445544555 net 1 80 0.002475247524753 0.19801980198019803 down 1 81 0.002475247524753 0.2004950495049505 something 1 82 0.0024752475247524753 0.2029702970297029702970297029702970297029					
routine 1 77 0.002475247524753 0.190594059406 took 1 78 0.002475247524753 0.1930693069307 fishing 1 79 0.0024752475247524753 0.1930693069307 fishing 1 80 0.0024752475247524753 0.195544554455455 net 1 80 0.0024752475247524753 0.1980198019803 down 1 81 0.0024752475247524753 0.2004950495049505 something 1 82 0.002475247524753 0.200495049505 something 1 83 0.002475247524753 0.200495049505 something 1 83 0.002475247524753 0.200495049505 happened 1 84 0.002475247524753 0.2054455445546546 happened 1 84 0.002475247524753 0.207920792079 begged 1 85 0.002475247524753 0.2103960396039604 spare 1 86 0.002475247524753 0.2103960396039604 spare 1 86 0.0024752475247524753 0.2103960396034 fife 1 87 0.002475247524753 0.2187128712871287 fife 1 88 0.0024752475247524753 0.21534653465346534 offered 1 88 0.0024752475247524753 0.2178217821782 return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.2227722772278 any 1 0.0024752475247524753 0.22524752475247525	one	1		0.0024/324/324/324/33	0.1030433043304
routine 1 77 0.002475247524753 0.190594059406 took 1 78 0.002475247524753 0.1930693069307 fishing 1 79 0.0024752475247524753 0.1930693069307 fishing 1 80 0.0024752475247524753 0.195544554455455 net 1 80 0.0024752475247524753 0.1980198019803 down 1 81 0.0024752475247524753 0.2004950495049505 something 1 82 0.002475247524753 0.200495049505 something 1 83 0.002475247524753 0.200495049505 something 1 83 0.002475247524753 0.200495049505 happened 1 84 0.002475247524753 0.2054455445546546 happened 1 84 0.002475247524753 0.207920792079 begged 1 85 0.002475247524753 0.2103960396039604 spare 1 86 0.002475247524753 0.2103960396039604 spare 1 86 0.0024752475247524753 0.2103960396034 fife 1 87 0.002475247524753 0.2187128712871287 fife 1 88 0.0024752475247524753 0.21534653465346534 offered 1 88 0.0024752475247524753 0.2178217821782 return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.2227722772278 any 1 0.0024752475247524753 0.22524752475247525	usual	1	76	0.0024752475247524753	0.18811881188118812
took 1 78 0.0024752475247524753 0.19306930693069307 fishing 1 79 0.0024752475247524753 0.19554455445544555 net 1 80 0.0024752475247524753 0.1980198019803 down 1 81 0.0024752475247524753 0.20049504950495049505 something 1 82 0.0024752475247524753 0.202970297029702978 unusual 1 83 0.0024752475247524753 0.20544554455445546 happened 1 84 0.0024752475247524753 0.2079207920792079 begged 1 85 0.0024752475247524753 0.2103960396039604 spare 1 86 0.0024752475247524753 0.2128712871287 life 1 87 0.0024752475247524753 0.21534653465346534 offered 1 88 0.0024752475247524753 0.2178217821782 return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.225772277227722772277					
fishing 1 79 0.0024752475247524753 0.195544554455455445555 net 1 80 0.0024752475247524753 0.19801980198019803 down 1 81 0.0024752475247524753 0.2004950495049505 something 1 82 0.002475247524753 0.20297029702970298 unusual 1 83 0.002475247524753 0.2054455445544566 happened 1 84 0.002475247524753 0.2079207920792079 begged 1 85 0.002475247524753 0.210396039604 spare 1 86 0.002475247524753 0.2128712871287 life 1 87 0.002475247524753 0.21534653465346534 offered 1 88 0.0024752475247524753 0.2178217821782 return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.22257722772277227722778 any 1 91 0.0024752475247524753 0.22554752475247525	routine	1	//	0.0024/324/324/324/33	0.1905940594059406
fishing 1 79 0.0024752475247524753 0.195544554455455445555 net 1 80 0.0024752475247524753 0.19801980198019803 down 1 81 0.0024752475247524753 0.2004950495049505 something 1 82 0.002475247524753 0.20297029702970298 unusual 1 83 0.002475247524753 0.2054455445544566 happened 1 84 0.002475247524753 0.2079207920792079 begged 1 85 0.002475247524753 0.210396039604 spare 1 86 0.002475247524753 0.2128712871287 life 1 87 0.002475247524753 0.21534653465346534 offered 1 88 0.0024752475247524753 0.2178217821782 return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.22257722772277227722778 any 1 91 0.0024752475247524753 0.22554752475247525	took	1	78	0.0024752475247524753	0.19306930693069307
net 1 80 0.0024752475247524753 0.19801980198019803 down 1 81 0.0024752475247524753 0.2004950495049505 something 1 82 0.0024752475247524753 0.20297029702970298 unusual 1 83 0.0024752475247524753 0.2054455445544566 happened 1 84 0.0024752475247524753 0.2079207920792079 begged 1 85 0.0024752475247524753 0.210396039604 spare 1 86 0.0024752475247524753 0.2128712871287 life 1 87 0.0024752475247524753 0.21534653465346534 offered 1 88 0.0024752475247524753 0.2178217821782 return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.222772277227722772278 any 1 91 0.0024752475247524753 0.22524752475247525					
down 1 81 0.0024752475247524753 0.2004950495049505 something 1 82 0.0024752475247524753 0.20297029702970298 unusual 1 83 0.0024752475247524753 0.2054455445544546 happened 1 84 0.002475247524753 0.2079207920792079 begged 1 85 0.002475247524753 0.210396039604 spare 1 86 0.002475247524753 0.2128712871287 life 1 87 0.002475247524753 0.21534653465346534 offered 1 88 0.0024752475247524753 0.2178217821782 return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.22277227722772278 any 1 91 0.0024752475247524753 0.22524752475247525	tisning	1	/9	0.0024/524/524/524/53	0.1955445544554
down 1 81 0.0024752475247524753 0.2004950495049505 something 1 82 0.0024752475247524753 0.20297029702970298 unusual 1 83 0.0024752475247524753 0.2054455445544546 happened 1 84 0.002475247524753 0.2079207920792079 begged 1 85 0.002475247524753 0.210396039604 spare 1 86 0.002475247524753 0.2128712871287 life 1 87 0.002475247524753 0.21534653465346534 offered 1 88 0.0024752475247524753 0.2178217821782 return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.22277227722772278 any 1 91 0.0024752475247524753 0.22524752475247525	net	1	80	0.0024752475247524753	0.19801980198019803
something 1 82 0.0024752475247524753 0.20297029702970298 unusual 1 83 0.002475247524753 0.20544554455445546 happened 1 84 0.002475247524753 0.2079207920792079 begged 1 85 0.002475247524753 0.210396039604 spare 1 86 0.002475247524753 0.2128712871287 life 1 87 0.002475247524753 0.21534653465346534 offered 1 88 0.0024752475247524753 0.2178217821782 return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.22277227722772278 any 1 91 0.0024752475247524753 0.22524752475247525					
unusual 1 83 0.0024752475247524753 0.205445544554456465465465465465465465465465	down	1	81	0.0024/524/5247524753	0.2004950495049505
unusual 1 83 0.0024752475247524753 0.205445544554456465465465465465465465465465	something	1	82	0.0024752475247524753	0.20297029702970298
happened 1 84 0.0024752475247524753 0.2079207920792079 begged 1 85 0.002475247524753 0.2103960396039604 spare 1 86 0.002475247524753 0.21287128712871287 life 1 87 0.002475247524753 0.21534653465346534 offered 1 88 0.002475247524753 0.21782178217821782 return 1 89 0.002475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.22277227722772278 any 1 91 0.0024752475247524753 0.22524752475247525					
begged 1 85 0.0024752475247524753 0.2103960396039604 spare 1 86 0.002475247524753 0.21287128712871287 life 1 87 0.002475247524753 0.21534653465346534 offered 1 88 0.0024752475247524753 0.21782178217821782 return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.22277227722772278 any 1 91 0.0024752475247524753 0.22554752475247525	unusual	1	83	0.0024/52475247524753	0.20544554455445546
begged 1 85 0.0024752475247524753 0.2103960396039604 spare 1 86 0.002475247524753 0.21287128712871287 life 1 87 0.002475247524753 0.21534653465346534 offered 1 88 0.0024752475247524753 0.21782178217821782 return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.22277227722772278 any 1 91 0.0024752475247524753 0.22554752475247525	hannened	1	84	0.0024752475247524753	0.2079207920792079
spare 1 86 0.0024752475247524753 0.21287128712871287 life 1 87 0.002475247524753 0.21534653465346534 offered 1 88 0.002475247524753 0.21782178217821782 return 1 89 0.002475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.22277227722778 any 1 91 0.0024752475247524753 0.22524752475247525		1			
spare 1 86 0.0024752475247524753 0.21287128712871287 life 1 87 0.002475247524753 0.21534653465346534 offered 1 88 0.002475247524753 0.21782178217821782 return 1 89 0.002475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.2227722772278 any 1 91 0.0024752475247524753 0.22524752475247525	begged	1	85	0.0024752475247524753	0.2103960396039604
life 1 87 0.0024752475247524753 0.21534653465346534 offered 1 88 0.0024752475247524753 0.21782178217821782 return 1 89 0.0024752475247524753 0.220297029703 grant 1 90 0.0024752475247524753 0.2227722772278 any 1 91 0.0024752475247524753 0.22524752475247525		1	86	0.0024752475247524753	0.21287128712871287
offered 1 88 0.0024752475247524753 0.21782178217821782 return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.22277227722778 any 1 91 0.0024752475247524753 0.22524752475247525		-			
return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.22277227722778 any 1 91 0.0024752475247524753 0.22524752475247525	life	1	87	0.0024752475247524753	0.21534653465346534
return 1 89 0.0024752475247524753 0.2202970297029703 grant 1 90 0.0024752475247524753 0.22277227722778 any 1 91 0.0024752475247524753 0.22524752475247525	offered	1	QQ.	0.0024752475247524752	0.21782178217921792
grant 1 90 0.002475247524753 0.2227722772278 any 1 91 0.002475247524753 0.2252475247525		1			
any 1 91 0.002475247524753 0.2252475247525			89	0.0024752475247524753	0.2202970297029703
any 1 91 0.002475247524753 0.2252475247525	return	1			
		_		0.0024752475247524753	0.22277227722772279
asked 1 92 0.002475247524753 0.2277227722773	grant	1	90		
52 0.0027/327/327/33 0.22//22//22//3	grant	1	90		
	grant any	1	90 91	0.0024752475247524753	0.22524752475247525

Слово	Сколько раз встречается	Место	Частота встречаемости	Место * Частота
nothing	1	93	0.0024752475247524753	0.2301980198019802
became	1	94	0.0024752475247524753	0.23267326732673269
irate	1	95	0.0024752475247524753	0.23514851485148516
related	1	96	0.0024752475247524753	0.2376237623762376
story	1	97	0.0024752475247524753	0.2400990099009901
him	1	98	0.0024752475247524753	0.24257425742574257
told	1	99	0.0024752475247524753	0.24504950495049505
wished	1	100	0.0024752475247524753	0.24752475247524752
found	1	101	0.0024752475247524753	0.25
fresh	1	102	0.0024752475247524753	0.2524752475247525
baked	1	103	0.0024752475247524753	0.25495049504950495
table	1	104	0.0024752475247524753	0.25742574257425743
then	1	105	0.0024752475247524753	0.2599009900990099
decided	1	106	0.0024752475247524753	0.2623762376237624
wanted	1	107	0.0024752475247524753	0.26485148514851486
more	1	108	0.0024752475247524753	0.26732673267326734
than	1	109	0.0024752475247524753	0.2698019801980198
just	1	110	0.0024752475247524753	0.2722772277227723
ask	1	111	0.0024752475247524753	0.2747524752475248
wasn	1	112	0.0024752475247524753	0.27722772277227725
t	1	113	0.0024752475247524753	0.27970297029702973
satisfied	1	114	0.0024752475247524753	0.2821782178217822
following	1	115	0.0024752475247524753	0.2846534653465347

Слово	Сколько раз встречается	Место	Частота встречаемости	Место * Частота
magic	1	116	0.0024752475247524753	0.2871287128712871
house	1	117	0.0024752475247524753	0.2896039603960396
ones	1	118	0.0024752475247524753	0.29207920792079206
before	1	119	0.0024752475247524753	0.29455445544554454
granted	1	120	0.0024752475247524753	0.297029702970297
again	1	121	0.0024752475247524753	0.2995049504950495
governor	1	122	0.0024752475247524753	0.30198019801980197
riches	1	123	0.0024752475247524753	0.30445544554455445
ordering	1	124	0.0024752475247524753	0.3069306930693069
about	1	125	0.0024752475247524753	0.3094059405940594
servants	1	126	0.0024752475247524753	0.3118811881188119
woman	1	127	0.0024752475247524753	0.31435643564356436
unhappy	1	128	0.0024752475247524753	0.31683168316831684
demanded	1	129	0.0024752475247524753	0.3193069306930693
eventually	1	130	0.0024752475247524753	0.3217821782178218
being	1	131	0.0024752475247524753	0.32425742574257427
satisfy	1	132	0.0024752475247524753	0.32673267326732675
last	1	133	0.0024752475247524753	0.3292079207920792
be	1	134	0.0024752475247524753	0.3316831683168317
ruler	1	135	0.0024752475247524753	0.3341584158415842
creatures	1	136	0.0024752475247524753	0.33663366336633666
who	1	137	0.0024752475247524753	0.33910891089108913
live	1	138	0.0024752475247524753	0.3415841584158416
made	1	139	0.0024752475247524753	0.3440594059405941
rags	1	140	0.0024752475247524753	0.34653465346534656
standing	1	141	0.0024752475247524753	0.349009900990099
by	1	142	0.0024752475247524753	0.35148514851485146
broken	1	143	0.0024752475247524753	0.35396039603960394
inside	1	144	0.0024752475247524753	0.3564356435643564
		145		



The Golden Fish

Once upon a time, in a land far far away lived a very poor couple in a shack not far from the edge of the sea. Their only means of food was the fish that the old man caught in the sea. One morning, as was his usual routine, the fisherman took his fishing net down to the sea. But on this day something unusual happened, on this day the fisherman caught the Golden Fish. The Golden Fish begged for the fisherman to spare his life, and offered in return to grant the fisherman any wish he would like. But the kindhearted fisherman asked for nothing, and returned the Golden Fish to the sea. However, the fisherman's wife was not so kindhearted, she became irate when he related the story to her, and sent him back to the sea to catch the Golden Fish and to wish for a loaf of bread. The fisherman did as he was told, he caught the fish and wished for a loaf of bread. When he returned home he found a fresh baked loaf of bread on the table.

The fisherman's wife then decided that she wanted more than just a loaf of bread. The next morning she sent her husband to ask for a new. He returned home to find his wife with a new washtub, but she still wasn't satisfied.

The following day the husband was sent to the sea to find the magic fish and to wish for a new house. This wish was, like the ones before it, granted to the fisherman. But the fisherman was sent back again the next day to wish that his wife would become governor. This time he returned home to find his wife dressed in riches and ordering about servants. But the woman was still unhappy, and demanded to become Queen of all the land.

Eventually, even being Queen of all the land did not satisfy the wife, and so she sent her husband once last time to the sea to catch the Golden Fish and to wish that she would be ruler of the sea and of all creatures who live in it. The fisherman caught the fish, and made the wish. However, when he returned home his wife was dressed in her old rags, standing by her old broken washtub, inside the old shack, with not even a loaf of bread to eat.