1

Сейчас работает модуль, который данные выбирает с диска, разбирает и формирует контекстные ключевые слова(сначала из текста статьи, затем расширяет список из словарей).

Доработать модуль, сохранив в бд MySql или еще куда-нибудь следующие данные (сделай пока бд как у Леоновой приблизительно):

- ФИО научного сотрудника;

- Труд/труды, опубликованные под именем данного сотрудника;

- Ссылки на литературу, использованную в данных научных трудах;

- Ключевые слова к каждому научному труду данного сотрудника;

- Ученая степень научного сотрудника;

- Ученое звание научного сотрудника;

- Контактная информация (в случае существования в открытом доступе).

О статьях храним следующие данные:

На рисунке 7 приведена модель “Документ”(7), которая позволяет описать различные виды документов в деятельности организаций (см. табл. 7):

|  |  |
| --- | --- |
| *Doc= < Ab, Art, AcArt, ConPap, EduArt, RewArt, PartDoc, UDC, Cont, Ref, Abstract, Author, KW, Link, Par, Th, Title, MatB>* | *(7)* |
|  |  |

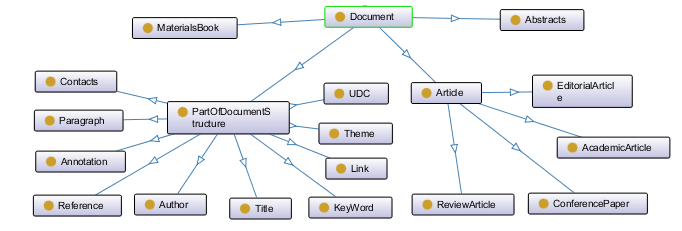


Рисунок 7- Модель “Документ”, разработанная в Protégé

Таблица 7. Классы модели “Документ”

|  |  |  |  |
| --- | --- | --- | --- |
| Параметр  модели | Имя класса | Класс в стандартной онтологии | Описание класса |
| *Ab* | Abstracts | bibo:Thesis | Тезисы |
| *Art* | Article | bibo:Article | Статья |
| *AcArt* | AcademicArticle | bibo:AcademicArticle | Академическая статья |
| *ConPap* | ConferencePaper | vivo:ConferencePaper | Статья на конференции |
| *EduArt* | EditorialArticle | vivo:EditorialArticle | Статья научно-популярная |
| *RewArt* | ReviewArticle | vivo:ReviewArticle | Обзорная статья |
| *PartDoc* | PartOfDocumentStructute | bibo:DocumentPart | Структура документа по частям |
| UDC | UDC |  | УДК |
| Cont | Contacts |  | Контакты |
| Ref | Reference | bibo:ReferenceSource | Единица списка литературы |
| Abstract | Abstract | dbo:Abstract | Аннотация |
| Author | Author |  | Автор |
| KW | KeyWord |  | Ключевое слово |
| Link | Link |  | Ссылка |
| Par | Paragraph |  | Параграф |
| Th | Theme |  | Тема документа |
| Title | Title |  | Заголовок |
| MatB | MaterialsBook | bibo:Collection | Сборник |

На рисунке 8 приведена модель “Образовательный ресурс"(8), которая позволяет описать научные мероприятия и события в деятельности организаций (см. табл. 8):

|  |  |
| --- | --- |
| *EduR = < Boo, Lec, Man, Per, Mag, NewP, Rep, GrRep, LabRep, PrRep, StResRep, TestRep, Slide, ManLR, ManKR, ManSR, Disc, Doc>* | *(8)* |
|  |  |

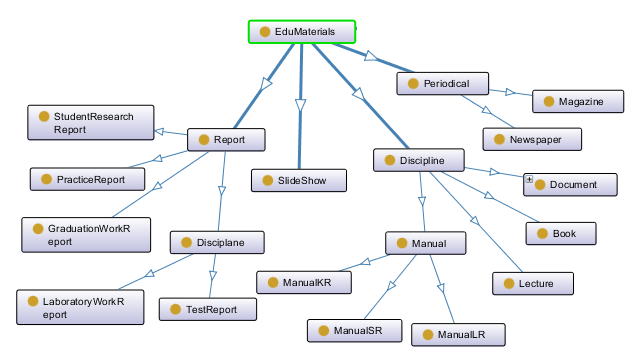


Рисунок 8- Модель “Образовательный ресурс", разработанная в Protégé

Таблица 8. Классы модели “Образовательный ресурс"

|  |  |  |  |
| --- | --- | --- | --- |
| Параметр  модели | Имя класса | Класс в стандартной онтологии | Описание класса |
| *Boo* | Book | bibo:Book | Книга |
| *Lec* | Lecture |  | Лекция |
| *Man* | Manual | bibo:Manual | Методичка |
| *Per* | Periodical | bibo:periodicl | Периодическое издание |
| *Mag* | Magazine | bibo:magazine | Журнал |
| *NewP* | Newspaper | bibo:journal | Газета |
| *Rep* | Report | bibo:Report | Отчет |
| *GrRep* | GraduationWorkReport |  | Отчет по ВКР |
| *LabRep* | LaboratoryWorkReport |  | Отчет по ЛР |
| *PrRep* | PracticeReport |  | Отчет по практике |
| *StResRep* | StudentResearchReport |  | Отчет по НИРС |
| *TestRep* | TestReport |  | Отчет по контрольной/тесту |
| *Slide* | SlideShow |  | Презентация |
| *ManLR* | ManualLR | bibo:Manual | Методичка по лабораторным работам |
| *ManKR* | ManualKR | bibo:Manual | Методичка по курсовой |
| *ManSR* | ManualSR | bibo:Manual | Методичка по самостоятельной работе |
| *Disc* | Discipline |  | Дисциплина (информатика, математика и т.д.) |
| *Doc* | Document | bibo:Document | Абстракция Документ |

2

Теперь надо сделать похожий кусок, только с поиском статей в интернет (паук начинает поиск с сайтов elibrary, google , известны даже конкретные страницы пользователей) и потом поиск с помощью поисковых машин, но сильно не углубляйся! Только надо сделать возможность корректировки этих ссылок в интерфейсе.

**Google Scholar**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| № п.п. | Фамилия Имя Отчество | Должность | Число публикаций автора | Число цитирований  (Все) | Индекс Хирша |
| 1 | Андриевская Наталия Климовна  <https://scholar.google.com.ua/citations?hl=ru&user=9aZ3OTcAAAAJ> | Ст. пр | 13 | 4 | 1 |
| 2 | Васяева Татьяна Александровна  [https://scholar.google.com.ua/citations?hl=ru&user=ue6WbGAAAAAJ](https://scholar.google.com.ua/citations?hl=ru&user=ue6WbGAAAAAJ%20%20%20) | доцент | 44 | 15 | 2 |
| 3 | Землянская Светлана Юрьевна  https://scholar.google.com.ua/citations?user=0bE2ZxkAAAAJ&hl=ru | Доцент | 37 | 20 | 3 |
| 4 | Мартыненко Татьяна Владимировна  <https://scholar.google.com.ua/citations?user=85KTpG0AAAAJ&hl=ru> | доцент | 31 | 18 | 2 |
| 5 | Новиков Денис Дмитриевич  <https://scholar.google.com.ua/citations?user=8qmUN_QAAAAJ&hl=ru> | ас | 1 | 0 | 0 |
| 6 | Поляков Александр Иванович  <https://scholar.google.com.ua/citations?user=B2ROfv4AAAAJ&hl=ru> | ст.пр. | 5 | 0 | 0 |
| 8 | Савкова Елена Осиповна  https://scholar.google.com.ua/citations?user=XbvDmOwAAAAJ&hl=ru | доцент | 46 | 36 | 4 |
| 9 | Светличная Виктория Антоновна  <https://scholar.google.com.ua/citations?hl=ru&user=j8ogF-IAAAAJ> | доцент | 40 | 17 | 3 |
| 10 | Секирин Александр Иванович  <https://scholar.google.ru/citations?hl=ru&user=2l0OyB4AAAAJ> | доцент | 39 | 58 | 5 |
| 12 | Теплова Ольга Валентиновна  <https://scholar.google.com.ua/citations?user=vaZfgWoAAAAJ&hl=ru> | ст.пр | 1 | 0 | 0 |
| 13 | Хмелевой Сергей Владимирович  <https://scholar.google.com.ua/citations?hl=ru&user=_RBRM3kAAAAJ> | Доцент | 28 | 14 | 3 |
| 16 | Воронова Алена Игоревна  [https://scholar.google.com.ua/citations?hl=ru&user=XViRSOMAAAAJ](https://clck.yandex.ru/redir/dv/*data=url%3Dhttps%253A%252F%252Fscholar.google.com.ua%252Fcitations%253Fhl%253Dru%2526user%253DXViRSOMAAAAJ%26ts%3D1476959563%26uid%3D1200653001455202488&sign=32f9e4a4ebbeb56f2bd056f8339f92d2&keyno=1) | ас | 2 | 0 | 0 |
| 17 | Пряхин Владимир Викторович  [https://scholar.google.com.ua/citations?hl=ru&user=hAhFNCoAAAAJ&view\_op=list\_works](https://clck.yandex.ru/redir/nWO_r1F33ck?data=NnBZTWRhdFZKOHRaTENSMFc4S0VQSXZfNUZYSkdfdG9uSmt5MGlzX0Q1QjBhNHZNN0RCbXlaXzloLXdZRmNHSWljSnphcEltcmZ1aTY5UTB1b0x4MXFyUkJYUWFOMzRQejlDSWtYdWZ6RVhyQ0FnOUNTbkp4Zm5yWUJuRlFibXB5V2JDTWhIM1UyT2VUcGhWbENOUnZhVHpYUnViYkkwWFF3dHpEcjBtYVcxN2RrNk5hSmY5MUt4NThEZmZhVUlTajJMZlpKZkFaVDQ&b64e=2&sign=e4c3b6bda63022b6786e5a537414d8ca&keyno=17) | ас | 1 | 0 | 0 |
|  |  |  | 567 | 755 |  |

**Science Index (РИНЦ)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| № п.п. | Фамилия Имя Отчество | Должность | **SPIN-код** | Число публикаций автора с учетом статей, найденных в списках литературы | Суммарное число цитирований автора | Индекс Хирша |
| 1 | Мартыненко Татьяна Владимировна  <http://elibrary.ru/author_profile.asp?authorid=848652> | доцент | 5656-3268 | 7 | 7 | 1 |
| 2 | Секирин Александр Иванович  <http://elibrary.ru/author_profile.asp?id=848805> | доцент | 1158-6791 | 8 | 18 | 2 |
| 3 | Светличная Виктория Антоновна  <http://elibrary.ru/author_profile.asp?id=848421> | доцент | 8377-9145 | 8 | 10 | 1 |
| 4 | Андриевская Наталия Климовна  <http://elibrary.ru/author_profile.asp?id=845456> | Ст. преп. | 4781-6958 | 2 | 1 | 0 |
| 6 | Поляков Александр Иванович  <http://elibrary.ru/author_profile.asp?id=412195> | ст.пр. | 3698-4878 | 0 | 0 | 0 |
| 7 | Савкова Елена Осиповна  http://elibrary.ru/author\_profile.asp?id=850798 | доцент | 4738-9733 | 11 | 26 | 1 |
| 8 | Землянская Светлана Юрьевна  http://elibrary.ru/author\_profile.asp?id=849754 | доцент | 2243-9770 | 10 | 11 | 1 |
| 9 | Теплова Ольга Валентиновна | ст.пр. | 5583-3298 | 0 | 0 | 0 |
| 10 | Новиков Денис Дмитриевич  <http://elibrary.ru/author_profile.asp?id=851224> | ас | 9470-2355 | 0 | 0 | 0 |
| 11 | Хмелевой Сергей Владимирович  <http://elibrary.ru/author_profile.asp?id=852981> | доцент | 6687-2560 | 5 | 3 | 1 |
| 12 | Васяева Татьяна Александровна  http://elibrary.ru/author\_profile.asp?id=852371 | доцент | 1372-4940 | 5 | 4 | 1 |
| 16 | Воронова Алена Игоревна | ас |  | 0 | 0 | 0 |
| 17 | Пряхин Владимир Викторович | ас |  | 0 | 0 | 0 |
|  |  |  |  | 59 | 92 |  |

**SciVerse Scopus**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| № п.п. | Фамилия Имя Отчество | Должность | **Scopus**  **Author ID** | Число публикаций автора | Число цитирований | Индекс Хирша |
| 1 | Андриевская Наталия Климовна  http://www.scopus.com/authid/detail.uri?authorId=57031846000 | Ст. пр. | 57031846000 | 1 | 0 | 0 |
| 2 | Васяева Татьяна Александровна  <http://www.scopus.com/authid/detail.uri?authorId=57188874934> | доцент | 57188874934 | 1 | 0 | 0 |
| 3 | Землянская Светлана Юрьевна  http://www.scopus.com/authid/detail.uri?authorId=57189025634 | доцент | 57189025634 | 1 | 0 | 0 |
| 4 | Мартыненко Татьяна Владимировна  http://www.scopus.com/authid/detail.uri?authorId=57188878717 | доцент | 57188878717 | 1 | 0 | 0 |
| 8 | Савкова Елена Осиповна  <http://www.scopus.com/authid/detail.uri?authorId=6504456802>  <http://www.scopus.com/authid/detail.uri?authorId=57197776034> | доцент | - | 2 | 0 | 0 |
| 9 | Светличная Виктория Антоновна | доцент | - | 0 | 0 | 0 |
| 10 | Секирин Александр Иванович  <http://www.scopus.com/authid/detail.uri?authorId=57189030062>  <http://www.scopus.com/authid/detail.uri?authorId=57188877026> | доцент | 57188877026,  57189030062 | 2 | 0 | 0 |
| 13 | Хмелевой Сергей Владимирович  http://www.scopus.com/authid/detail.uri?authorId=57188879888 | доцент | 57188879888 | 1 | 0 | 0 |
|  |  |  |  | 23 | 3 |  |

Разобрать то, что там есть: получить автора **Author ID** Число публикаций Число цитирований Индекс Хирша автора,

кол страниц, Аннотацию, название. сборник и тд

и сформировать список ключевых слов по названию и тексту и аннотации, также расширить его терминами из словаря. Организовать хранение, как и для первого модуля.

3

Организовать интеллектуальный поиск по своему хранилищу для вывода материалов пользователю.

1 Поиск по категориям

2 Поиск по контекстным словам (с расширением из словаря)

4

Аутентиф и регистрация. Внедрить все модули в единую систему.

5

Модуль классификации элементов

Когда список слов известен, необходимо определить, куда документ помещать. Для примера возьмем следующие тематические рубрики(для пробы несколько не близких 4-5 рубрик хватит я думаю):

1. языки программирования
2. тестирование программ
3. Безопасность программ и данных
4. технологии разработки ПО
5. Системный анализ
6. Современные инструменты типа CASE
7. Программирование мобильных систем
8. Технологии программирования на платформах .Net и Java
9. Web-программирование
10. Моделирование систем
11. Базы данных
12. Операционные системы
13. Компьютерные сети
14. Методы и системы искусственного интеллекта
15. Нейронные сети
16. Нечёткие модели и «мягкие» вычисления
17. Методы обработки изображений и распознавание образов
18. Машинное обучение и интеллектуальный анализ данных
19. Технологии обработки больших данных (Data Mining и Big Data)

Каждая из рубрик к этому моменту уже имеет свой багаж терминов. Надо выбрать наиболее семантически близкую рубрику(описать методы, основанные на исследовании текста, и попробовать несколько, чтобы выбрать лучший!).

Вначале, пока данных еще нет, надо предусмотреть возможность админу ввести и получить из словарей начальный набор терминов, т.е. должен поддерживаться полуавтоматический режим.

И при помещении статьи в категорию предусмотреть диалог с пользователем(вывести рубрики в порядке ранга, дополнить авторской рубрикой(по фамилии) поставить птички (можно дублировать в разные подходящие категории) и только потом сохранять.