

Приложение D.

Задания дорожки контекстно- зависимого аннотирования

В этом приложении приведены запросов, использовавшихся для оценки результатов аннотирования. В качестве расширенных описаний запросов использовались описания, созданные в рамках РО-МИП'2004.

D.1 Задания по основе Веб-запросов.

1. Запрос *rbwa1932: Биография Куприна*
Описание: Страница должна содержать информацию о биографии русского писателя Куприна, где и когда родился, где провел детство, отрочество, его взгляды на жизнь, творчество.
2. Запрос *rbwa1933: биография Ленина*
Описание: Соответствующие запросу документы должны содержать информацию о личности Владимира Ильича Ульянова-Ленина, его жизни и политической деятельности. Особенно полезными считаются документы содержащие исследования на эту тему.
3. Запрос *rbwa1939: биография петра*
Описание: Необходимо найти информацию о биографии Петра I, об основных событиях его жизни, деятельности. Полезны страницы, содержащие исследования на эту тему.
4. Запрос *rbwa2001: блюда из кальмаров*
Описание: Необходимо найти рецепты холодных и горячих блюд, одним из основных ингредиентов которых, являются кальмары. Дополнительными плюсами является такая информация об этих блюдах, как их полезность, кулинарные особенности их приготовления.
5. Запрос *rbwa2038: болезни сердца*
Описание: Страница должна содержать подробное описание болезней сердца, их обзор или классификацию, описание заболеваний, связанных с нарушением работы сердца, их профилактики и лечения.
6. Запрос *rbwa2156: бухгалтерские программы*
Описание: Необходимо собрать информацию о доступных

бухгалтерских программах и их возможностях. Простое упоминание существования программы или реклама возможности разработки таких программ на заказ релевантны-ми не являются.

7. Запрос gbwa2490: *витамины в пище человека*
Описание: Искомая страница должна содержать информацию о витаминах, присутствующих в пище человека, их пользы и влиянии на здоровье человека, обзор тех продуктов питания, в которых они содержатся. Дополнительными плюсами является указание на то, сколько и каких витаминов приходится на единицу того или иного продукта.
8. Запрос gbwa2504: *вкусные сложные салаты*
Описание: Релевантная страница должна содержать рецепт хотя бы одного салата.
9. Запрос gbwa2775: *все про массаж*
Описание: Соответствующие запросу документы должны содержать информацию о различных техниках массажа. Лечебный, расслабляющий, укрепляющий... Наиболее релевантными являются документы, содержащие подробное описание техник массажа.
10. Запрос gbwa2812: *Вузы московской области*
Описание: Требуется составить список высших учебных учреждений в Московской области. Релевантными являются документы, содержащие название вуза и некоторые данные позволяющие его найти (ссылка на сайт, адрес (хотя бы город) или другая контактная информация).
11. Запрос gbwa3235: *горный университет*
Описание: Соответствующие запросу документы должны содержать подробную информацию о Горном Университете, истории университета, его месте нахождения и т.д. Особенно полезной является информация, связанная с обучением, условиями поступления, специальностями, кафедрами и т.д.
12. Запрос gbwa3669: *детские лагеря в Подмоскowie*
Описание: Страница должна содержать информацию о наличии и расположении детских лагерей в Подмоскowie, их статус и программа мероприятий, проводимых в них, каковы условия отправления ребенка в лагерь.
13. Запрос gbwa5749: *кластерный анализ*
Описание: Необходимо найти информацию о кластерном

анализе с подробным описанием этого метода, предназначения, применения его в различных научных областях.

14. Запрос gbwa5892: *код России*

Описание: Необходимо найти информацию о международных кодах России как на уровне страны, так и на уровне отдельных субъектов федерации. Полезной является информация, содержащая список/справочник телефонных кодов регионов или городов РФ, коды мобильных операторов России, штрих код РФ.

15. Запрос gbwa6896: *Любовные истории*

Описание: Соответствующая запросу страница может содержать как рассказы, связанные с любовными переживаниями, так и подробную информацию о группе "Любовные истории". В последнем случае полезной является информация об участниках группы, истории создания, исполняемых песнях, концертах. Отзывы фанатов группы не относятся к исчерпывающему ответу на запрос.

16. Запрос gbwa7240: *мерседес из германии*

Описание: Страница должна содержать информацию о местах, где можно купить автомобиль марки Мерседес в Германии с указанием стоимости и описанием технических данных этих машин.

17. Запрос gbwa7831: *Напиток богов амброзия*

Описание: Необходимо найти растолкование понятия "напиток богов" или "амброзия": история происхождения этого выражения, когда оно употребляется, в каком контексте, какую смысловую нагрузку оно несет.

18. Запрос gbwa8080: *нобелевская премия*

Описание: Соответствующая запросу страница должна содержать подробную информацию о нобелевской премии. Полезной является информация об истории создания премии, создателе, научных областях, по которым присуждается премия, ее лауреатах. Упоминание о присуждении нобелевской премии конкретному человеку не является исчерпывающим ответом на запрос.

19. Запрос gbwa8710: *Основные философские взгляды Аристотеля*

Описание: Страница должна содержать информацию об основных философских взглядах Аристотеля, его идеологии. Косвенные упоминания взглядов или философии Аристотеля не делают страницу релевантной.

20. Запрос gbwa9234: *Песнь о Нибелунгах*
Описание: Страница должна содержать информацию об эпическом произведении "Песнь о Нибелунгах". Полезной информацией считается краткое содержание, история создания, указание автора, персонажей, места действия, временной линии и т.д. Полезными страницами также являются страницы, содержащие критические статьи или рефераты на эту тему.
21. Запрос gbwa9700: *породы кошек классификация*
Описание: Страница должна содержать подробную классификацию пород кошек. Страницы, содержащие описание отдельных пород кошек, не являются релевантными.
22. Запрос gbwa9938: *предсказания будущего*
Описание: Страница должна содержать либо конкретные предсказания будущего, либо методы их получения (например, способы гадания), либо детальное обсуждение возможности делать предсказания будущего (например, анализ предсказаний Ванды). Страницы с предложениями услуг по предсказанию или перечнем предсказателей - не релевантны.
23. Запрос gbwa11225: *рецепты для микроволновой печи*
Описание: Требуется найти рецепты приготовления пищи для микроволновой печи. Релевантными являются документы, содержащие описание хотя бы одного рецепта приготовления пищи для микроволновой печи.
24. Запрос gbwa12347: *составление резюме*
Описание: Искомая страница должна содержать правила и рекомендации по составлению резюме на русском или английском языке. Примеры написания резюме, шаблоны или образцы считаются соответствующими запросу.
25. Запрос gbwa12638: *старинные русские праздники*
Описание: Необходимо найти обзор русских народных обрядов, праздников, обычаев, с подробным описанием каждого.
26. Запрос gbwa13149: *телевизоры, домашний кинотеатр*
Описание: Страница должна содержать информацию о телевизорах, подходящих для организации домашнего кинотеатра, о существующих на данный момент таких моделях, их характеристиках, стоимости, возможности приобретения, а так же о будущих разработках.

27. Запрос gbwa13402: *тональные средства*
Описание: Соответствующая запросу страница должна содержать подробную информацию о характеристиках тональных средств, инструкции по применению. Дополнительными плюсами являются указания, как правильно выбирать тональное средство в зависимости от различных факторов (например, типа кожи). Также полезными страницами являются страницы, посвященные обзорам и рейтингам тональных средств различных марок косметики.
28. Запрос gbwa13502: *Тринадцатый воин*
Описание: Релевантная страница должна содержать информацию о фильме "Тринадцатый воин" - например, историю создания, список актеров и т.п. Предложения купить касету, скачать mp3 или сходить в кинотеатр - не релевантны.
29. Запрос gbwa13761: *упражнения для улучшения зрения*
Описание: Соответствующие запросу документы должны содержать информацию о приемах и упражнениях для снятия усталости глаз, улучшения остроты зрения. Полезной также является информация о различных методиках по этой теме.
30. Запрос gbwa14193: *Формула 1*
Описание: Требуется найти информацию о чемпионате по автомобильным гонкам "Формула 1" - местах проведения гонок, правилах чемпионата. Страницы, содержащие информацию только об одном пилоте, одной модели болида и т.п. не являются релевантными из-за узкости фокуса.
31. Запрос gbwa15541: *что такое любовь*
Описание: Соответствующие запросу документы должны содержать информацию о таком философском понятии, как любовь. Полезной информацией являются рассуждения на эту тему, различные высказывания великих людей.
32. Запрос gbwa15582: *закон об обществах с ограниченной ответственностью*
Описание: Документы должны содержать информацию о Законе об обществах с ограниченной ответственностью, комментарии, поправки к нему. Особо важными являются документы содержащие текст Закона.
33. Запрос gbwa15631: *как накачать мышцы*
Описание: Соответствующие запросу документы должны содержать информацию о методиках увеличения мышечной

массы. Особое значение имеют документы с точными указаниями на методику и расписания занятий.

34. Запрос gbwa16729: *с чего начать бизнес*
Описание: Необходимо найти информацию о том как создать собственное дело. Полезной информацией являются набор шагов по созданию собственной фирмы, законодательные акты о регистрации предприятия, советы начинающим бизнесменам, советы по организации работы и менеджмента, кредитование бизнеса и т.д.
35. Запрос gbwa16982: *что взять в роддом*
Описание: Страница должна содержать информацию о том, что необходимо взять в роддом роженице или что необходимо, чтобы забрать ее при выписке из роддома. Полезная подробная информация о необходимых документах и вещах для роженицы и младенца.
36. Запрос gbwa17542: *что такое lol*
Описание: Документы должны содержать информацию о том, что означает и когда, в каком смысле применяется сокращение lol, которое обычно используется при обмене сообщениями в интернете: чатах, ICQ, IRQ, и т. д.
37. Запрос gbwa18064: *как приготовить пиццу*
Описание: Страница должна содержать рецепты приготовления пиццы с указанием ингредиентов и подробным описанием процесса приготовления. Дополнительным плюсом является сравнительный анализ пицц по принадлежности кухням различных стран, их основные различия и характеристики.
38. Запрос gbwa19044: *о порядке выезда из российской федерации и въезда в российскую федерацию*
Описание: Документы должны содержать информацию об условиях въезда и выезда из Российской Федерации. Особо важными являются документы, содержащие нормативные акты по этой теме (например, закон о регистрации иностранных граждан).
39. Запрос gbwa23612: *что такое доменное имя*
Описание: Искомые документы должны содержать информацию о таком понятии из теории информации, как доменное имя. Дополнительными плюсами является информация о доменах, регистрации ДИ. и т. д.

D.2 Задания на основе запросов из журналов ИПС по нормативным коллекциям.

1. Запрос rbla285: *акцизы*
2. Запрос rbla559: *бандитизм*
3. Запрос rbla589: *банкротство предпринимателей*
4. Запрос rbla1260: *федеральный закон о рекламе*
5. Запрос rbla1427: *функции государства*
6. Запрос rbla1598: *государственный внутренний долг*
7. Запрос rbla1719: *индексация ставки земельного налога*
8. Запрос rbla2194: *коллективные трудовые споры*
9. Запрос rbla2608: *льготная пенсия*
10. Запрос rbla2921: *надбавка за выслугу лет*
11. Запрос rbla2985: *налог на имущество*
12. Запрос rbla3575: *о государственной экспертной деятельности*
13. Запрос rbla3776: *о представлении отчетности аудиторских организаций*
14. Запрос rbla3779: *о применение контрольно-кассовых машин*
15. Запрос rbla3885: *о страховании гражданской ответственности авиаперевозчиков*
16. Запрос rbla3926: *о внесении изменений в уголовный кодекс*
17. Запрос rbla3979: *о членах коллегии федерального органа исполнительной власти*
18. Запрос rbla4455: *оптовая торговля*
19. Запрос rbla4625: *отпуск*
20. Запрос rbla4757: *памятники архитектуры*
21. Запрос rbla5046: *платные стоянки*
22. Запрос rbla5247: *положение о ведении общероссийских классификаторов*
23. Запрос rbla5927: *принципы международных коммерческих договоров*
24. Запрос rbla5941: *приобретение гражданства*

25. Запрос rbla5985: *приватизация жилья*
26. Запрос rbla6024: *признании отцовства*
27. Запрос rbla6055: *продажа акций*
28. Запрос rbla6497: *реализация путевок фонды социального страхования*
29. Запрос rbla6523: *реформа ЖКХ*
30. Запрос rbla6580: *регистрация предприятия*
31. Запрос rbla6731: *розничная продажа*
32. Запрос rbla6970: *социальное страхование+путевки*
33. Запрос rbla7501: *судебные издержки*
34. Запрос rbla7774: *транспортный налог*
35. Запрос rbla7889: *уклонение от заключения договора*
36. Запрос rbla8557: *выкуп земли*
37. Запрос rbla8576: *выплаты при рождении ребенка*
38. Запрос rbla8727: *закон о переписи населения#57-ФЗ*
39. Запрос rbla8731: *закон о предпринимательской деятельности*
40. Запрос rbla8741: *закон о статусе судьи*
41. Запрос rbla9260: *Бартерные сделки*
42. Запрос rbla10010: *Конвенция рыболовство*
43. Запрос rbla10490: *О государственной поддержке молодежных и детских общественных объединений*
44. Запрос rbla10857: *О валютном регулировании и валютном контроле*
45. Запрос rbla10882: *О внесении изменений и дополнений в Уголовный кодекс РФ*
46. Запрос rbla10934: *О защите прав иностранных инвесторов*
47. Запрос rbla11054: *Об ограничении курения табака и потребления табачных изделий.*
48. Запрос rbla11085: *Об освобождении от налога на добавленную стоимость*
49. Запрос rbla11189: *Об учебно-методических объединениях*

50. Запрос rbla11409: *Перечень производств, профессий и работ с тяжелыми и вредными, особотяжелыми и особо вредными условиями труда предприятий и организаций*
51. Запрос rbla11499: *Положение о паспорте гражданина РФ*
52. Запрос rbla11506: *Положение о порядке присвоения ученых званий*
53. Запрос rbla11509: *Положение о порядке сдачи квалификационного экзамена и оценки знаний претендентов*
54. Запрос rbla11565: *Понятие банковская тайна*
55. Запрос rbla12595: *Учредительные документы общества с ограниченной ответственностью*
56. Запрос rbla12742: *Закон о едином налоге*