



## ПРЕДЛОЖЕНИЕ

за провеждане на докторантура в редовна форма на обучение  
по реда на чл. 21, ал. 7 от ЗВО

в Института по математика и информатика при Българска академия на науките

професионално направление: 4.6. Информатика и компютърни науки  
докторска програма: „Информатика“

тема на докторантурата:

### **Инфраструктури и отворени решения за колекциите като данни в българските институции за културно наследство**

Предложил темата: доц. д-р Милена Добрева

## Актуалност

Концепцията Collections as Data предлага фундаментална промяна в начина, по който културните институции организират своите дигитални сбирки - не само като обекти за разглеждане и съхранение, а като данни, които могат да бъдат анализирани, обработвани и повторно използвани чрез изчислителни методи. Тази промяна на парадигмата поставя нови изисквания към институциите - по отношение на качеството на метаданните, стандартизацията, правните условия за достъп и повторна употреба, и техническата инфраструктура за публикуване.

В България GLAM институциите (галерии, библиотеки, архиви, музеи) разполагат с богати колекции, но нивото на дигитална зрелост варира значително. Много от тях поддържат частично дигитализирани каталози с хетерогенни формати на метаданни, неясни или рестриктивни лицензионни режими и ограничена техническа инфраструктура за публикуване на отворени данни. Липсва задълбочен анализ на готовността на институциите за преход към модела Collections as Data, както и стратегия за този преход.

На европейско ниво нормативната рамка създава благоприятна среда - Директивата за отворени данни и повторна употреба на информация от обществеността (EU 2019/1024), Регламентът за данните (Data Act, 2023), препоръката на Комисията за обща европейска рамка за данни за културното наследство (2021), както и инициативи като EOSC, European Data Space for Cultural Heritage и Европейския облак за отворена наука - но прилагането им на институционално ниво в България остава фрагментирано и недостатъчно изследвано. Паралелно, Wiki-екосистемата (Wikidata, Wikimedia Commons) се утвърждава като де факто инфраструктура за свързани отворени данни за културно наследство, а националните и европейските портали за отворени данни (data.government.bg, data.europa.eu) предлагат платформи за публикуване, които остават слабо използвани от GLAM институциите.

Същевременно нарастващото приложение на изкуствен интелект в културния сектор поставя нови изисквания към качеството, структурата и достъпността на институционалните данни. ИИ е едновременно инструмент за обработка и обогатяване на данните (автоматична класификация, разпознаване на обекти, генериране на метаданни) и основен консуматор на добре структурирани свързани отворени данни - което прави въпроса за институционалната готовност за отваряне още по-належащ.

Предложената докторантура е пряко свързана с приоритетите на проекта FOCUS, по-специално с направленията „Datafication of cultural heritage” и „Open knowledge infrastructures”. Докато други изследвания в рамките на проекта се фокусират върху техническите аспекти на интеграцията и обогатяването на данни, тази докторантура адресира институционалната, правната и организационната страна на процеса - без която техническите решения не могат да бъдат устойчиво внедрени.

## Цел на докторантурата

Целта на докторантурата е разработването на методологическа рамка за оценка на готовността на GLAM институции за отваряне и публикуване на данни за културно наследство, както и изработването на експериментални технологични решения, които подпомагат тоди преходформулиране на пътна карта за прехода към модела Collections as Data.

## Научни задачи и методи за изпълнение

- Разработване на модел за оценка на институционалната готовност (maturity model), включващ измерения като: качество и пълнота на метаданните, използвани стандарти, ниво на дигитализация, техническа инфраструктура, организационен капацитет и политики за достъп;
- Оценка на качеството на данните в избрани институции чрез приложение на разработения модел, включително анализ на пълнота, съгласуваност, гранулярност, интероперабилност и съответствие с принципите FAIR;
- Идентифициране на бариери и катализатори за отварянето на институционални данни - технически, организационни, правни и културни;
- Изследване на интеграцията с Wiki-екосистемата (Wikidata, Wikimedia Commons) като стратегия за отваряне и повишаване на видимостта на институционалните колекции, включително анализ на успешни практики от европейски GLAM институции;
- Оценка на готовността за публикуване чрез национални и европейски портали за отворени данни (data.government.bg, data.europa.eu), включително анализ на техническите и организационни предпоставки за устойчиво публикуване;
- Анализ на предпоставките за приложение на ИИ в GLAM институциите - качество и пълнота на данните като необходимо условие за ефективно използване на ИИ инструменти за обработка, обогатяване и достъп до колекции;
- Валидиране на рамката чрез пилотно прилагане в конкретни институции и анализ на резултатите с участието на заинтересовани страни.
- Подбор и евентуално разработка на допълнителни софтуерни инструменти, които подпомагат прехода от дигитални сбирки към Collection as Data.

## Очаквани резултати

В резултат от изследванията се очакват:

- модел за оценка на институционалната готовност за Collections as Data (maturity model), приложим към GLAM институции с различен профил и мащаб, вкл.на правни аспекти;
- емпирични данни от оценка на конкретни институции, илюстриращи приложимостта на модела и разнообразието от институционални ситуации;
- модел за преход на институции с натрупано дигитално съдържание към предлагане на Collections as Data;
- анализ на възможностите за интеграция с Wiki-екосистемата и портали за отворени данни като канали за публикуване и видимост;
- софтуерни инструменти, които подпомагат културните институции за прехода към Collections as Data.

## Въздействие

Очакваните резултати ще допринесат пряко за подобряването на институционалния капацитет за публикуване на отворени данни за културно наследство в България. Моделът за оценка и пътната карта ще предоставят на институциите практически инструменти за планиране и изпълнение на прехода към Collections as Data.

На политическо ниво, анализът на правната рамка и идентифицираните бариери ще подпомогнат формулирането на по-ефективни политики за отворени данни в културния сектор, в контекста на европейските инициативи за общо европейско пространство за данни за културното наследство.

От информатична гледна точка, разработеният maturity model и метриците за оценка на качеството на данни имат потенциал за приложение и в други предметни области, където институционалните данни преминават през процес на отваряне.

В по-широк план, резултатите ще подпомогнат интеграцията на българското културно наследство в европейски и международни инфраструктури (EOSC, Europeana, Wikidata) и ще подкрепят прилагането на принципите на отворената наука в хуманитарните и социалните науки. Изследването ще допринесе и за подготовката на институциите за ефективно използване на ИИ технологии - като осигури необходимата база от качествени, стандартизирани и достъпни данни, без която потенциалът на изкуствения интелект в културния сектор не може да бъде реализиран.

## Място на зачисляване

Секция „Софтуерни технологии и информационни системи” (ИМИ-БАН)

## Използвана научна инфраструктура

КЛаДА-БГ

HEMUS supercomputer