

ПРЕПОРЪКИ ЗА РАБОТА С БЪЛГАРСКИЯ ПОРТАЛ ЗА ОТВОРЕНА НАУКА И НАЦИОНАЛНОТО ХРАНИЛИЩЕ КЪМ НЕГО

Настоящият документ съдържа в обобщен вид основна информация и препоръки за работа с Българския портал за отворена наука¹ и националното хранилище² към него като част от инициативата за развитие на отворената наука в България.

ОБЩА ИНФОРМАЦИЯ ЗА СИСТЕМАТА

Българският портал за отворена наука и националното хранилище към него са част от информационна система, която се поддържа от министъра на образованието чрез Националния център за информация и документация на основание чл. 2, ал. 24 от Правилника за устройството и дейността на НАЦИД в изпълнение на Препоръка (ЕС) 2018/790 на Комисията от 25 април 2018 година относно достъпа до научна информация и нейното съхранение.

Българският портал за отворена наука осигурява единна входна точка за лесен и бърз достъп до научна информация и резултати от научни изследвания, финансирани с публични средства, включително до ресурси на институционални хранилища с отворен достъп в Република България.

Хранилището е създадено, за да подпомогне най-вече учени и институции, които не разполагат със свои собствени хранилища. В него те могат да съхраняват и разпространяват научни публикации и резултати от своите изследвания. Към него могат да се свързват и автоматично да се публикуват и данни от други външни системи – вече изградени институционални хранилища.

Системата осигурява свързаност с европейските е-инфраструктури за отворен достъп до научна информация. При спазване на последните стандарти за

¹ <https://bpos.bg/>

² <https://pub.bpos.bg/login>

оперативна съвместимост, информацията от Българския портал за отворена наука се индексира от OpenAIRE³. Изграждането на Портала и свързването с OpenAIRE е важна стъпка към активното приобщаване на българската научна общност към Европейския облак за отворена наука (EOSC)⁴.

Българският портал за отворена наука и националното хранилище са изградени при спазване на указанията и насоките на OpenAIRE⁵, които представляват опростени спецификации на метаданни, спазването на които гарантира съвместимост на хранилището с OpenAIRE. В изпълнение на тези насоки хранилището е регистрирано в глобалната директория за академични хранилища с отворен достъп OpenDOAR⁶.

Ресурсите в хранилището се депозират при спазване на FAIR принципите за свободно разпространение на научноизследователски резултати и данни, според които научните ресурси следва да бъдат:

- Откриваеми (Findable) – описани с подробни метаданни и постоянни идентификатори;
- Достъпни (Accessible) – за машини и хора по стандартен протокол чрез постоянните си идентификатори, при съответните ограничения;
- Оперативно съвместими (Interoperable) – използвани са стандартни формати и схеми за данни и метаданни, което позволява обмена им между различните системи;
- Преизползваеми (Reusable) – с ясно обозначен лиценз и права за достъп.

ПРЕПОРЪКИ ЗА ОРГАНИЗИРАНЕ И ОПИСАНИЕ НА НАУЧНАТА ИНФОРМАЦИЯ

Препоръките представляват практически насоки, почиват на FAIR принципите, проучените добри практики и препоръки от утвърдени европейски инфраструктури като OpenAIRE и инициативи като EOSC и GOFAIR.

Целта им е да подпомогнат висшите училища и научните организации да организират научната си информация така, че да бъде постигната оперативна съвместимост между системите на национално и наднационално ниво и да се осигури безпроблемното събиране на метаданните и индексиране на публикациите от институционалните хранилища.

³ <https://www.openaire.eu/>

⁴ <https://www.eosc-portal.eu/about/eosc>

⁵ <https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/latest/index.html>

⁶ <https://v2.sherpa.ac.uk/opensoar/>

1. Всеки ресурс в хранилището има постоянен идентификатор

Постоянните идентификатори осигуряват постоянна, непроменяща се с времето препратка към стабилен в дългосрочен план ресурс. Важно е идентификаторът да е уникален, т.е. да няма възможност да бъде повторен и назначен към друг ресурс, и постоянен, т.е. да осигурява дългосрочна връзка и откриваемост на ресурса. Най-широко използвани са идентификатори като DOI и Handle. Националното хранилище поддържа функционалност за създаване на уникални идентификатори за публикуваните обекти и назначава идентификатори във формат Handle.Net.

Препоръчително е всяко институционално хранилище да поддържа функционалност за служебно назначаване на постоянен идентификатор на ресурсите, които се депозират при тях. Когато даден ресурс има предварително назначен идентификатор (например защото вече е публикуван в електронно списание), в институционалното хранилище се посочва вече назначеният от списанието и не се присвоява нов.

Назначаваните идентификатори от институционалните хранилища следва да отговарят на посочените по-горе изисквания и да са работещи, за да може да бъде осъществено правилно селективното събиране на метаданни /harvesting/ от Българския портал за отворена наука и индексирането им от OpenAIRE.

2. Ресурсите са описани с възможно най-подробни метаданни, като за всеки ресурс е налице задължителния минимален набор от метаданни

При депозиране на ресурс задължително се посочват:

- **автор/и** (имена и персонален идентификатор с оглед еднозначното разграничаване между отделните учени и сигурна връзка между тях и научните им резултати);
- **заглавие**;
- **дата на публикуване** (задължително се посочва година на публикуване);
- **тип ресурс** (напр. научна статия, част от книга, доклад от конференция, рецензия, отчет и други);
- **идентификатор на ресурса** (напр. DOI, Handle);
- **права за достъп** (отворен достъп, достъп само до метаданни, ембарго период, ограничен достъп).

Силно препоръчително е да бъдат попълнени **възможно най-подробни метаданни** за ресурса, включително: **източника на финансиране, описание, ключови думи и лиценз** (напр. *Creative Commons Attribution 4.0 International*).

Подробните метаданни повишават видимостта на научните резултати, правят ги лесно откриваеми и увеличават възможностите за използването и преизползването им. Попълването на минималния набор от задължителни метаданни позволява тяхното публикуване в портала и лесно индексирание от OpenAIRE.

3. Метаданните и данните следват утвърдени стандартни схеми и речници

Метаданни за ресурсите в националното хранилище се описват съгласно схемите Dublin Core⁷ и DataCite Metadata Schema⁸ при съобразяване с OpenAIRE Guidelines for Literature Repository Managers.

Използвани са COAR Controlled Vocabularies⁹, които представляват набори от термини и фрази, предназначени за индексирание и/или извличане на съдържание. Чрез тези речници се осигурява еднозначност на използваните термини - депозиращият и извличащият информация, както и свързаните системи, влагат едно и също значение в един и същ термин.

Препоръката за използване на утвърдени стандартни схеми и речници е свързана с изискванията за оперативната съвместимост, която позволява системите да разчитат и обменят информация помежду си без специализирани ad hoc алгоритми, преводи и речници за съответствие (mapping).

4. Уеднаквяване на използваната терминология за описване на типовете ресурси от всички хранилища и е-инфраструктури

С оглед улесняване на събирането на метаданни и индексирането от OpenAIRE е препоръчително уеднаквяването на използваните термини от всички хранилища и е-инфраструктури.

Типовете ресурси в номенклатурата на националното хранилище са стойности от COAR Resource Type Vocabulary¹⁰, като някои от най-използваните типове са: journal article (статия в списание), doctoral dissertation (докторска дисертация), report to funding agency (отчет до финансираща организация) , research article (научна статия), book (книга), book part (част от книга), review article (обзорна статия).

⁷ <https://www.dublincore.org/specifications/dublin-core/dces/>

⁸ <https://schema.datacite.org/meta/kernel-4.1/>

⁹ <https://www.coar-repositories.org/news-updates/what-we-do/controlled-vocabularies/>

¹⁰ http://vocabularies.coar-repositories.org/documentation/resource_types/

5. Насърчават се регистрирането на уникален персонален идентификатор на всеки учен и обозначаването му в метаданните към депозираните ресурси

Един от най-използваните видове персонални идентификатори е ORCID ID¹¹ - безплатен постоянен цифров идентификатор, който осигурява еднозначна връзка между учения и резултатите от научната му дейност. Други примери за такива идентификатори са WoS ResearcherID, ScopusID, ISNI.

6. Ресурсите са с ясно обозначени права за достъп и лиценз

Насърчава се публикуването с отворен достъп и лиценз Creative Commons Attribution 4.0¹² (CC BY), които в максимална степен съответстват на концепцията за отворен достъп до научна информация.

Посочването на правата за достъп и лиценза за използване на ресурсите е от съществено значение. Неяснота по отношение на тези два аспекта може сериозно да възпрепятства използването и преизползването на ресурсите и да създаде затруднения при индексирването им от други системи и инфраструктури, което от своя страна да намали видимостта на научните резултати.

7. Метаданните се извличат по стандартен протокол за комуникация

Протоколът следва да е отворен, безплатен, универсално изпълним и да позволява автентификация и оторизация, когато е необходимо.

Българският портал за отворена наука поддържа функционалност за селективно събиране на метаданни (selective harvesting) от институционални хранилища с отворен достъп в Република България, които поддържат протокола The Open Archives Initiative Protocol for Metadata Harvesting¹³. За всяко включено институционално хранилище се определят:

- адрес на OAI-PMH услугата (OAI-PMH URL);
- metadataPrefix – формат, в който да бъдат върнати данните, напр. oai_dc;
- setSpec и setName – идентификатор и име на колекция, към което принадлежат обектите, които трябва да бъдат събрани („ожънати“). Ще бъдат „ожънати“ само обектите, принадлежащи на посочената колекция и нейните подколекции;
- период на „жънене“.

¹¹ <https://orcid.org/register>

¹² <https://creativecommons.org/licenses/by/4.0/>

¹³ <https://www.openarchives.org/OAI/openarchivesprotocol.html>

При определяне на колекциите от институционалните хранилища следва да имате предвид, че в националното хранилище се съхраняват публикации, одобрени за публикуване или вече публикувани (accepted manuscripts, published versions), които са създадени от учени. Материали, създадени от студенти по време на обучението им, като дипломни и курсови работи, не попадат в обхвата на националното хранилище.

РОЛЯ НА УНИВЕРСИТЕТИТЕ И НАУЧНИТЕ ОРГАНИЗАЦИИ

Насочена изцяло в полза на научноизследователска дейност в България и преодоляване на предизвикателствата при провеждане на научни изследвания, успехът на системата и инициативата за отворена наука зависи от активността и колективното усилие на цялата научноизследователска общност.

Университетите и научните организации със съдействието на държавата имат важна роля в процеса по приобщаване на учените към инициативата за отворена наука, създаване на предпоставки и насърчаване на публикуването с отворен достъп. Учените и изследователите следва да бъдат насърчавани да качват своите резултати от научни изследвания, финансирани изцяло или частично с публични средства в сигурни хранилища с отворен достъп във възможно най-кратък срок след публикуване, или да проучват възможностите да публикуват резултатите си в издания с отворен достъп.

Залагането на отворения достъп, споделянето и повторното използване на резултатите от научни изследвания, включително ползите от тях, в политиките на национално и институционално ниво е първа стъпка към насърчаването на отворения достъп до научни информация и превръщането му в стандартна практика при провеждане на научни изследвания.