

СОФИЙСКИ
УНИВЕРСИТЕТ



„СВ. КЛИМЕНТ
ОХРИДСКИ“
ОСНОВАН 1888 г.

СОФИЙСКИ УНИВЕРСИТЕТ

„СВ. КЛИМЕНТ ОХРИДСКИ “

ФАКУЛТЕТ ПО МАТЕМАТИКА И ИНФОРМАТИКА

Автореферат

на дисертация на
тема

УСЛУГИ НА ИЗСЛЕДОВАТЕЛСКИ ИНФОРМАЦИОННИ
СИСТЕМИ ЗА ОБЛАК ЗА ОТВОРЕНА НАУКА

Научен труд за присъждане на образователна и научна степен “доктор”
в професионално направление "4.6 Информатика и компютърни науки",
докторската програма "Информационни технологии" – Компютърни мрежи

Докторант:

СИЛВЕСТЕР ХАСАНИ

Научен ръководител:

Проф. д-р ЕЛИЗА СТЕФАНОВА

В резюме

Отворената наука е движение, което прави научните изследвания и тяхното разпространение достъпни за всички нива на нашето общество, било то любителско или професионално. Системата за Отворената наука се смята за бъдещето на научните изследвания и всички изследователски институции преминават към нейното използване.

В Европа Европейският облак за отворена наука насърчава този преход, като създава набор от правила и насоки, които трябва да се следват, за да направят данните за научните изследвания достъпни за всички изследователи от ЕС чрез оперативно съвместими данни FAIR. От всички европейски страни, преминаващи към Отворена наука, балканските страни се присъединяват съвсем наскоро към този преход. Това важи особено за албанските институции.

В тази дисертация Отворената наука е анализирана подробно. Започвайки от това какво представлява Отворена наука и достигайки до Отворен достъп са обяснени всички ключови компоненти на Отворената наука. Отворените данни са описани и сравнени с FAIR данните. Описан е и софтуерът с отворен код, а концепцията за цифровите хранилища е подробно представена като ключов елемент при съхранението на данни от изследванията. Съхранението на всички данни от изследванията е една от важните задачи на Отворената наука.

След това се въвежда европейският възглед за Отворената наука заедно с Европейският облак за Отворена наука (EOSC), преминавайки към най-използваните цифрови хранилища на изследователските данни. EOSC определя насоките, които трябва да се следват от научните организации в Европа, за да бъдат техните научни резултати съвместими за бъдещо сътрудничество в Open Science. Необходимо е пълно разбиране и съответствие към принципите и стандартите на EOSC, за да имаме реално приложение на Open Science в Европа. В резултат на това се установява настоящото състояние за преминаване към Отворената наука в балканските страни, като се акцентира върху идентифицирането на етапите в прехода към Отворена наука и основните проблеми, пред които са изправени балканските университети. За да се направите това, е направено проучване, като е изпратен въпросник до научни

институции и резултатите от него са анализирани подробно. Открити са ключови проблемни области и са планирани следващите стъпки.

Внимателният анализ илюстрира необходимостта от Албански облак за Отворена наука (AOSC). Целта на AOSC е да направи албанската наука отворена и да помогне на Албания да се присъедини към инициативата на EOSC. Представен е прототип за внедряване в Албания, следвайки европейските стандарти, определени от Европейският облак за Отворена наука. Албанският CRIS е хранилището, което ще помогне за изграждането на Албански облак за Отворена наука и за проследяване на прехода към Отворена наука и за присъединяване ѝ към EOSC. Структурата на данните в хранилището е илюстрирана. Анализът също така показва необходимостта от Отворена научна политика, която да се прилага в албанските университети. Политиката е представена и се взема в предвид от албанските изследователски институции за изпълнение. Политиката за отворен достъп цели прехода към Отворената наука и го вижда като ключов елемент за увеличаване на въздействието на научните изследвания върху албанската общност. Крайната цел на първата политика в областта на Отворената наука и дигиталното хранилище с отворен достъп в Албания е да се сдобие с данни от изследвания за постъпването им в FAIR и да се улесни достъпа до тези данни от всички албански изследователи.

В заключение, тази дисертация описва подробно преходът на албанските университети в Отворената наука и следващите стъпки, предприети за насърчаване на бъдещето на изследванията.

Съдържание

В резюме	2
Съдържание	4
Списък на фигурите	6
Списък с таблици	7
Глава 1. Въведение	8
1.1. Отворена наука	8
1.2. Отворен достъп	9
1.3. Отворени данни	10
1.4. FAIR данни	12
1.5. Софтуер с отворен код	14
1.6. Цифрови хранилища и управлението им	15
1.7. Цели на дисертацията	16
1.8. Публикации, върху които се базира дисертацията	17
1.9. Структура на дисертацията	18
Глава 2. Състояние на отворената наука в Европа.....	20
2.1. Европейски облак за отворена наука	20
2.2. Дигитални библиотеки.....	24
2.3. Хранилища за управление на научни данни.....	26
2.3.1. DSpace-CRIS.....	26
2.3.2. Zenodo -Invenio.....	28
2.3.3. Fedora.....	29
2.4. Изводи	30
Глава 3. Текущото състояние на Отворената наука на Балканите	32
3.1. Състояние	32
3.2. Методология	33
3.3. Изводи	35

Глава 4. Албанско цифрово хранилище за Отворена наука	38
4.1. Анализ на изискванията	38
4.2. Инсталиране на DSpace-CRIS.....	42
4.4. Модел на данните, внедрен в Албанския-CRIS.....	44
4.5. Политика за Отворена научна	46
4.6. Изводи	46
Глава 5. Заключение и бъдещи научни разработки	48
5.1. Изводи	48
5.2. Приноси	50
5.2.1. Научни приноси	50
5.2.2. Приложни приноси.....	50
5.2.3. Научноприложни приноси.....	50
5.3. Научни публикации, свързани с дисертационния труд	51
5.4. Бъдещо развитие	52
Библиография	53

Списък на фигурите

Фигура 2-а: Предложени различни модели EOSC	22
Фигура 2-б: Действия на федеративен модел на EOSC	23
Фигура 2-с: Демонстрация на интерфейса DSpace –CRIS	27
Фигура 3-а: Профили на респондентите в България	35
Фигура 4-а: албански DSpace -CRIS.....	39
Фигура 4-б: Пример от списъка на университетите в албанското цифрово хранилище с отворен достъп	42
Фигура 4-с: ORCID интеграция и вход	43
Фигура 4-д: Влезте с ORCID	43
Фигура 4-е: Албански -CRIS основната структура.....	45

Списък с таблици

Таблица 1: Сравнение на хранилища с отворен код.....	30
Таблица 2:Списък на институциите в Албания	39

Глава 1. Въведение

В тази глава е направен обзор на концепцията за Отворена наука, както и на нейните два основни елемента - Отворен достъп и Отворени данни. Въведени са определения за най-важните теми, свързани с Отворената наука и е дискутирана задълбочено всяка една от тях.

1.1. Отворена наука

Отворената наука напоследък привлича голямо внимание. Отворената наука е движението да се направят научните изследвания (включително публикации, данни, физически проби и софтуер) и разпространението им достъпни за всички нива на нашето общество, любителско или професионално [1]. Тя се счита и за наука, която е прозрачна и достъпна, която се споделя и развива чрез съвместни мрежи [2].

Отворената наука се състои от шест принципа: отворени данни, софтуер с отворен код, отворена методология, отворено рецензиране, отворен достъп и отворени образователни ресурси. През годините бяха видими няколко предимства и недостатъци на Отворената наука /ОН/, както е показано в [3]. Предимствата на ОН включват: Строга взаимна проверка, тъй като публикуваните доклади са на свободен достъп и данните от изследванията са публично финансирани, по-прозрачни, възпроизводими, и въздействащи върху наука. Недостатъците включват: потенциална злоупотреба, общо неразбиране от страна на обществеността, и също така нискокачествена наука, тъй като основната грижа е просто тя да бъде отворена.

Като част от движението Отворена наука, Отворения достъп (ОД) и изследвания за управление на данните (RDM) също са много важни теми за разбиране. Ключова стъпка за направата на науката отворена и прозрачна е отвореният достъп до научноизследователски данни, известни още като отворени данни. Тъй като изследванията са все повече управлявани от процеса на събиране на данни, напредъкът

в научните познания става все по-тясно свързан с тяхното добиване. Политиката за отворени данни дава възможност на изследователите да използват съществуващите знания свързани с иновациите по лесно достъпен начин [3]. В следващите раздели са обяснени понятията отворен достъп и отворени данни.

1.2. Отворен достъп

Както вече беше споменато по-горе, отвореният достъп е много важен за Отворената наука. Дефиницията за отворения достъп е набор от принципи и практики, чрез които резултатите от изследванията се разпространяват онлайн, безплатно или без други бариери за достъп. С отворения достъп се прилагат и отворени лицензи за авторски права [4], чрез които се избягват или отпадат напълно строго определените бариери за защита на интелектуалния труд и повторната употреба на данни. Свободният достъп до данни се определя също така като широко международно движение, чиято основна цел е да предостави безплатен и отворен онлайн достъп до отворената академична информация, като публикации и данни от направените изследвания. Докато някои публикации, могат също така да бъдат публикувани по различен начин, публикациите с отворен достъп са определени като публикации, към които няма финансови, правни, или технически пречки за достъп до тях. Това означава, че всеки може да чете, изтегля, копира, разпространява, отпечатва, търси и издирва информацията, която го интересува, или да я използва във връзка с образованието или по друг начин, но само в рамките на правните споразумения. Когато използването на данните е свързано с научните изследвания, отвореният достъп се явява, като публикуваща платформа за научна комуникация, която прави научноизследователската информация достъпна за читателите безплатно, за разлика от традиционния абонаментен модел, при който читателите имат достъп до научната информация, само при заплащане на абонамент [5].

Политиката за отворен достъп е приета от изследователски институции, изследователски спонсори или правителства, която изисква всички изследователи, част от организацията, да предоставят с отворен достъп своите публикации, статии от списания и доклади от конференции, по един от следните два начина. Първият начин е чрез самостоятелно архивиране на техните окончателни изследвания и трудове в

свободно достъпно институционално хранилище или хранилище, което се нарича МЕТОД НА ЗЕЛЕНИЯ ОТВОРЕН ДОСТЪП. Вторият начин е чрез публикуването им в списание с отворен достъп, наречено ЗЛАТЕН ОТВОРЕН ДОСТЪП. Понякога се използват и двата начина.

Институциите, които поддържат Отворената наука и Отворения достъп, имат политика на отворен достъп. Това е политика, приета от научноизследователски институции, спонсори на изследвания или правителства, които възлагат направата на изследвания, които обикновено включват университетски преподаватели, научноизследователски персонал или получатели на изследователско финансиране, за да направят своите публикации, рецензирани статии в списания и доклади от конференции в отворен достъп чрез метода на самостоятелното архивиране на техните окончателни изследвания и проекти в свободно достъпно институционално хранилище или хранилището на Зеления път (Green Route) или чрез публикуването им в списание с отворен достъп Златен път (Gold Route).

1.3. Отворени данни

Отворените данни са основен компонент на Отворената наука. Определението за отворени данни е данни и съдържание, които могат да бъдат свободно използвани, модифицирани и споделяни от всеки за всякакви цели [6]. Когато тази информация е полезна, използваема и употребявана Отворените данни се считат за Отворено знание. Основните свойства на Отворените данни са:

- Данните трябва да са достъпни като цяло и на не повече от разумната им цена за възпроизвеждане, достъпът е за предпочитане чрез изтегляне през интернет. Данните също трябва да бъдат достъпни в удобна и в преобразуваща се форма.
- Данните трябва да бъдат предоставени при условия, които позволяват повторна употреба и преразпределение, включително смесването с други набори от данни. Те трябва също да могат да се четат машинно.

- Данните може да са достъпни за всички. Те трябва да могат да се използват еднократно и многократно, както и да се преразпределят. Не е желателно да се проявява дискриминация спрямо областите, в които ще се използва, както и спрямо отделни лица или групи.

Има много причини, поради които искаме данните да бъдат отворени, но според Фондация „Отворени знания“/Open Knowledge Foundation/ основните причини са [7]:

- Прозрачност: Данните трябва да бъдат отворени за достъп за всички желаещи, трябва да бъдат споделени с всички и трябва да могат да се използват повторно във всякаква форма необходима за това.
- Осъзната социална и търговска стойност: Отворените данни могат да помогнат за създаването на иновативни бизнеси и услуги, които да доведат до тяхната социална и търговска стойност.
- Участие и ангажираност: данните не само трябва да бъдат прозрачни, но също така трябва да мотивират правителството, бизнеса и организациите да отворят своите данни за всички. По този начин могат да се вземат по-добри решения.

От друга страна, данните трябва да бъдат не само отворени, но и готови за многократно използване, за което се внедряват някои стандарти. Отвореният стандарт за данни е набор от спецификации и изисквания за това как някои данни трябва да бъдат публично достъпни. Тези стандарти трябва да бъдат отворени към приноса на всички и включват стандарти за:

- Схематичност: Схематичните стандарти определят структурата на данните, които ще бъдат публикувани. Това включва имената, описанията и типовете данни на полета или колони с данни. Схематичните стандарти също могат да включват как един набор от данни е свързан с друг.
- Семантичност: Семантичните стандарти определят терминологията или езика в публикуваните данни.
- Атомарност: Атомарните стандарти определят как трябва да бъдат представени основните елементи на данните. Атомарните стандарти могат да представляват отделни стойности на данни или комбинация от стойности на данните.

1.4. FAIR данни

Една от основните характеристики на Отворените данни се нарича FAIR. FAIR са данни, които могат да бъдат намерени, могат да са достъпни, съвместими са и могат да се използват многократно. Всички Отворени данни са FAIR данни. Понякога се случва, че някои FAIR данни, не са отворени, затова се казва че не всички FAIR данни са отворени данни.

Данните, които са FAIR, следват принципите на FAIR, които са обяснени по -долу [8]:

- Откриваемост:
 - F1. (Мета) данните имат глобален уникален и постоянен идентификатор (ID). (Metadata - Метаданните са „данни, които предоставят информация за други данни“)
 - F2. Данните са добре описани от метаданните.
 - F3. (Мета) данните трябва да бъдат регистрирани или индексирани като ресурс за търсене.
 - F4. Метаданните ясно посочват идентификатора на данните (ID).
- Достъпност:
 - A1. Възможно е да се получат (мета) данни чрез идентификатор (ID), използвайки стандартизиран комуникационен протокол.
 - A1.1 Протоколът е отворен за обществеността, безплатен и неограничен в използването.
 - A1.2 Протоколът трябва да може да предоставя методи за идентификация и оторизация, ако е необходимо.
 - A2. Метаданните трябва да са на разположение, дори ако данните станат недостъпни.

- Оперативна съвместимост:
 - I1. Използвайте добре дефиниран, достъпен, споделен и широко приложим език за описание и представяне на научните (мета) данни.
 - I2. (Мета) данните използват речник, който следва принципа на FAIR.
 - I3. (Мета) данните трябва да съдържат идентифицируема справочна информация към други (мета) данни
- Принцип за многократна употреба:
 - R1. Мета (данните) имат изобилие от точно свързани атрибути.
 - R1.1 (мета) данните се публикуват с ясен и достъпен лиценз за използване на данните.
 - R1.2 (мета) данните са свързани с първоизточника си.
 - R1.3 (мета) данните отговарят на общностните стандарти за всяка дисциплина.

Има две концепции, свързани с FAIR данните: Метаданните и глобално уникален и постоянен идентификатор. Метаданните се дефинират като данни, които предоставят информация за други данни, така че това са данни за данните [9]. Самите метаданни имат много функции, те спомагат за организирането на данните, осигуряват цифрова идентификация, архивиране и съхранение. Те също така помагат при откриването на подходящи сходни източници, както и за разграничаване на данни, които не са еднакви.

Глобалият уникален и постоянен идентификатор е връзката с документ, файл или друг обект, който е уникален в световен мащаб и не е обвързан с нито една институция. В Open Science ORCID е световно използван глобално уникален и постоянен идентификатор. ORCID е идентификатор на изследовател. Идентификаторът представлява буквено -цифров код, който уникално идентифицира научни и академични автори, като същевременно следва принципите на отворения достъп, когато е непатентован код [10].

В Open Science данните трябва да бъдат не само отворени, но и FAIR. Това означава, че данните, които не са FAIR, трябва да бъдат трансформирани в FAIR данни.

1.5. Софтуер с отворен код

Вече споменахме на няколко места терминът софтуер с отворен код. Като софтуер с отворен код се счита софтуер, който не само предлага достъп до първоизточника с неговия код, но трябва да отговаря и на следните критерии, изброени по -долу [11]:

- Безплатно преразпределение - никоя страна не може да бъде ограничена от продажбата или отдаването на софтуера като компонент от съвкупното разпространение, съдържащо програми от няколко различни източника;
- Изходен код - програмата трябва да включва изходния код и трябва да позволява разпространението на изходния код, както и специален формуляр;
- Производни/вторични трудове - модификации и производни/вторични научни работи/разработки трябва да бъдат разрешени;
- Достойно използване на изходния код на автора - може да се изисква, новите творби да носят различно име или номер на версията от оригиналния софтуер;
- Да не се прави дискриминация на лица или групи;
- Да няма дискриминация срещу областите на развитие;
- Разпространение на лиценз – правата отнасящи се за ползване на програмата, трябва да се прилагат за всички, на които програмата се препредава, без да е необходимо изваждането на допълнителен лиценз от тези страни;
- Лицензът не трябва да бъде специфичен за продукт;
- Лицензът не трябва да ограничава други софтуери;
- Лицензът трябва да бъде технологично неутрален;

1.6. Цифрови хранилища и управлението им

За съхраняването на данни са необходими цифрови хранилища. Тези цифровите хранилища са информационни системи, които събират, съхраняват, управляват, запазват и осигуряват достъп до цифровото съдържание [12]. Цифровите хранилища се използват и за съхранението на данни от различни изследвания. Едно от най-популярните цифрови хранилища е DSpace. Първоначално то е проектирано от изследователите в MIT Libraries и HP Labs и в момента се използва от над 250 институции. DSpace е безплатна софтуерна платформа с отворени източници за изграждане на хранилища на цифрови активи, с акцент върху лесен достъп до тези активи, както и тяхното дългосрочно запазване. Първоначално софтуерът е проектиран, като се има предвид конкретен модел на услугата: този на институционални хранилища от изследователски материали и по-специално научни статии, които се произвеждат от академични изследователски институции [13].

Друга популярна система за хранилища е Fedora, тя е система за съхранение на цифрови обекти с отворени източници. Използвайки стандартно базирана архитектура, платформата Fedora предоставя разширяема рамка от сервизни компоненти за поддръжка на функции като интеграция в търсачките, съобщения, работен поток, преобразуване на формати, масово събиране и други. Освен това към услугата на хранилището Fedora се включват функции като удостоверяване, упражняване на контрол за достъпа, определяне версията на съдържанието, репликация, проверка на целостта, динамични изгледи на цифровите обекти и други [14].

Предимствата на хранилищата с отворен достъп са:

- Отваряне на резултатите от институциите за световната аудитория;
- Като резултат на отворения достъп се получава максимална видимост и въздействие;
- Показване на институцията пред заинтересовани групи като бъдещ персонал, бъдещи студенти и други заинтересовани страни;
- Събиране и съхранение на цифрови данни;

- Управление и измерване на изследователска и преподавателска дейност;
- Осигуряване на работно пространство за незавършена работа и за съвместни или мащабни проекти;
- Активиране и насърчаване на интердисциплинарни подходи към изследванията;
- Улесняване на разработването и споделянето на цифрови учебни материали и помагала;
- Подкрепа на студентските начинания, осигуряване на достъп до дипломни работи и дисертации, както и място за разработване на електронни портфейли.

Най-често използваните хранилища на софтуери за съхранение със свободен достъп в съответствие с OpenDOAR са Digital Commons, DSpace и EPrints. Други примери за това са arXiv, bioRxiv, Dryad, Figshare, или Zenodo.

1.7. Цели на дисертацията

Основната цел на тази дисертация е да се анализира ситуацията и да се предложат съответни мерки в подкрепа на въвеждането на Отворената наука в балканските университети. Основният фокус е върху идентифициране на главните проблеми и предлагането на подходящи решения за въвеждане на Отворената науката в Албания, която досега не е постигнала голям напредък в тази област.

Следните по-конкретни цели са поставени в дисертацията:

- Анализ на развитието на Отворените науки в Европа. Проучване на европейския възглед за Отворената наука - Европейски облак за отворена наука.
- Идентифициране на етапите в прехода към Отворената наука, както и основните проблеми, пред които са изправени балканските университети. Разработката на подробен въпросник и разпространението му до съответни балкански институции. Анализ на основните ключови констатации, които са идентифицирани. Планиране на следващите стъпки за преминаване към Отворената наука на балканските институции.
- Разработване на прототип на отворено хранилище за научни изследвания, следвайки европейските стандарти, определени от European Open Science Cloud.

Разработването на този прототип е в подкрепа на прехода към Отворената наука в Албания

- Разработване на политика за Отворена наука, която да се прилага в Албания
- Описване на следващите стъпки за изпълнение и внедряване на Отворена наука в Албания.

1.8. Публикации, върху които се базира дисертацията

По-долу е даден списък на публикации, фокусирани върху Отворената наука в Европа и за напредъка на Отворената наука в балканските институции. Тези публикации в научни списания и/или международни конференции или научни доклади са в основата на дисертационния труд:

- Prompting an EOSC in practice - Final report and recommendations of the Commission 2nd High Level Expert Group on the European Open Science Cloud (EOSC), 2018. Directorate – General for Research and Innovation (European Commission) doi: 10.2777/112658. Available at: <https://op.europa.eu/en/publication-detail/-/publication/5253a1af-ee10-11e8-b690-01aa75ed71a1/language-en>
- Relevance of the EOSC initiative and FAIR principles in the realm of Open Science and implementation phases of the EOSC, Michel Schouppe Jean-Claude Burgelman. Proceedings of the XX International Conference “Data Analytics and Management in Data Intensive Domains” (DAMDID/RCDL’2018), Moscow, Russia, October 9-12, 2018. Available at: https://ec.europa.eu/research/openscience/pdf/eosc-fair_paper_schouppe-burgelman_2018.pdf#view=fit&pagemode=none
- Georgiev, Atanas; Stefanov, Krassen (2019) Bulgarian Open Science Digital Library - First Prototype, in Proc. of the 9th International Conference on Digital Presentation and Preservation of Cultural and Scientific Heritage (DiPP), Volume: 9, Pages: 251-258. WOS:000487853900025, ISSN: 1314-4006, eISSN: 2535-0366. Available at: http://dipp.math.bas.bg/images/2019/251-258_8_3.8_sDiPP2019-66_f_v.1.F_20190908.pdf

- Stanchev, Peter; Stefanov, Krassen (2019) Bulgarian Open Science Cloud, in Proc. of the 9th International Conference on Digital Presentation and Preservation of Cultural and Scientific Heritage (DiPP), Volume: 9, Pages: 259-264. WOS:000487853900026, ISSN: 1314-4006, eISSN: 2535-0366 Available at: http://dipp.math.bas.bg/images/2019/259-264_6_3.9_sDiPP2019-65_f_v.1.F_20190908.pdf
- Grigorov, A., Georgiev, A., Petrov, M., Stefanov, K., Varbanov, S. (2009) Building a Knowledge Repository for Life-long Competence Development, in IJCEELL V19 N4/5/6 2009, Special issue "Stimulating Personal Development and Knowledge Sharing", eds. R. Koper, K. Stefanov and D. Dicheva, pp.300-312. Available at: https://www.researchgate.net/publication/254324536_Building_a_Knowledge_Repository_for_Lifelong_Competence_Development

1.9. Структура на дисертацията

Тази дисертация е структурирана по следния начин:

Настоящото въведение представя целите, задачите и структурата на диесртационния труд.

Следващата глава представя обзор на литературата за Отворената наука в Европа. Тя дава подробен анализ на развитието на Отворената наука в Европа. Изследват се основните характеристики на Европейския отворен научен облак. Представени са и са описани различни изследователски хранилища. Направено е и сравнение между хранилищата.

Глава 3 определя етапите на прехода към отворена наука в Албания и България. Представен е въпросник, който е използван за проучване на състоянието по отношение на и прехода към Отворена наука в балканските институции. Направена е оценка на текущото състояние. Описан е основният елемент от прототипа на цифрово хранилище с отворен достъп в Албания, като първа стъпка към албанския облак за отворена наука. Разработена е архитектурата на такъв прототип.

Глава 4 представя прототипа и обяснява с подробности основните характеристики и подробности за албанското хранилище. След това се представя необходимия софтуер и изграждането на цифровото хранилище с помощта на DSpace-CRIS. Обръща се внимание и подробно се описва модела на данните и дизайна на цифровата библиотека, съвместима с EOSC. И накрая се разработва политика за отворена наука, която ще се прилага на практика.

В глава 5 се обсъждат заключенията и бъдещите работи, а в последния раздел е показана библиографията.

Глава 2. Състояние на отворената наука в Европа

В тази глава се анализира европейското виждане за Отворената наука. Първо, Европейският облак за отворена наука е въведен като модел, който Европа следва за Отворената наука. След това се илюстрират и сравняват цифрови хранилища за управление на изследователски данни. Въз основа на този анализ се предлага решение за избор на хранилище за разработване на Open Science Digital Repository.

2.1. Европейски облак за отворена наука

Европейският облак за отворена наука (ЕООН, European Open Science Cloud /EOSC/) е надеждна дигитална платформа за научната общност, осигуряваща безпроблемен достъп до данни и оперативно съвместими услуги, които се отнасят до целия цикъл от изследователски данни, от откриване до съхранение, управление, анализ и повторна употреба [15] [16].

ЕООН беше пуснат във Виена през ноември 2018 г. от Европейската комисия за създаване на европейска инфраструктура за данни, интегрираща облачни решения с голям капацитет и разширяване на услугите, предлагани на Open Science [17].

При пускането в краткосрочен план се поставя за цел до края на 2020 г. да са достъпни жизнеспособна платформа с официален статут, състояща се от правилата за участие, структура на управление, FAIR данни и оперативно съвместими услуги, първи приложения. Подкрепяйки политиката на ЕС за отворена наука, EOSC има за цел да даде на Европейския съюз (ЕС) глобален напредък в управлението на данните от научните изследвания и да гарантира, че европейските учени ще извлекат пълните ползи от науката, управлявана от данни.

EOSC е част от "Европейската облачна инициатива за изграждане на конкурентна икономика на данните и знанието в Европа", която стартира през 2016 г. Тя обединява съществуващите и нововъзникващите инфраструктури за данни, за да се предостави на европейска науката, промишлеността и държавните органи три основни характеристики [15]: инфраструктура за данни от световна класа за съхранение и управление на данни,

високоскоростна връзка за пренос на данни и мощни високопроизводителни компютри за обработка на данни.

За да се реализират амбициозните цели на EOSC, през 2018 г. беше създадена структура за управление, която обединява публични финансиращи организации, включително държавите-членки, асоциираните държави и Европейската комисия; и заинтересованите страни от EOSC, които са изследователи, доставчици на услуги, индустрия, публичен сектор и други потребители. Трестепенната структура включва:

- Управителен съвет, състоящ се от представители на държавите-членки и Европейската комисия, които заедно решават бъдещето на EOSC;
- Изпълнителен съвет, състоящ се от експерти от ключови европейски общности в областта на изследователската и електронната инфраструктура, които съветват относно прилагането на EOSC и пътя напред;
- Форум на заинтересованите страни, който дава глас на по-широката общност от потребители, доставчици на услуги, индустрия и публичен сектор.

EOSC ще позволи универсален достъп до данни и нови равни условия за изследователите от ЕС. На изследователите ще бъде предоставен списък с услуги:

- Лесен достъп чрез универсална точка за достъп за всички европейски изследователи.
- Интердисциплинарният достъп до данни отприщва потенциала на интердисциплинарните изследвания.
- Услугите и данните са оперативно съвместими; FAIR данни.
- Данните, финансирани с публични пари, по принцип са отворени; данните са възможно най-отворени, колкото е необходимо затворени.
- EOSC ще спомогне за увеличаване на признанието за интензивни данни изследвания и наука за данни в Европа.

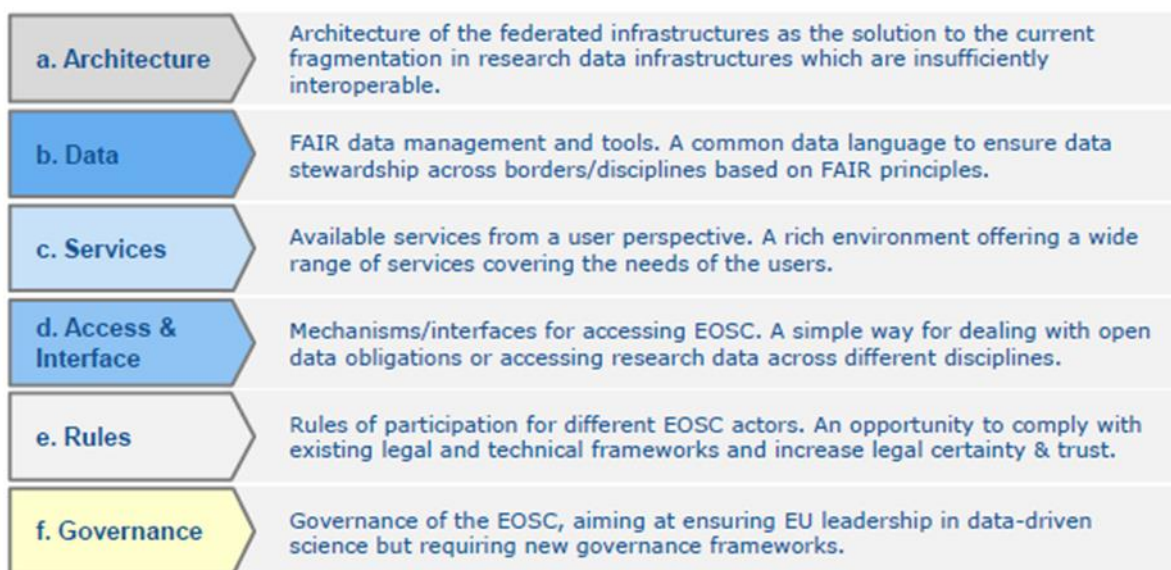
Основната цел на EOSC е да позволи универсален достъп до данни за всички изследвания в ЕС чрез точка за лесен достъп, тоест интердисциплинарен достъп до всички данни, използвайки услуги и данни, които са оперативно съвместими FAIR данни. Това означава, че EOSC ще помогне за разпознаването на отворен достъп и Отворена

наука в цяла Европа. За да се постигне това, бяха въведени няколко модела, както е показано във фигура 2-а.

	Current model	Federated	Centralized model
Resources	Fragmentation of resources and access to them	Integrated access to federated resources for ALL researchers	Access to a single data device (storage; single supercomputer?) for ALL researchers
Services	Varying quality of services currently	Service standards for all federated resources	Centrally decided standards
Interoperability	Varying levels of interoperability standards	Common standards for all federated resources	Centrally decided interoperability standards
Governance	Fragmented across 100+ institutions	Layered governance for EOSC participants – balanced stakeholders, MS, EC - with specific rules	Single body, centralized governance set up by EC
Costs & time to implement	Baseline	Marginally higher than baseline	Substantially higher than baseline
MS and stakeholder acceptance	Low	High	Extremely low

Фигура 2-а: Предложени различни модели EOSC

EOSC избира да интегрира федеративен модел, състоящ се от 6 основни действия, както е показано на Фигура 2-b [18]. Федералният модел има за цел да осигури интегриран достъп до всички услуги, предоставяни от EOSC. Тези услуги следват FAIR данните. EOSC предоставя списък с FAIR инструменти за данни, стандарти, услуги и каталози. Той също така предоставя списък с услуги за всички изследователи, включително услуги за намиране, съхраняване, достъп, повторна употреба и разпространение на данни, генерирани от други изследователи. EOSC предоставя портал, където всички изследователи могат лесно да получат достъп до всички споменати инструменти и информация, като имат правила за участие и управителен съвет, който да реши бъдещето на EOSC, отворен достъп и отворена наука в Европа.



Фигура 2-б: Действия на федеративен модел на EOSC

Архитектурата на федералната инфраструктура е да помогне за установяването на 4 цели от изследователската визия на Европейския съюз. Първо, това ще помогне за организирането и функционирането на компонентите на федеративното ядро на EOSC; включително споделени ресурси на EOSC, платформа EOSC, рамка за съответствие; за улесняване, наблюдение и регулиране при необходимост на ежедневните транзакции в цялата федерация.

Второ, ще изследва национални, хоризонтални и тематични инфраструктури за данни; включително хранилища на данни за изследвания, доставчици на услуги, научни мрежи, RI и eInfras; да установят своя интерес, желание и способност да се присъединят към фазата на развитие на EOSC.

Трето, ще дефинира изискванията за компоненти на EOSC и акредитирани/сертифицирани от FAIR инфраструктури за данни, въз основа на спазването на Правилата за участие на EOSC.

И накрая, това ще помогне за създаването и актуализирането на регистъра на обединените инфраструктури за данни, за да се следи и налага спазването на правилата и общите правила чрез схемата за акредитация/сертифициране на FAIR. Той ще продължи да надгражда проектите „EOSC Hub“, „OpenAIRE“, „FREYA“ и „EOSC pilot“; Проекти MERIL и RICH.

EOSC ще предостави на всеки изследовател пет вида услуги:

- Идентификация и удостоверяване, както и точка за достъп и система за маршрутизиране към ресурсите на EOSC;
- Защитена и персонализирана работна среда/пространство, включително дневник, настройки, запис за съответствие и предстоящи проблеми;
- Достъп до съответна информация, включително състоянието на EOSC, списък на инфраструктурите, свързана с политиката информация, рамка за съответствие и конкретни насоки за това как да направим данните FAIR, да сертифицираме хранилище или услуга, да осигурим съвместни услуги;
- Услуги за намиране, достъп, съхраняване, повторна употреба и анализи, включително анализи, сливане/сливане на данни, извличане на данните, генерирани от други, каталогизирани по подходящ начин;
- Услуги, за да направят данните си FAIR, да ги съхраняват и да осигурят дългосрочно запазване.

2.2. Дигитални библиотеки

Цифровите библиотеки с отворен достъп се наричат още платформи с отворен достъп. Тези платформи са хранилища за научни изследвания, които са съвместими с Open Access. Те са цифрови библиотеки със следните компоненти [19]:

- активи за научни и образователни знания в цифров вид, като публикации, книги, филми, колекции от данни, които автоматично се индексират, търсят и имат достъп до тях с помощта на информационни технологии;
- всички активи на знанието са класифицирани в тематични колекции и йерархии, като се използват различни класификационни схеми и онтологии;
- всички активи на знанието са описани със съответния формат на метаданни, в съответствие с основните стандарти за метаданни, използвани в цифровите библиотеки, и в съответствие със стандарта OAI-PMH (The Open Archives Initiative Protocol за събиране на метаданни). Всички такива активи се съхраняват в цифров вид, оценяват се чрез мрежата и следват правилата за отворен достъп.

EOSC предложи формат на модел на данни, който да се използва - Current Resource Information System (CRIS). CRIS е база данни или друга информационна система за съхраняване, управление и обмен на контекстуални метаданни за изследователската дейност, финансирана от спонсори на изследвания или проведена в изследователската организация. Моделът на данните позволява представяне на изследователски обекти, техните дейности/взаимовръзки (изследване) и техните резултати (резултати), както и висока гъвкавост с формални (семантични) връзки, позволява качествена поддръжка, архивиране, достъп и обмен на научноизследователска информация и подпомага трансфера на знания към лицата, вземащи решения, за оценка на изследванията, мениджърите на научните изследвания, стратегите, изследователите, редакторите и широката общественост.

Използваният формат е CERIF (Common European Research Information Format, Общ европейски информационен формат за научни изследвания). Основните характеристики на CERIF са:

- Концепция за изследователските обекти и техните взаимоотношения;
Спецификация (концептуално ниво)
- Описание на изследователските организации и техните взаимоотношения;
Модел (логическо ниво)
- Формализиране на изследователски обекти и техните взаимоотношения;
Скриптове за база данни (физическо ниво)

CRIS може да бъде реализиран, като се използва подмножество или супермножество на пълния модел CERIF за проекти, хора, организации, публикации, патенти, продукти, услуги и съоръжения (по-специално оборудване) с роли, временно обвързани взаимоотношения. Предимствата на CERIF са: неговата неутрална архитектура; моделът на данните може да бъде реализиран като реляционен, обектно-ориентиран или информационен модел за извличане на данни; моделът на процеса може да бъде реализиран като СУБД и заявка; HTML уеб/ събиране/ IR-заявка или като усъвършенстван технологичен процес, базиран на знания.

Но взаимодействието изисква структурирана схема. CERIF се използва като модел за внедряване на самостоятелен CRIS (но готов за взаимодействие), като модел за

дефиниране на обвивката около наследствения non-CERIF CRIS, за да позволи хомогенен достъп до хетерогенни системи и като дефиниция на формат за обмен на данни към създадено общо хранилище на данни от няколко CRIS.

2.3. Хранилища за управление на научни данни

Управлението на изследователските данни /Research Data Management/ е дигитална библиотека, която се използва за изследователски проекти и данни от изследванията. Управлението на изследователските данни (RDM) описва организацията, съхранението, запазването и споделянето на данни, събрани и използвани в изследователски проекти [20].

Има много хранилища за управление на изследователски данни, но три са най-разпространените. Това са DSpace-CRIS, Fedora и Zenodo (Invenio). По-долу е описан всяко едно от тези хранилища. Важно е да се отбележи, че всички тези хранилища са съвместими с OpenAIRE и със софтуера с отворен код. Те са близки до европейската визия за Отворената наука и работят в тясно сътрудничество с принципите на данните на FAIR.

2.3.1. DSpace-CRIS

DSpace-CRIS е първото безплатно разширение на DSpace с отворени източници за разработването на изследователски данни и управление на информация [21]. DSpace е софтуерна платформа с отворени източници, която позволява на организациите да:

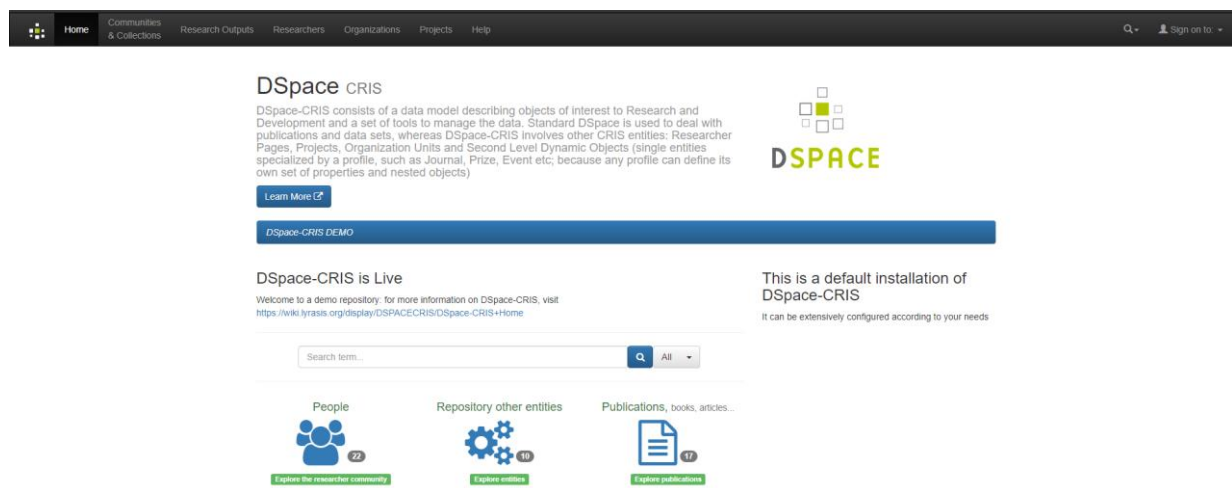
- събират и описват дигитален материал с помощта на модул за подаване на работен процес или различни опции за програмно запълване;
- разпространяват цифрови активи на организацията в мрежата чрез система за търсене и извличане;
- запазват цифровите активи в дългосрочен план.

От гледна точка на архитектурата на приложението, DSpace е приложение състоящо се от база данни, мениджър на хранилището и уеб интерфейс. Архитектурата включва

специфичен модел на данни с конфигурируеми схеми за метаданни, работни процеси, и да сърфирате или функция за търсене.

По подразбиране DSpace използва схема за метаданни, базирана на Qualified Dublin Core (QDC). Институциите могат да разширят тази базова схема или да добавят персонализирани схеми, подобни на QDC. DSpace може да импортира или експортира метаданни от други основни схеми за метаданни като MARC или MODS.

Фигура 2-в показва интерфейса DSpace-CRIS, който включва демонстрация в dspace-cris.4science.cloud [22].



Фигура 2-с: Демонстрация на интерфейса DSpace –CRIS

Основната характеристика на DSpace-CRIS е неговият гъвкав модел на данни, който позволява събиране и управление на данни и информация от изследванията, характерни за системата CRIS, за да бъдат дефинирани обекти и атрибути с техните взаимоотношения. А информационна система за научни изследвания CRIS е база данни или друга информация за съхранение, управление и обмен на контекстни метаданни за изследователската дейност, финансирана от спонсори или проведена в изследователска организация.

DSpace-CRIS може да съхранява публикациите, които са много важна част от изследването, но също така и проекти, патенти, профили на изследвания, грантове и организационни единици. Той дава възможност за запълване, показване, управление на данни и метаданни за всички изследователски обекти, споменати по-горе.

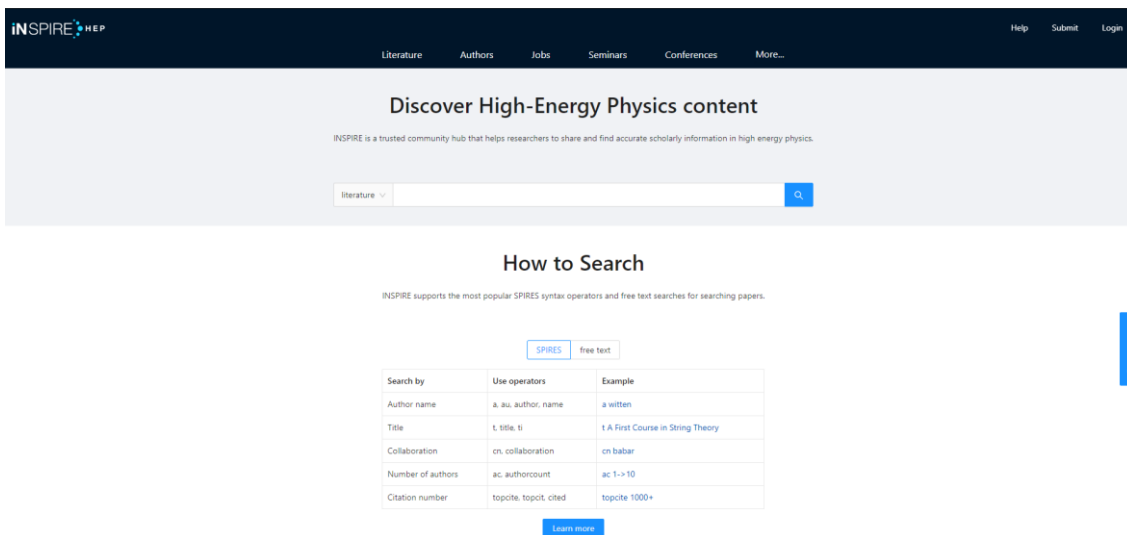
DSpace-CRIS се състои от модел на данни, описващ обекти, представляващи интерес за научноизследователска и развойна дейност, и набор от инструменти за управление на данните и метаданните на изследователските обекти: Страници за изследователи, Проекти, Организационни единици и Динамични обекти от второ ниво, специализирани по профил, като журнал, награда или събитие, защото всеки профил може да дефинира свой собствен набор от свойства и вложени обекти.

2.3.2. Zenodo -Invenio

Zenodo и Invenio винаги се разглеждат заедно и има причина за това. Zenodo е добър пример за услуга за управление на изследователски данни, която се хоства от CERN, но не е хранилище, което можете да стартирате като свое приложение. Това е Invenio, което е приложение за хранилище, което може да ви напомни на Zenodo [23] [24]. Invenio е проект на софтуер с отворен код, първоначално разработен от CERN. Zenodo се захранва от Център за данни на CERN и рамката за цифрова библиотека Invenio и е изцяло работеща с продукти с отворен софтуерен код.

Zenodo предлага услуги на Open Science в Европа от CERN като част от мисията си да направи резултатите от работата си достъпни.

Пример на Zenodo - интерфейс Invenio е наличен в inspirebeta.net [24] е показано на фигура 2- г. Inspire е високоенергийната информационна система по физика, която съчетава успешното съдържание на базата данни SPIERS, курирано в DESY, Fermilab и SLAC в продължение на десетилетия в CERN.



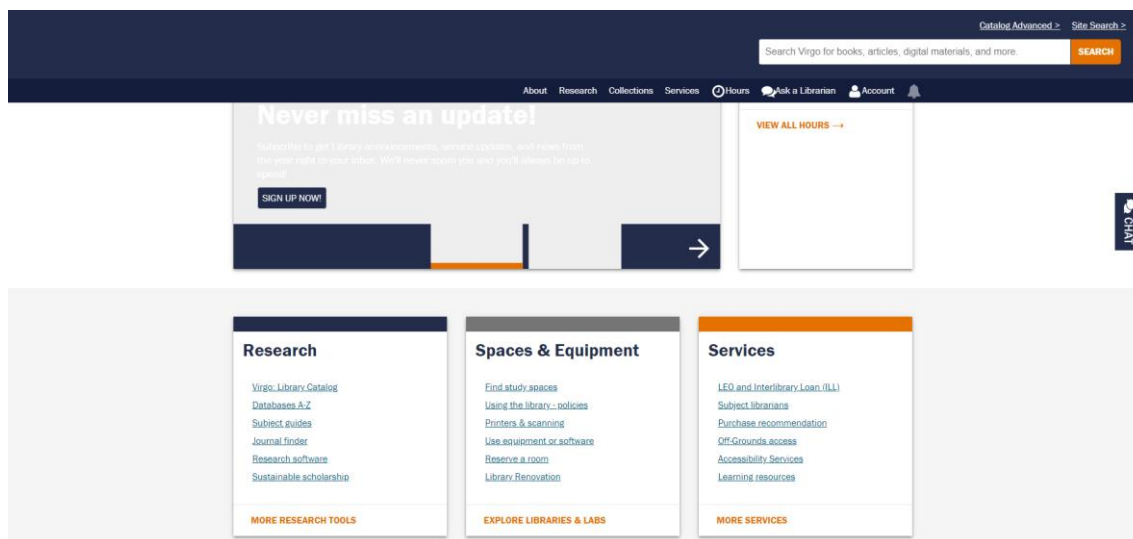
Фигура 2-г: Invenio интерфейс на Inspirebeta.net

От друга страна, Invenio е основата на Zenodo и сега предлага три вида услуги, включително проект за управление на данни за изследвания, който е с отворени източници, наречен InvenioRDM. Той идва с предварително конфигурирани профили на хранилища за институционални хранилища и системи за управление на данни от изследвания.

2.3.3. Fedora

Fedora (гъвкава разширяема архитектура на хранилище за цифрови обекти) е модулна архитектура, изградена на принципа, че оперативната съвместимост и разширяемостта се постигат най-добре чрез интегриране на данни, интерфейси и механизми като ясно определени модули [25].

Fedora е архитектура за управление на цифрови активи, върху която са изградени много видове цифрови библиотеки и системи за институционални хранилища. Fedora е основната архитектура за цифрово хранилище и не е цялостна система за приложение за управление, индексирание, откриване и доставка. Тя е базирана на Java, така че е потенциално приложима за всяка платформа. Fedora е съвместима с FAIR, но не успява да предостави директно хранилище за директно управление на данни от изследвания. Някои изследователски институции и университети използват персонализирано хранилище, базирано на Fedora, за управление на данни като Университета на Вирджиния (<https://www.library.virginia.edu/>), показано на Фигура 2-д.



Фигура 2-д: базиран на Fedora интерфейс в Университета на Вирджиния

2.4. Изводи

Хранилища с отворен достъп са обяснени по-горе с всичките им характеристики. За да обобщим, трите представени хранилища са сравнени в Таблица 1 по-долу.

Таблица 1: Сравнение на хранилища с отворен код.

	Fedora	DSpace-CRIS	Invenio RDM
Предоставен потребителски интерфейс	Да	Да	Да
Динамично генерирани страници	Да	Да	Не
Софтуер с отворен код	Да	Да	Да
Изключена конфигурация за RDM	Не	Да	Да
Изследователски публикации управление на субекта	Да	Да	Да
Управление на други обекти на изследователски данни	Не	Да	Да
Съвместим с OpenAIRE	Да	Да	Да
FAIR съвместим	Да	Да	Да
Гъвкава архитектура	Да	Не	Не
Съвместимост с множество формати на метаданни	Да	Да	Да
Персонализирани формати на метаданни	Да	Не	Да

Всички горепосочени софтуери имат потребителски интерфейс, съвместими са с отворените източници, OpenAIRE и FAIR и поддържат множество формати на метаданни в техните обекти. Когато обаче става въпрос за изследователски обекти, само две предоставят конфигурации, подходящи за изследователски организации. В заключение, DSpace-CRIS и Invenio са по-подходящи за хранилища за управление на изследователски данни в изследователски институции и университети, тъй като са предварително конфигурирани за изследователски организации. Това не означава, че Fedora не може да се използва. Както е посочено в горния раздел, някои университети използват Fedora и основната причина за това е фактът, че тя е гъвкава. Нейната архитектура позволява интеграция с други системи, присъстващи в техните институции, и използвайки Fedora, те добавят практиките за управление на изследователските данни към такива системи.

За дисертацията ще се използва DSpace-CRIS, а основната причина е познатият потребителски интерфейс, с който са свикнали изследователите. Освен това той предоставя всички необходими обекти за управление на изследователски данни, необходими за пълно управление на изследователските данни в хранилището и страниците се генерират динамично. Добър пример за този избор е първият прототип на Bulgarian Open Science Cloud (BOSC) [26], който използва DSpace-CRIS и е съвместим с OpenAIRE и CRIS. Избраният модел на данни разчита на основните единици на модела на Общия европейски информационен изследователски формат (CERIF) [27]. CERIF е стандартът, който ЕС препоръчва на своите държави членки за записване на информация за изследователската дейност. Той е разработен и поддържан от EuroCRIS. EuroCRIS е създаден, за да се справи с проблемите на настоящите изследователски информационни системи в световен мащаб с акцент върху Европа. Едно от предимствата на InvenioRDM е възможността да се работи с персонализирани формати на метаданни, но това предимство не е от значение в тази дисертация, тъй като използваният формат ще бъде съвместим с CERIF.

Следвайки модела на BOSC, DSpace-CRIS е системата с отворен код, използвана за създаването на първия прототип на хранилището с отворен достъп в Албания.

Глава 3. Текущото състояние на Отворената наука на Балканите

В тази глава етапите на прехода към Отворена наука и основните проблеми са идентифицирани чрез проучване, което беше направено в Албания и България. Представено е сравнение на основните разлики между тези страни и други европейски страни. Основните характеристики на първия прототип на портал за Отворена наука за Албания са описани в края на главата.

3.1. Състояние

Има проучвания, които улавят настоящото състояние на Отворената наука в балканските страни. Едно от тях е National Open Science Initiatives in Europe (NI4OS-Europe), което се фокусира върху отворената наука в страните от Югоизточна Европа [28]. Разпространява се в 15 страни, включително Албания и България. Въпреки това и двете страни са с ниско участие от 1,91% в Албания и 7,48% в България. В него се обсъждат критериите, правилата и политиките за финансиране на Отворената наука; осведоменост относно EOSC, FAIR и отворените данни. То обаче не включва въпроси, свързани с хранилища за изследвания или технически аспекти на тези системи.

Важна инициатива, която обхваща настоящото състояние на страните от Централна Европа, е проучването „Национални инициативи“ [29]. То е подробно и улавя прехода към Отворена наука на Централна и Северна Европа. Проучването обаче е насочено по специално към EOSC групите и не включва балканските страни.

Най-разпространеното проучване е EUA „Оценка на изследванията в прехода към отворена наука“ [30]. То е направено през 2019 г. в 32 държави с 260 участника. Нито Албания, нито България обаче са част от това проучване. Освен това, самото проучване не обсъжда технически подробности, а се фокусира повече върху осведомеността за отворената наука и нейната подкрепа в институциите.

Могат да се споменат и други национални инициативи, като Националната научна програма „Информационни и комуникационни технологии за единен дигитален пазар в науката, образованието и сигурността“ (ИКТвНОС), която е инициатива, финансирана от Министерството на образованието и науката на РБългария, която се фокусира върху разширяването на участието на българската изследователска общност в Европейското изследователско пространство чрез изграждане на силна връзка в науката, образованието и обществото [31].

Проучването, направено в рамките на дисертационния труд, обхваща настоящото състояние на Албания и България по отношение на Отворената наука и включва и теми, свързани с технологични аспекти на въпроса.

3.2. Методология

Проучването, представено в Приложение А, се опитва да улови прехода към Отворена наука в България и Албания. То е подобно на EUA, което не е разпространено в тези страни. Включва също така някои технически подробности, които не са били част от други проучвания, разпространени в България и Албания. Проучването се опитва да постигне по-голямо участие в сравнение с NI4OS-Европа, което е най-разпространеното проучване в Албания. Това проучване се използва за оценка на Албанската Отворена наука и за планиране на съответни обучителни и изследователски дейности в помощ на Албания, за да може да се присъедини към инициативата EOSC. Проучването показва настоящото състояние на Отворената наука на албански език и го сравнява с България и Европа.

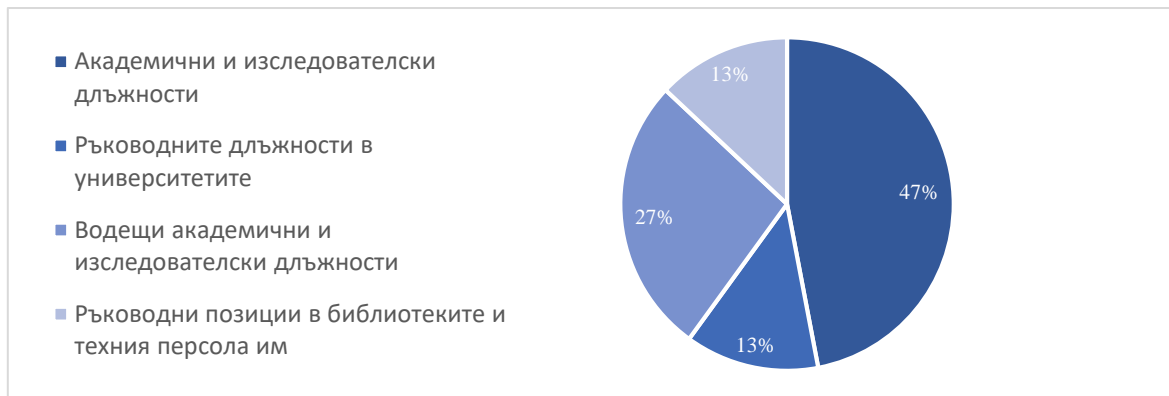
Анкетата е структурирана с 32 въпроса. Включена е комбинация от отворени въпроси, класиране, въпроси с няколко избора и въпроси с един избор. Проучването обхваща по-голямо разнообразие от теми от Отворената наука от доклада на EUA, който се състои от 20 въпроса. Темите включват текущото състояние на оценката на научните изследвания в българските и албанските университети, както и начина, по който институциите разглеждат своите практики. Освен това, проучването е разширено с технически въпроси относно техническите аспекти, като например внедряване на институционални

хранилища (IR), протоколи, стандарти, постоянни идентификатори и използвани показатели.

Въпросите относно настоящите процедури за оценка на изследванията са подобни на доклада на EUA. Това означава, че в първия раздел имаше 3 секции, обхващащи изследователската кариера и оценката на работата на изследователските звена, отворения научен подход и разпределението на финансиране за научни изследвания в рамките на институцията са във втория раздел, а след това техническите аспекти се изследват в третия раздел. Целта е да се провери дали се прилага текущата изследователска информационна система, използвана в университетите, и кои протоколи, стандарти и речници се използват. Прави се преглед на други институционални хранилища, използвани в университетите, и накрая - прилагат ли се на институционално ниво политиките за отворен достъп.

Респондентите са от академичен и изследователски състав, заемачи ръководни позиции в университети, хора на ръководни позиции в библиотеки, както и персонал в тях. Докладът съдържа 53 отзива от 25 различни институции в България и Албания. Има 13 отговора от 6 албански институции и 40 отговора от 19 български институции. Делът на извадката е в близост до 2 отговора на институт. Албанският брой респонденти е сравнително нисък, но всички участници са с надеждата, че тяхното виждане отразява напълно състоянието на преход към Отворена наука на албанските институции. Резултатите показват, че 11 от анкетираните (85%) са на водещи академични или изследователски позиции, с изключение на 1 отговор от научния персонал и 1 отговор от управленски персонал, работещ в библиотека.

По-голямата част от анкетираните в България са академични и изследователски служители - 47%, 27% са водещи академични и изследователски длъжности (директори, ръководители на катедри), следвани от 13% от ръководните длъжности в университетите (декани, ректори, проректори), а 13% от отговорите са от ръководни позиции и персонал на библиотеки, както е показано на фигура 3-а.



Фигура 3-а: Профили на респондентите в България

3.3. Изводи

Проучването предоставя преглед на настоящото състояние на прехода към Отворена наука в българските и албанските институции. Изводите от проучването показват, че българските и албанските институции подкрепят Open Science. Някои ключови констатации включват:

- По-голямата част от българските институции имат публично достъпна политика за оценка на изследванията, подобна на другите институции на ЕС, докато Албания изостава с повече от половината институции без политика за оценка на изследванията;
- Нито един от албанските университети няма политика за отворен достъп или анкетираните не знаят дали такъва политика съществува;
- Албанските институции нямат CRIS. Дори в българските институции само 10% имат такива системи и нито една от тях не е съвместима с CERIF.

От проучването става ясно, че както българските, така и албанските институции са все още в ранните етапи на преход към Отворена наука, с някои малки изключения в България. България вече е започнала да мисли и да работи за отваряне на Българската наука към света. Планиран е Българският отворен научен облак (BOOSC), базиран на същите принципи, стандарти и технологии на EOSC и напълно съвместим с него. Те също така имат първия прототип на Българското отворено научно цифрово хранилище.

Проучването ни насочва към необходимостта от Албанския отворен научен облак (AOOSC). Целта на AOOSC е да направи албанската наука отворена и да помогне на Албания да се присъедини към инициативата EOSC. В следващата глава е показано разгръщането

на първия прототип на AOSC, създаден по начин, който да е напълно съвместим с EOSC. Албанският отворен научен облак ще осигури следните принципи:

- Единна изследователска инфраструктура, която да осигури по-добра култура на управление на изследователските данни сред учените в Албания;
- Данни, които са достъпни, оперативно съвместими и за многократна употреба (FAIR данни);
- Инструменти за осигуряване на надеждна среда за FAIR данни, услуги за преобразуване на FAIR данни;
- Интерфейс за работа с данни от различни дисциплини, както и входна точка с уникална услуга за идентификация и удостоверяване;
- Информация за политиките с отворени източници, информация за FAIR данни, насоки за съответствие с EOSC и OpenAIRE;
- Услуги за намиране, достъп, повторно използване и анализ на изследователските данни, генерирани от други изследователи с подходящи набори от данни, съвместими с EOSC.

Първата стъпка към Албанския облак за отворена наука е да се предостави цифрово хранилище за изследвания с отворен достъп, за да се запазят изследователските резултати и активи, в съответствие с EOSC. Това е дигитална библиотека с отворен източници, която ще бъде първата стъпка към пълноправен Албански облак за отворена наука. Албанското цифрово хранилище с отворен достъп следва да има следните характеристики:

- За разработване на такъв елемент ще се използва софтуер за цифрово хранилище с отворени източници DSpace-CRIS.
- След допълнителна проверка на данните, съхранявани в албанските университети, само имената на автори и публикации могат да бъдат извлечени от съществуващите данни. Това означава, че се изправяме пред важно предизвикателство само по себе си. От друга страна, съвместимите с EOSC и OpenAIRE метаданни могат да се използват от самото начало, без да се сблъскват с трудностите при трансформирането на оригиналните данни в FAIR данни.

- Метаданните, които ще се използват, са модел на данни CERIF, който е съвместим с OpenAIRE и EOSC. Използва се само подмножество от модела CERIF, тъй като моделът с метаданни CERIF включва по-широк кръг обекти, които не са от значение за OpenAIRE.
- За разработване на такъв елемент ще се използва софтуер за цифрово хранилище с отворен код DSpace-CRIS.

За попълването на дигиталната библиотека ще започнем с проверка на данните за Министерството на образованието, спорта и младежта в Албания и създаване на списък с всички албански университети и изследователи. Използвайки този списък, ще се търсят външни библиографски ресурси и ресурси от набори от данни (като WoS, Scopus), за да се предостави списък с публикации или метаданни. След това данните, внесени от външни библиографски ресурси, се изпращат до университети и изследователи, за да могат те да изчистят или предоставят допълнителна информация за метаданните. Също така в интернет ще се търси допълнителна информация като алтметрика и ще се добавя към различните раздели. Публични хранилища се отварят за достъп и се изтегля информация.

Следващата стъпка е намирането на постоянни цифрови идентификатори (като DOI, ORCID). Това се извършва чрез използване на регулярни изрази с CrossRef REST API за намиране на DOI и ORCID Public API за намиране на ORCID идентификатори за всяко от изследванията в публикацията DOI. Използвайки услугите за разрешаване на DOI, се добавя допълнителна информация за обекта (като URL адреси) за обогатяване на библиографските метаданни на такъв обект.

И накрая, всички данни са нормализирани, за да отразяват модела на данни, използван в това хранилище, за да съхраняват изцяло данните, следвайки указанията на EOSC и OpenAIRE за метаданни.

Глава 4. Албанско цифрово хранилище за Отворена наука

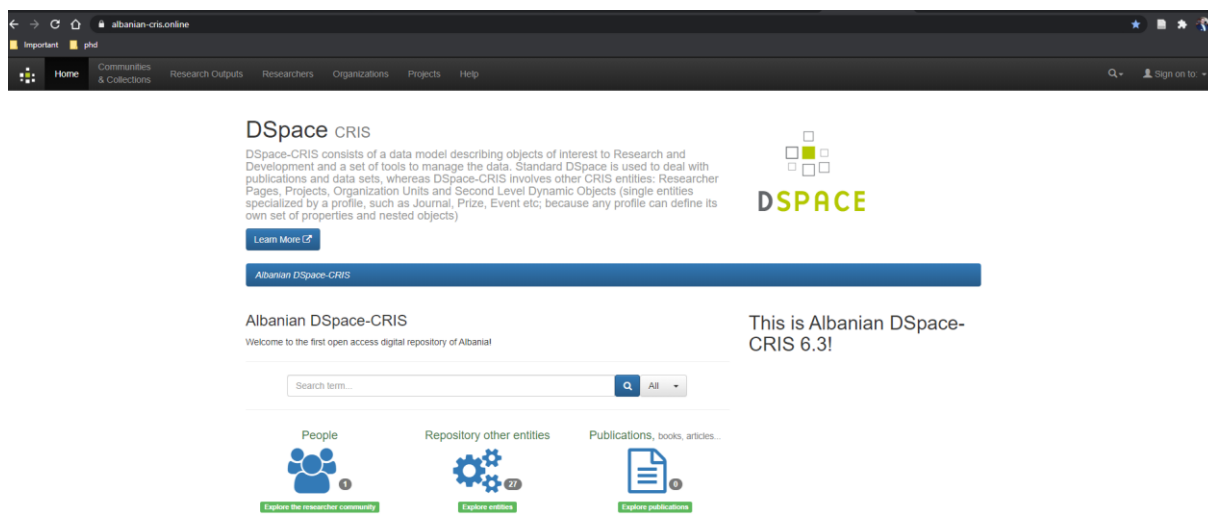
В тази глава е описано създаването на прототип на албанската информационна система за отворена наука, напълно съвместима с EOSC. Направен е анализ на изискванията, които трябва да се спазват. След това е представен необходимият софтуер и изграждането на цифровото хранилище с помощта на DSpace-CRIS. Илюстриран е и моделът на данни и дизайна на цифровата библиотека, съвместима с EOSC. Накрая е представена разработена политика за Отворената наука, която да бъде прилагана.

4.1. Анализ на изискванията

Изводите от предишната глава показват, че албанските изследователи подкрепят Отворената наука. Повечето институции обаче нямат дигитално хранилище, а да не говорим и за съвместимо хранилище за Open Science. За това Open Source Digital Repository е следващата стъпка към Отворената науката в Албания. За да се изгради такава една система, която да бъде напълно съвместима с EOSC, трябва да се вземат някои важни решения.

Първо, системата е изградена върху DSpace-CRIS. Това е първото цифрово хранилище с отворени ресурси за изследователски институции. Основната причина, поради която се избира такава система, е фактът, че тя е в съответствие с международните стандарти, използвани в EOSC. Тези стандарти включват улесняване на оперативната съвместимост и трансфера на данни. Приетите стандарти са ORCID, OpenAIRE насоки за управление на хранилища на литература, PlanS by Coalition S и FAIR принципи за данни.

DSpace-CRIS позволява внедряване на различни модели данни, като по този начин могат да се използват съвместими с EOSC метаданни. Основните субекти на DSpace-CRIS включват изследователи, проекти, общности и колекции, организации и изследователски продукти. Фигура 4-а показва текущите елементи на DSpace-CRIS, внедрени в Албания. Системата е достъпна на адрес <https://albanian-cris.info/> [32, р.].



Фигура 4-а: албански DSpace -CRIS

Както бе посочено по-горе, в Албания до момента не е била налична система, което означава, че всички данни ще трябва да се попълват ръчно. Най-добрият наличен вариант е да се създаде списък на всичките образователни институции в страната. Такъв списък е показан в таблица 2 по-долу с всичките известни данни от такива институции. Информацията се проверява в Министерството на образованието, спорта и младежта в Албания.

Таблица 2: Списък на институциите в Албания

Име на институция	Описание	Вид на институцията	Авторско право
Academy of Albanological Studies	ASA	Public	© Akademia E Studimeve Albanologjike
Agricultural University of Tirana	UBT	Public	© Universiteti Bujqesor i Tiranës
Albanian University	albanianuniversity	Private	© Albanian University
Aldent University	UAL	Private	© Universiteti Aldent
Aleksandër Moisiu University of Durrës	UAMD	Public	© Universiteti Aleksandër Moisiu
Aleksandër Xhuvani University	UNIEL	Public	© Universiteti i Elbasanit "Aleksandër Xhuvani"
Armed Forces Academy	AFA	Public	© Akademia e Forcave të Armatosura

Bedër University	BU	Private	© Kolegji Universitar Bedër
Canadian Institute of Technology	CIT	Private	© CIT
Epoka University	EU	Private	© Epoka University
Eqrem Çabej University	UOGJ	Public	© UOGJ
European University of Tirana	UET	Private	© UET
Fan Noli University	UNKORCE	Public	© Universiteti "Fan S. Noli", Korçë
Luarasi University	LUARARI UNIV	Private	© Universiteti Luarasi
Marin Barleti University	UMB	Private	© Universiteti Barleti - Tiranë
Mediterranean University of Albania	UMSH	Private	© UMSH
Metropolitan University of Tirana	UMT	Private	© Metropolitan University of Tirana
Our Lady of Good Counsel University	UNIZKM	Private	© Our Lady of Good Counsel - Catholic University
Polis University	universitetipolis	Private	© UNIPOLIS
Polytechnic University of Tirana	UPT	Public	© UPT
Sports University of Tirana	UST	Public	© Universiteti I Sporteve Të Tiranës
University of Arts	UART	Public	© Universiteti i Arteve
University of Medicine, Tirana	UMT	Public	© UMT
University of New York Tirana	UNYT	Private	© University of New York Tirana
University of Shkodër "Luigj Gurakuqi"	UNISHK	Public	© University of Shkodër "Luigj Gurakuqi"
University of Tirana	UNITIR	Public	© UNITIR
University of Vlorë "Ismail Qemali"	UNIVLORA	Public	© Universiteti i Vlorës

Първата стъпка за попълване на дигиталната библиотека започва с проверката на данните от Министерството на образованието, спорта и младежта в Албания и създаване на списък с всички албански изследователи от университетите. Използвайки този списък, се изследват външни библиографски ресурси и ресурси от набори от данни (като WoS, Scopus), за да се предостави списък с публикации или метаданни. След това данните, внесени от външни библиографски ресурси, се изпращат до университетите и изследователите, за да се предоставят допълнителна информация за метаданните. С помощта на търсене в интернет ще се попълва и допълнителна информация като алтметрика, която ще се бъде добавена там. Също така ще се осъществи достъп и ще се изтеглят данни от хранилища за публикации с пълнотекстова информация.

Намирането на постоянни цифрови идентификатори (като DOI, ORCID) е следващата стъпка. Това се извършва чрез използване на регулярни изрази с CrossRef REST API за намиране на DOI и ORCID Public API за намиране на ORCID идентификатори за всяко от изследванията в публикацията DOI. Използвайки услугите за разрешаване на DOI, допълнителна информация за обекта (като URL адреси) се изследва и добавя, когато е възможно. Това ще обогати библиографските метаданни.

И накрая, всички данни са нормализирани, за да отразяват модела на данни, използван в това хранилище, за да се съхраняват напълно данните, следвайки указанията на EOSC и OpenAIRE за метаданни.

Тези университети се съхраняват и създават като общности, а също и изследователски организации в цифровото хранилище с отворен достъп, както е показано на фигура 4-b. След това за всеки университет се създава списък с изследователи. След това списъкът се изпраща до всеки университет за двойна проверка на данните.

Communities and Collections

Shown below is a list of communities and the collections and sub-communities within them. Click on a name to view that community or collection home page.

The screenshot displays a list of four academic institutions, each with a logo on the left and text on the right. Below the text is a search bar. The entries are: 1. Academy of Albanological Studies (ASA) with a logo of two overlapping circles. 2. Agricultural University of Tirana (UBT) with a green circular logo featuring a building. 3. Albanian University (albanianuniversity) with a circular logo containing a book and the text 'ALBANIAN UNIVERSITY'. 4. Aldent University (UAL) with a red circular logo containing the text 'UAL' and 'UNIVERSITETI ALDENT'.

Фигура 4-б: Пример от списъка на университетите в албанското цифрово хранилище с отворен достъп

4.2. Инсталиране на DSpace-CRIS

На сървъра е инсталиран DSpace-CRIS 6.3. Избраният DSpace-CRIS е най-новата версия към момента на внедряване. По подразбиране DSpace-CRIS може да се използва за конфигуриране и поддържане на изследователски данни от потребителския интерфейс. Изследователските данни, които могат да бъдат въведени, са на организационните единици, изследователи, публикации, проекти и общности. По подразбиране дори възможността за въвеждане на оборудване е възможна от потребителския интерфейс. Това са предварително конфигурираните CRIS обекти, които могат да бъдат персонализирани и свързани заедно.

Друга функция, която по подразбиране е възможността за оценка, отчитане и анализ на цялата система за хранилищата. Това означава, че общностите, организационните единици, изследователите или публикациите могат да показват броя на изтеглянията, броя на цитатите, а също и най-добрите/глобални изтегляния. Налице е и глобално търсене и индексирание, което помага с анализите.

И накрая, DSpace-CRIS може да предложи оперативна съвместимост с външни системи и бази данни. Скриптовете се предоставят за периодично запитване на библиографски бази данни на Scopus и Web of Science. Също така, ORCID интегрира както за публични,

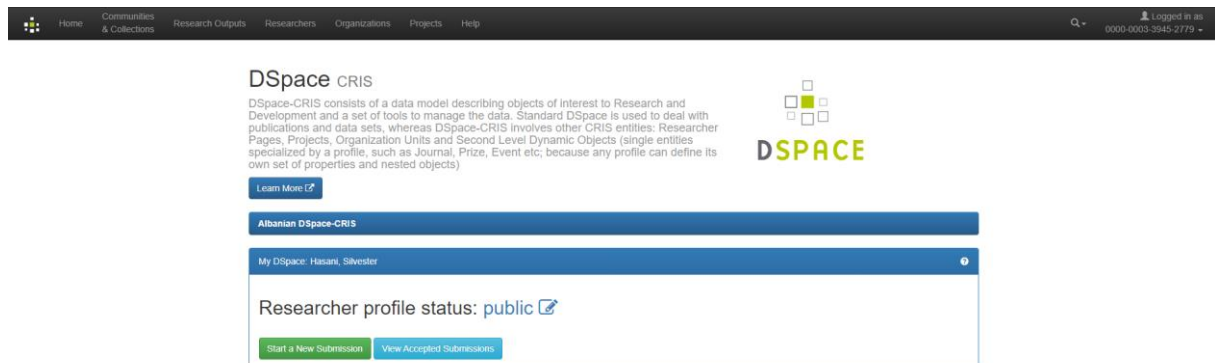
така и за членове на API. Това може да бъде видно на Фигура 4-с, където е показана конфигурирана интеграция с ORCID.



The screenshot shows the 'Log In to DSpace' interface. On the left, there is a login form with fields for 'E-mail Address' (containing 'hasani@uni-sofia.bg') and 'Password' (masked with dots). Below the password field is a link 'Have you forgotten your password?' and a green 'Log In' button. On the right, there is a section titled 'Use your ORCID or create a new one'. The text explains that ORCID provides a persistent digital identifier and supports automated linkages. Below this text is an 'ORCID Login' button with the ORCID logo.

Фигура 4-с:ORCID интеграция и вход

Възможно е да влезете, като използвате акаунт, създаден от администратора на системата, или нов потребител, регистриран в системата, но също и вход с вашия цифров идентификатор ORCID. Фигура 4-d показва влизане с помощта на профила ORCID. След като сте влезли с профила ORCID, можете да управлявате профила на изследователя, създаден в хранилището. Ако решите вашите данни да се показват в хранилището, трябва да настроите профила си на изследовател за публичен.



The screenshot shows the user profile page in DSpace CRIS. The top navigation bar includes links for Home, Communities & Collections, Research Outputs, Researchers, Organizations, Projects, and Help. The main content area is titled 'DSpace CRIS' and includes a brief description of the system. Below this, there is a 'Learn More' link. The user's profile is identified as 'Albanian DSpace-CRIS' and 'My DSpace: Hasani, Silvester'. The 'Researcher profile status' is set to 'public'. At the bottom, there are two buttons: 'Start a New Submission' and 'View Accepted Submissions'.

Фигура 4-d:Влезте с ORCID

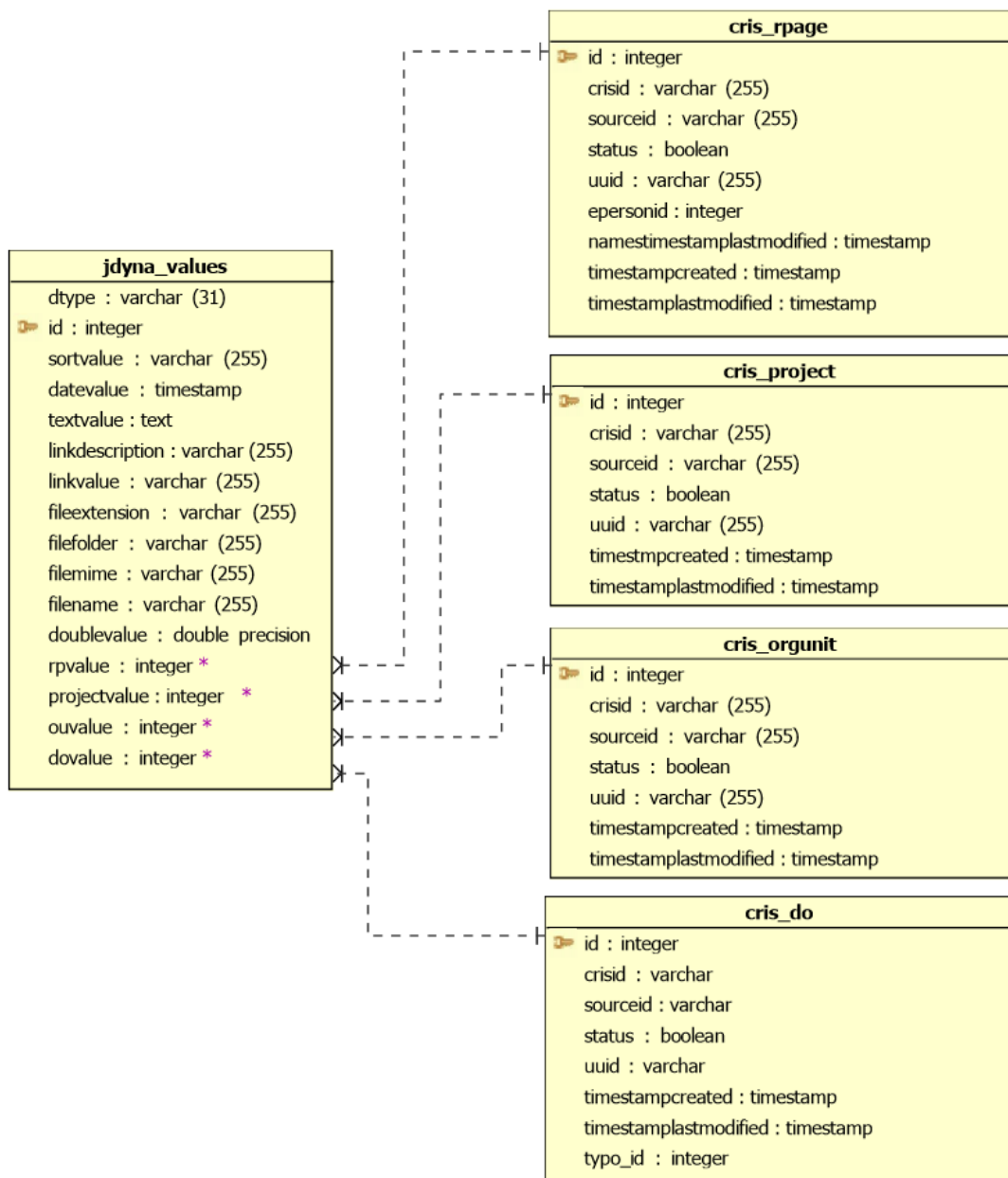
И накрая, както е показано по-горе, системата за цифрово хранилище на DSpace-CRIS е конфигурирана с интеграция на ORCID, където изследователите могат лесно да влизат, да променят и да претендират за своя изследователски профил с помощта на своя ORCID. След завършване на интеграцията, се позволява импортиране на база данни на

библиографски записи от Scopus и Web of Science. Албанският CRIS е конфигуриран така, че импортирането е в реално време на библиографските записи, търсещи във външната база данни или Web of Science и Scopus, се извършват с помощта на идентификатори, автори или заглавия на изследователския обект. След това се прави периодично сканиране на външната база данни за извличане на институционални публикации за импортиране. Това става чрез скриптовете, налични в системата за хранилища DSpace-CRIS, достъпни на официалния уебсайт на конфигурацията на системата [32].

4.4. Модел на данните, внедрен в Албанския-CRIS

Основната характеристика на DSpace-CRIS е фактът, че всяка структура от данни на обект може да бъде конфигурирана чрез потребителски интерфейс, чрез добавяне на прости и сложни полета. Могат да се създадат връзки между нови и стари обекти, конфигурирани в системата на хранилището. След като моделът на данни е конфигуриран, всички обекти имат публична страница, където може да се показва част или всяка информация, и те могат да бъдат търсени и преглеждани. Също така, всяка връзка, присъстваща в модела на данни на системата за хранилища, може да бъде проучена, тъй като те са свързани заедно. По този начин можете да преминете от профила на изследователя към организационни единици, проекти, публикации и обратно, като използвате връзката, създадена в системата на хранилището.

Фигура 4- е показва ER диаграмата на DSpace-CRIS, разположен в Албанския CRIS, който е ядрото по подразбиране на DSpace-CRIS. JDYNA_VALUES се използва за управление на всички взаимоотношения между всички изследователски обекти, конфигурирани в системата. CRIS_RPPAGE показва модела на данните и полетата на профила на изследователя, CRIS_PROJECT е моделът на данните за проектите, управлявани в системата за хранилища, CRIS_ORGUNIT е моделът на данните на организационните единици, а CRIS_DO е моделът на данни на изследователските обекти, който включва публикации, оборудване и всички други изследователски организации, свързани с разпространението на научноизследователска информация.



Фигура 4-е: Албански -CRIS основната структура

Всички области на всеки изследователски обект и общата структура осигуряват опростено прилагане на ключовите субекти и концепции на CERIF. Профилът на изследователя е субект Хора в CERIF формата за картографиране, Организационното звено е OrgUnit, а Публикациите са обектите на публикацията, част от DSpace-CRIS.

4.5. Политика за Отворена научна

В този документ е представена и първата политика за Отворена наука, която ще се прилага в Албания. Резултатите от проучването посочват, че в Албания не са приложени политики за отворен достъп. В сътрудничество с албанските университети е създадена политика, която трябва да се приложи, както е посочено по-долу. Целта на тази политика за отворен достъп е да следва посоката на Отворената наука в Албания. Тези университети в момента са в процес на прилагане на такава политика.

Предлаганата политика е в 5 подраздела, които определят общите ѝ принципи и след това преминават към различни аспекти на Отворената наука. Включва принципи на публикуване на отворен достъп, принципи на FAIR данни в изследването, принципи на използвания софтуер с отворен достъп и накрая се фокусира върху ролята и отговорностите на албанските изследователи. Като част от политиката е споменато и хранилище, което е цифровото хранилище, въведено в предишния раздел.

Политиката е насочена към Отворената наука като ключов елемент за увеличаване на въздействието на научните изследвания върху албанската общност. Очаква се тя да направи изследователските данни FAIR и да улесни достъпа до знания.

4.6. Изводи

В тази глава е предоставено описание на Албанското цифрово хранилище за отворена наука. Освен това е показано, че разработеното цифрово хранилище за отворена наука е напълно съвместимо със стандартите за метаданни на OpenAIRE и се поддържа от EOSC. Хранилището е достъпно на <https://albanian-cris.info> и в момента базата данни е в процес на запълване.

Всички албански институции, които ще бъдат въведени в базата данни на хранилището, са показани подробно и е завършен анализа на изискванията. След това се показва разгръщането на системата с целия необходим софтуер и се илюстрира конфигурацията, за да стане напълно функционална. Обяснява се и интеграцията с външни бази данни,

особено интеграцията с ORCID, която помага на изследователите да влизат лесно в системата.

И накрая, използваният модел на данни разчита на набор от основни обекти, дефинирани от модел на Общоевропейския информационен формат за научни изследвания CERIF. Там са показани взаимоотношенията на изследователските обекти. Показани са и полетата, включени в изследователските обекти.

Това дигитално хранилище с отворени ресурси, разположено в Албания, е първата стъпка към Отворената наука в страната. Албания е част от балканските страни, които все още правят първите си стъпки към Отворената наука. Албанският CRIS е хранилището, което ще помогне на Албанския Open Science Cloud да направи прехода към Open Science и да се присъедини към EOSC.

В заключение, политика на отворен достъп е разработена и в момента се приема от албанските университети, които подкрепят политиката за Отворена наука.

Глава 5. Заключение и бъдещи научни разработки

В тази глава са обобщени ключовите аспекти в този дисертационен труд. Първо се идентифицират етапите на прехода към Отворена наука и основните проблеми, пред които са изправени балканските университети. След това са представени следващите стъпки за развитие и внедряване на Отворена наука в Албания, приносите и списък с публикации, в които те са отразени. На финала се дискутират бъдещите стъпки, които могат да бъдат извършени в тази област.

5.1. Изводи

Основните приноси в това изследване са описани в главите и могат да бъдат разделени в следните основни групи.

Първо, описани са подробно ключовите елементи на концепцията за Отворена наука, започвайки от това какво е включва Отворената наука и достигайки до Отворен достъп. Отворените данни също са описани и сравнени с FAIR данните. Всички споменати концепции са тясно свързани с концепцията за софтуер с отворен код, затова тя също е разгледана подробно. Накрая, концепцията за цифровите хранилища е широко засегната, като ключов елемент при съхранението на данни от изследванията. Съхранението на данни от изследванията е важна задача за Отворена наука.

На второ място, се анализира европейското развитие на Отворената наука. Като се започне от Европейския отворен научен облак, който създава набор от насоки и правила, които трябва да се следват, за да направят данните за научните изследвания достъпни за всички изследователи от ЕС чрез оперативно съвместими данни FAIR. След това са сравнени най-използваните цифрови хранилища на изследователски данни. Накрая е показана посоката, която трябва да се следва за Отворената наука.

Трето, идентифицирани са етапите на прехода към Отворената наука и основните проблеми, пред които са изправени балканските университети. Фокусирайки се върху настоящото състояние на прехода към Отворената наука на балканските университети,

е разработен въпросник за балканските институции, проведено е изследване сред тях и са анализирани резултатите. Проучването даде възможност да се разбере, че Албания няма текуща информационна система за научни изследвания и е в ранните стъпки към отворената наука. Изследователите намират отворената наука за важна за бъдещето на тяхната страна, но нямат стимули. Следвайки случая на България, е създаден първият прототип на Албански облак за отворена наука (AOSC) и е разработен първият албански облак за отворена наука. DSpace-CRIS се използва за хранилището и достъпът до ORCID идентификатора е разрешен, за да позволи на всички изследователи свободен и лесен достъп до системата.

На финала е представена разработена политика за отворена наука, която да се прилага в албанските университети. Целта на тази политика за отворен достъп е да бъде следва тенденцията към Отворената наука в Албания. Някои албански университети в момента са в процес на преразглеждане и прилагане на такива политики. Политиката е насочена към Отворената наука и я вижда като ключов елемент за увеличаване на въздействието на научните изследвания върху албанската общност. Крайната цел на първата политика в областта на Отворената наука в Албания е да събере данните от изследванията за FAIR и да улесни достъпа до тези знания от всички албански изследователи.

5.2. Приноси

В тази секция са обобщени приносите на дисертационния труд.

5.2.1. Научни приноси

1. Изследвана е подробно областта на Отворената наука и всички нейни съставни елементи. Направеното изследване може да служи като изчерпателно ръководство за Отворена наука.
2. Предложена архитектура на прототип на албанския облак за отворена наука, който може да помогне на Албания да се присъедини по-бързо към Европейския облак за отворена наука. Той може да осигури Отворена наука за всички изследователски институции в страната. Този принос е представен в статия на конференцията ICERI2020 [33].

5.2.2. Приложни приноси

1. Разработено е първото албанско отворено цифрово хранилище за изследователски данни. То допринася за албанския облак за отворена наука, като предоставя FAIR услуги на изследователите. Представена статия на конференцията ICAI 2021 за първите стъпки в Албания към отворена наука [34].

5.2.3. Научноприложни приноси

1. Подготвено, разпространено и анализирано проучване в България и Албания, за да се установи текущото състояние на Отворената наука в региона. Проучването помогна да се направи извод, че Албания е в ранните етапи на отворената наука. Резултатите от проучването са представени в статия на конференцията ICAI2020 [35].
2. Разработена политика за отворен достъп за приемане на Отворената наука в Албания. Политиката може да помогне за ускоряване на прехода към отворена наука за всички изследователски институции в страната. Тя е представена в статия на конференцията EDULEARN21 [36].

5.3. Научни публикации, свързани с дисертационния труд

1. **S. Hasani**, E. Stefanova, K. Stefanov, A. Georgiev, ARE WE READY FOR OPEN SCIENCE - THE ANSWER OF THE BALKAN UNIVERSITIES, ICERI2020 Proceedings, Publisher: IATED, 2020, pages: 1947-1953, ISSN (print): 2340-1095, ISBN: 978-84-09-24232-0, doi: 10.21125 / iceri.2020.0481, International 2020 [33].
2. **S. Hasani**, E. Stefanova, A. Georgiev, K. Stefanov, Current State of Open Science in Balkan Universities, Proceedings of the International Conference Automatics and Informatics (ICAI2020), Publisher: IEEE, 2020, pages: 1-6, ISBN: 978-1-7281-9309-0, doi: 10.1109 / ICAI50593.2020.9311337, Ref, IEEE Xplore, International, PhD [35].
3. **S. Hasani**, E. Stefanova, A. Georgiev, K. Stefanov, OPEN SCIENCE IN ALBANIA: FIRST OPEN-SCIENCE POLICY IMPLEMENTED IN ALBANIA, EDULEARN21 Proceedings, Conference name: 13th International Conference on Education and New Learning Technologies, Dates: 5-6 July, 2021, Pages: 12267-12271, ISBN: 978-84-09-31267-2, ISSN: 2340-1117, doi: 10.21125/edulearn.2021.2575 [36].
4. **S. Hasani**, E. Stefanova, A. Georgiev and K. Stefanov, "First Steps towards Open Science in Albania" 2021 International Conference Automatics and Informatics (ICAI), Varna, Bulgaria, 2021, pp. 235-238, doi: 10.1109/ICAI50593.2020.9311337, ISBN: 978-1-6654-2662-6, e-ISBN: 978-1-6654-2661-9, 10.1109/ICAI52893.2021.9639622 [34].

5.4. Бъдещо развитие

Резултатът от работата върху тази дисертация може е основа за бъдещо развитие.

Най-важно е развитието на цифровото хранилище с отворен достъп. То е в процес на разгръщане и все още се нуждае от много внимание. Не само това, но тъй като системата ще се разширява, има нужда от обработка на метаданните за изследванията.

Второ, когато системата стане много по-използваема, ще е необходима миграция към по-добър сървър. В момента системата е внедрена в двуядрен процесор само с 2GB RAM. В момента, в който системата се нуждае от много едновременен потребителски достъп, може да се наложи по-добра система за обработка на всички заявки към сървъра.

На последно място, разгърнатото хранилище и политиката за отворени науки се оценяват от различни университети в Албания като лични инициативи. В бъдеще сътрудничеството с Министерството на образованието, спорта и младежта в Албания ще помогне на страната по-скоро да се присъедини към EOSC.

Библиография

- [1] M. Woelfle, P. Olliaro, and M. H. Todd, "Open science is a research accelerator," *Nat. Chem.*, vol. 3, no. 10, Art. no. 10, Oct. 2011, doi: 10.1038/nchem.1149.
- [2] R. Vicente-Saez and C. Martinez-Fuentes, "Open Science now: A systematic literature review for an integrated definition," *J. Bus. Res.*, vol. 88, pp. 428–436, Jul. 2018, doi: 10.1016/j.jbusres.2017.12.043.
- [3] "What is Open Science?," *ORION Open Science*, Sep. 27, 2017. <https://www.orion-openscience.eu/resources/open-science> (accessed May 05, 2020).
- [4] Peter Suber, "Open Access Overview," May 04, 2020. <http://legacy.earlham.edu/~peters/fos/overview.htm> (accessed May 04, 2020).
- [5] "What is open access?," May 04, 2020. <https://www.openaccess.nl/en/what-is-open-access> (accessed May 04, 2020).
- [6] "The Open Definition - Open Definition - Defining Open in Open Data, Open Content and Open Knowledge," May 06, 2020. <https://opendefinition.org/> (accessed May 06, 2020).
- [7] "What is open?," May 06, 2020. <https://okfn.org> (accessed May 06, 2020).
- [8] "FAIR Principles," *GO FAIR*, Jun. 15, 2020. <https://www.go-fair.org/fair-principles/> (accessed Jun. 15, 2020).
- [9] "Metadata | Definition of Metadata by Merriam-Webster," Jun. 15, 2020. <https://www.merriam-webster.com/dictionary/metadata> (accessed Jun. 15, 2020).
- [10] "ORCID," Jun. 15, 2020. <https://orcid.org/> (accessed Jun. 15, 2020).
- [11] "The Open Source Definition | Open Source Initiative," May 04, 2020. <https://opensource.org/osd> (accessed May 04, 2020).
- [12] "Digital Repository - an overview | ScienceDirect Topics," May 05, 2020. <https://www.sciencedirect.com/topics/computer-science/digital-repository> (accessed May 05, 2020).
- [13] M. Smith *et al.*, "DSpace: An Open Source Dynamic Digital Repository," *-Lib Mag.*, vol. 9, no. 1, Art. no. 1, Jan. 2003, doi: 10.1045/january2003-smith.
- [14] T. Staples, R. Wayland, and S. Payette, "The Fedora Project: An Open-source Digital Object Repository Management System," *-Lib Mag.*, vol. 9, no. 4, Apr. 2003, doi: 10.1045/april2003-staples.

- [15] mattste, "European Open Science Cloud," *Shaping Europe's digital future - European Commission*, Oct. 28, 2016. <https://ec.europa.eu/digital-single-market/en/european-open-science-cloud> (accessed Jun. 18, 2020).
- [16] "EOSC," *EOSC Portal*, Jul. 04, 2018. <https://eosc-portal.eu/about/eosc> (accessed Mar. 14, 2022).
- [17] "EOSC Launch," May 04, 2020. <https://eosc-launch.eu/home/> (accessed May 04, 2020).
- [18] European Commission, "EOSC Strategic Implementation Roadmap," *EOSC Strategic Implementation Roadmap*, May 2018. https://ketlib.lib.unipi.gr/xmlui/bitstream/handle/ket/1170/eosc_strategic_implementation_roadmap_large.pdf?isAllowed=y&sequence=1 (accessed Dec. 28, 2020).
- [19] A. Georgiev and K. Stefanov, "Bulgarian Open Science Digital Library - First Prototype," in *Digital Presentation and Preservation of Cultural and Scientific Heritage. Conference Proceedings*, Burgas, Bulgaria, Sep. 2019, p. 8. [Online]. Available: https://dipp.math.bas.bg/images/2019/251-258_8_3.8_sDiPP2019-66_f_v.1.F_20190908.pdf
- [20] J. Coombs, "LibGuides: Research Data Management: RDM Explained," Jun. 19, 2020. <https://library.dmu.ac.uk/rdmguide/rdmexplained> (accessed Jun. 19, 2020).
- [21] "DSpace-CRIS Home - DSpace-CRIS - LYRASIS Wiki," Jun. 19, 2020. <https://wiki.lyrasis.org/display/DSPACECRIS/> (accessed Jun. 19, 2020).
- [22] "Home | DSpace-CRIS DEMO," May 04, 2020. <https://dspace-cris.4science.it/> (accessed May 04, 2020).
- [23] "InvenioRDM — inveniosoftware.org," Jun. 19, 2020. <https://inveniosoftware.org/products/rdm/#faq> (accessed Jun. 19, 2020).
- [24] "Zenodo - Research. Shared.," Jun. 19, 2020. <https://about.zenodo.org/infrastructure/> (accessed Jun. 19, 2020).
- [25] "Fedora - The Flexible, Modular, Open-Source Repository Platform," *Duraspace.org*, Jun. 19, 2020. <https://duraspace.org/fedora/> (accessed Jun. 19, 2020).
- [26] P. Stanchev and K. Stefanov, "Bulgarian Open Science Cloud," in *Digital Presentation and Preservation of Cultural and Scientific Heritage. Conference Proceedings*, Burgas, Bulgaria, 2019, vol. 9, p. 6. [Online]. Available:

https://dipp.math.bas.bg/images/2019/259-264_6_3.9_sDiPP2019-65_f_v.1.F_20190908.pdf

- [27] “Main features of CERIF | euroCRIS,” Jun. 20, 2020. <https://www.eurocris.org/cerif/main-features-cerif> (accessed Jun. 20, 2020).
- [28] “NI4OS- Europe – National Initiatives for Open Science in Europe,” Jul. 08, 2020. <https://ni4os.eu/> (accessed Jul. 08, 2020).
- [29] A. Bodlos *et al.*, “EOSC Pillar ‘National Initiatives’ Survey (SUF edition).” AUSSDA, 2019. doi: 10.11587/VOSVGK.
- [30] B. Saenen, R. Morais, V. Gaillard, L. Borrell-Damián, and F. Tobon, “Data for Research Assessment in the Transition to Open Science. 2019 EUA Open Science and Access Survey Results.” Zenodo, Jan. 07, 2020. doi: 10.5281/ZENODO.3600122.
- [31] “Home | ICT,” Jul. 08, 2020. <https://npict.bg/> (accessed Jul. 08, 2020).
- [32] “Home | Albanian DSpace-CRIS.” <https://albanian-cris.info/> (accessed Oct. 06, 2021).
- [33] S. Hasani, E. Stefanova, K. Stefanov, and A. Georgiev, “ARE WE READY FOR OPEN SCIENCE – THE ANSWER OF THE BALKAN UNIVERSITIES,” in *ICERI2020 Proceedings*, Online Conference, Nov. 2020, pp. 1947–1953. doi: 10.21125/iceri.2020.0481.
- [34] S. Hasani, E. Stefanova, A. Georgiev, and K. Stefanov, “First steps towards Open Science in Albania,” in *2021 International Conference Automatics and Informatics (ICAI)*, Sep. 2021, pp. 235–238. doi: 10.1109/ICAI52893.2021.9639622.
- [35] S. Hasani, E. Stefanova, A. Georgiev, and K. Stefanov, “Current State of Open Science in Balkan Universities,” in *2020 International Conference Automatics and Informatics (ICAI)*, Varna, Bulgaria, Oct. 2020, pp. 1–6. doi: 10.1109/ICAI50593.2020.9311337.
- [36] S. Hasani, E. Stefanova, K. Stefanov, and A. Georgiev, “OPEN SCIENCE IN ALBANIA: FIRST OPEN-SCIENCE POLICY IMPLEMENTED IN ALBANIA,” *EDULEARN21 Proc.*, pp. 12267–12271, 2021.

Декларация за оригиналност

Декларирам, че представената дисертация на тема " УСЛУГИ НА ИЗСЛЕДОВАТЕЛСКИ ИНФОРМАЦИОННИ СИСТЕМИ ЗА ОБЛАК ЗА ОТВОРЕНА НАУКА " е мой труд, създаден по време на докторантската ми програма в Софийския университет "Св. Климент Охридски" в периода от 2019 до 2022 година.