

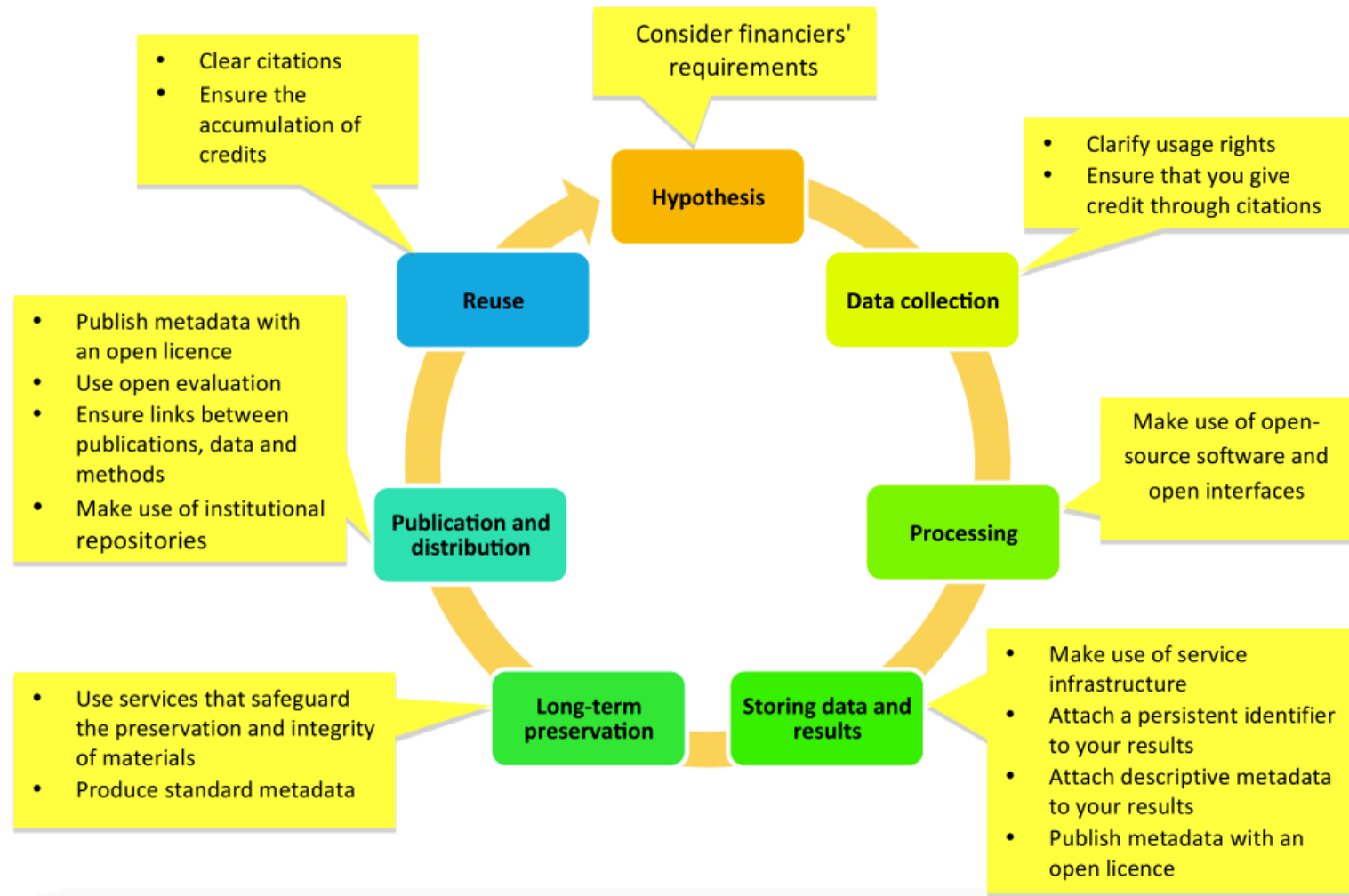
# Работа со складишта за отворени податочни множества

Capacity Building Training  
Skopje, 15.10.2021

Anastas Mishev, UKIM



- ❑ Основи принципи на отворената наука
- ❑ Складишта на отворени податочни множества како основа за проверлива и повторлива наука
- ❑ Типови на складишта на податоци
- ❑ Примери на складишта на отворени податочни множества
- ❑ Податочно складиште на репозиториумот на УКИМ
  - ❑ Практичен пример за поставување на податочно множество
- ❑ Иднината на научните публикации
- ❑ Референци



<https://www.fosteropenscience.eu/sites/default/files/pdf/3986.pdf>

- ❑ Складишта на отворени податочни множества се клучни за проверлива и повторлива наука
- ❑ Складиште на податоци или податочни множества е генерички термин кој означува архива во која се зачувани податоци за анализа, дистрибуција или известување
- ❑ Можат да бидат во строго тесна научна област или мултидисциплинарни
- ❑ Квалитет на податоците
  - ❑ Начини на проверка на квалитетот
- ❑ Опис на податоците
  - ❑ клучен за нивно повторно искористување
  - ❑ “data curation”
- ❑ Неопходен чекор кај одредени издавачи
  - ❑ Препорачан или овозможен кај други

“ - Researchers, on the whole, do not do replication tests as part of their own work, or even as part of the peer review process. In the future, they may be incentives for researchers to do so, and post-publication [crowd-sourced peer review](#) in the mold of Wikipedia, as promoted by Edward Curry, may prove to be a successful model.

“ - Academic institutions, and their libraries, are increasingly involved in the data management process, but are not involved in replication as a matter of course (note [some calls](#) for [libraries](#) to take a more [active](#) role in this regard).

- ❑ Складиштата на податоци, уште наречени и податочни библиотеки или податочни архиви, можат да бидат од различен тип
- ❑ Data warehouse – големи складишта кои најчесто складираат податоци од различни извори или сегменти, каде податоците можат да бидат и неповрзани меѓусебно.
- ❑ Data lake – големи складишта на неструктурирани податоци опишани со соодветни мета-податоци
- ❑ Data marts – подмножества од data warehouse кои содржат податоци релевантни за одредена област
- ❑ Складишта на мета податоци - содржат информации за модели на податоци, описи на изворите, значење на податоците итн.
- ❑ Data cubes – повеќедимензионални податоци складирани во табели, најчесто содржат временски секвенци од податоци.

## Projected figures 2025

---



**530%**

increase of global  
data volume  
From 33 zettabytes  
in 2018 to 175  
zettabytes



**€829  
billion**

value of data  
economy in the  
**EU27**  
From €301 billion  
(2.4% of EU GDP)  
in 2018



**10.9  
million**

data  
professionals in  
the **EU27**  
From 5.7 million in  
2018



**65%**

Percentage of EU  
population with  
basic digital skills  
From 57% in 2018

# Примери за складишта на податоци

## General repositories

- [4TU.ResearchData](#)
- [ANDS contributing repositories](#)
- [Dryad Digital Repository](#)
- [Figshare](#)
- [Harvard Dataverse](#)
- [Mendeley Data](#)
- [Open Science Framework](#)
- [Science Data Bank](#)
- [Zenodo](#)
- [Code Ocean](#)

## Taxonomic and Species Diversity

- [Global Biodiversity Information Facility \(GBIF\)](#)
- [MycoBank](#)
- [NCBI Taxonomy](#)
- [Integrated Taxonomic Information System \(ITIS\)](#)

## Biomedical Sciences

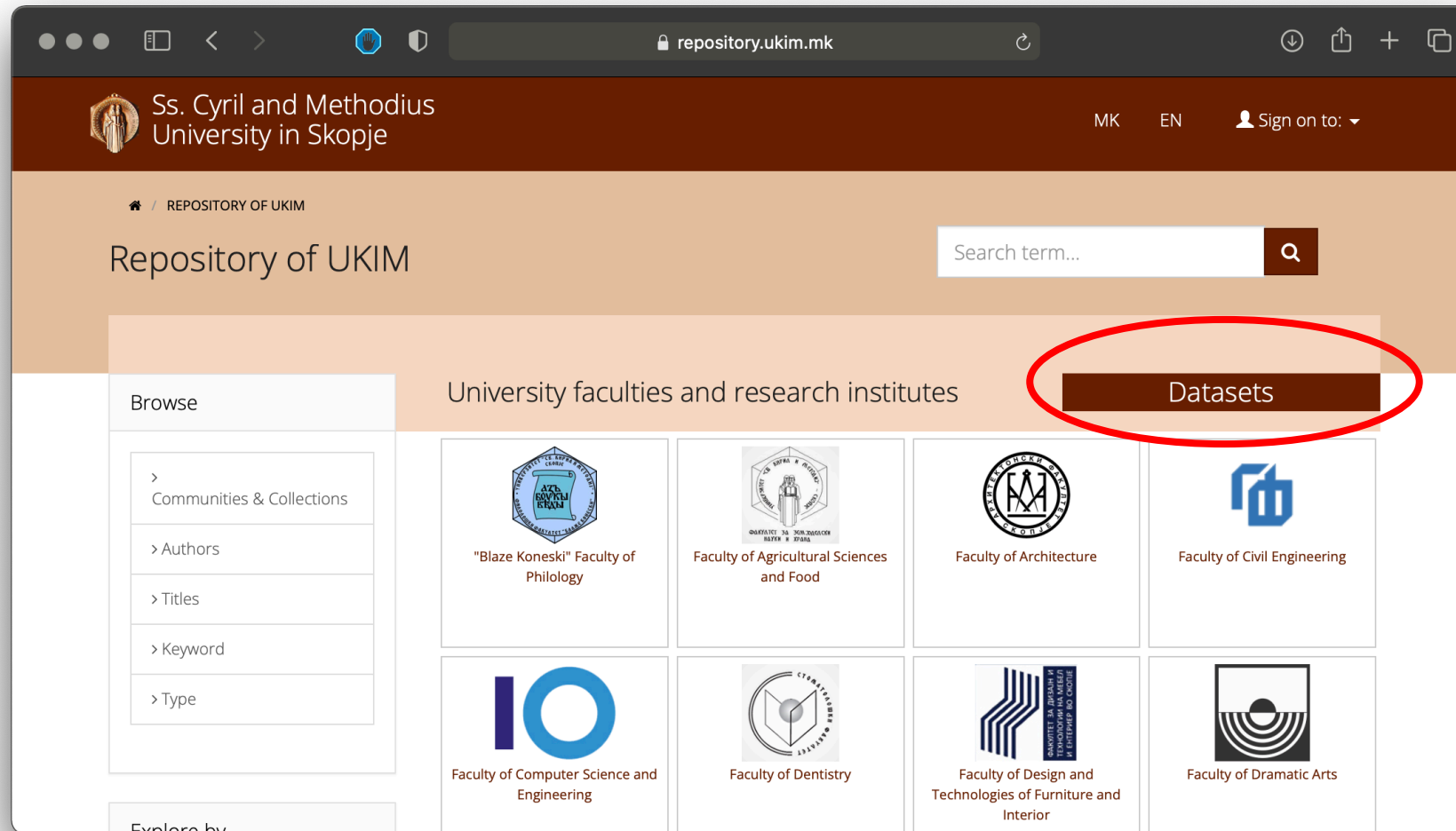
- [Eukaryotic Pathogen Database Resources \(EuPathDB\)](#)
- [Influenza Research Database \(IRD\)](#)
- [The Cancer Imaging Archive \(TCIA\)](#)

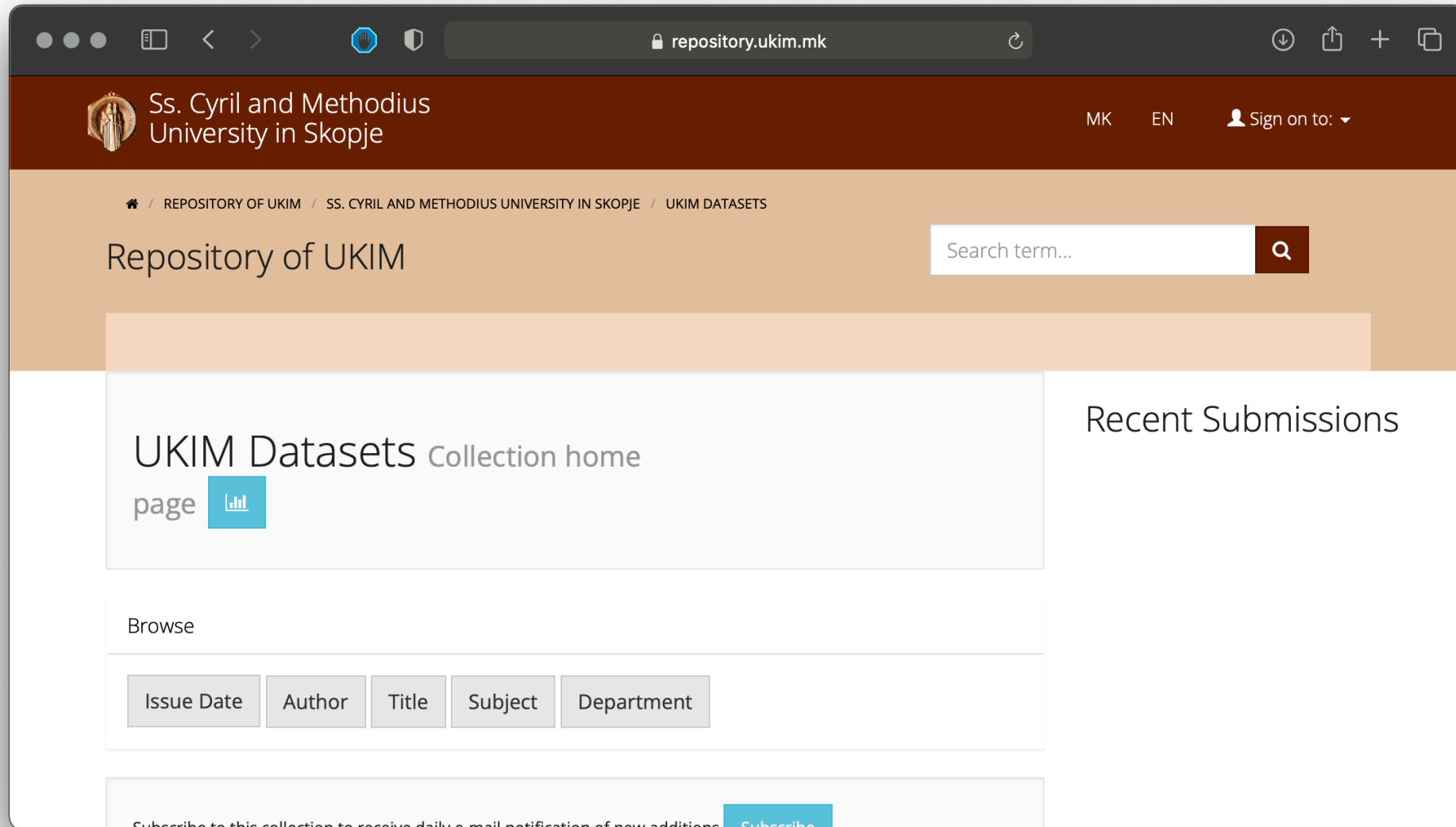
## Sequencing

- [dbSNP](#)
- [dbVar](#)
- [Database of Genomic Variants archive \(DGVa\)](#)
- [DNA DataBank of Japan \(DDBJ\)](#)
- [European Nucleotide Archive \(ENA\)](#)
- [GenBank](#)
- [MGnify](#)
- [miRBase](#)
- [NCBI Sequence Read Archive \(SRA\)](#)
- [NCBI Trace Archives](#)
- [UniProt](#)

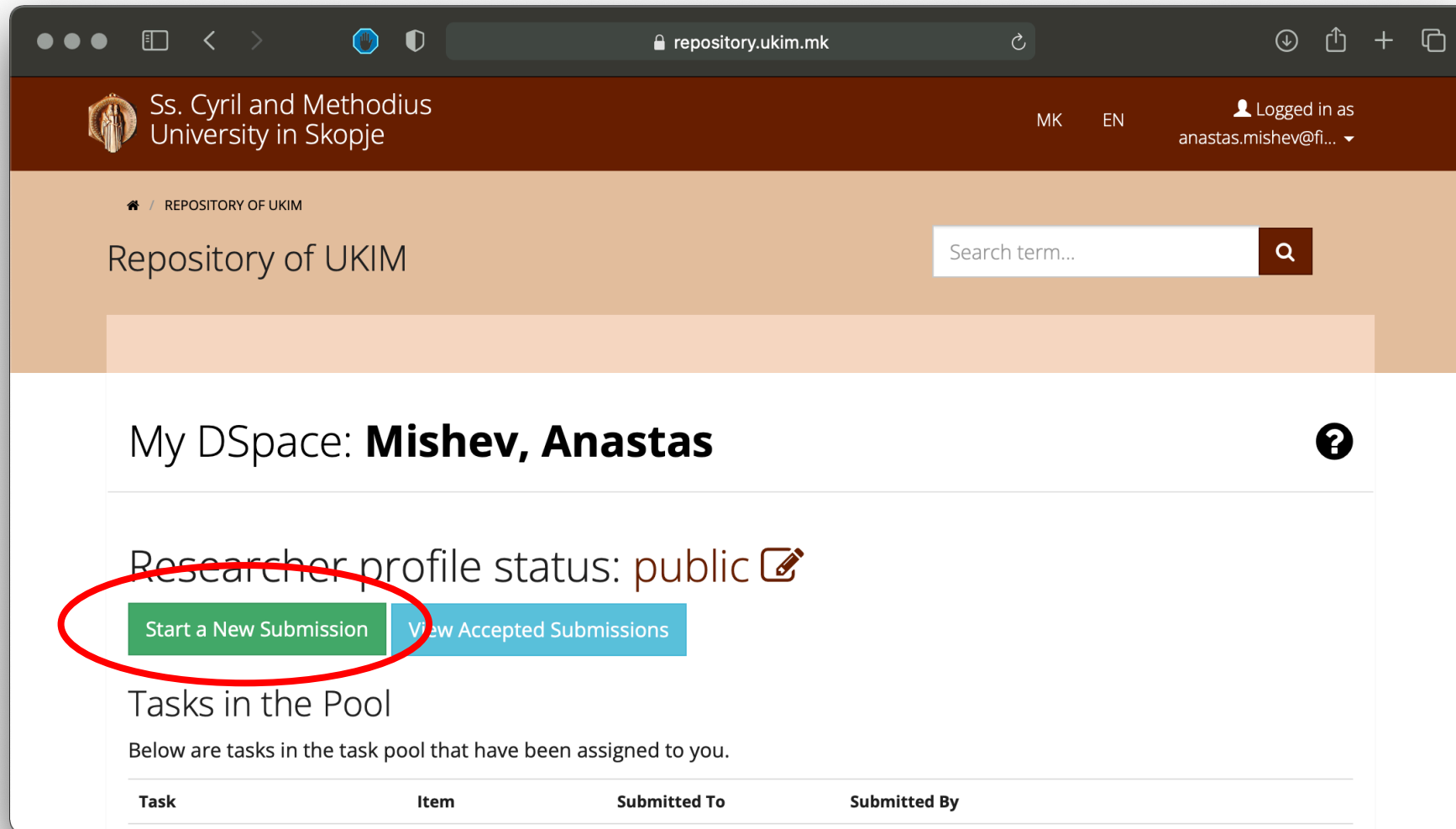


# Податочно складиште на репозиториумот на УКИМ





The screenshot shows a web browser window with the URL `repository.ukim.mk`. The page header features the logo of Ss. Cyril and Methodius University in Skopje and navigation links for MK, EN, and a sign-on option. The breadcrumb trail reads: `REPOSITORY OF UKIM / SS. CYRIL AND METHODIUS UNIVERSITY IN SKOPJE / UKIM DATASETS`. The main heading is "Repository of UKIM" with a search bar containing "Search term...". Below this, the page is divided into two columns. The left column displays "UKIM Datasets Collection home" with a "page" indicator and a bar chart icon. Underneath is a "Browse" section with filters for "Issue Date", "Author", "Title", "Subject", and "Department". At the bottom of this column, there is a "Subscribe" button. The right column is titled "Recent Submissions".



repository.ukim.mk

Ss. Cyril and Methodius University in Skopje

MK EN

Logged in as anastas.mishev@fi... ▾

REPOSITORY OF UKIM

Repository of UKIM

Search term... 🔍

My DSpace: **Mishev, Anastas** ?

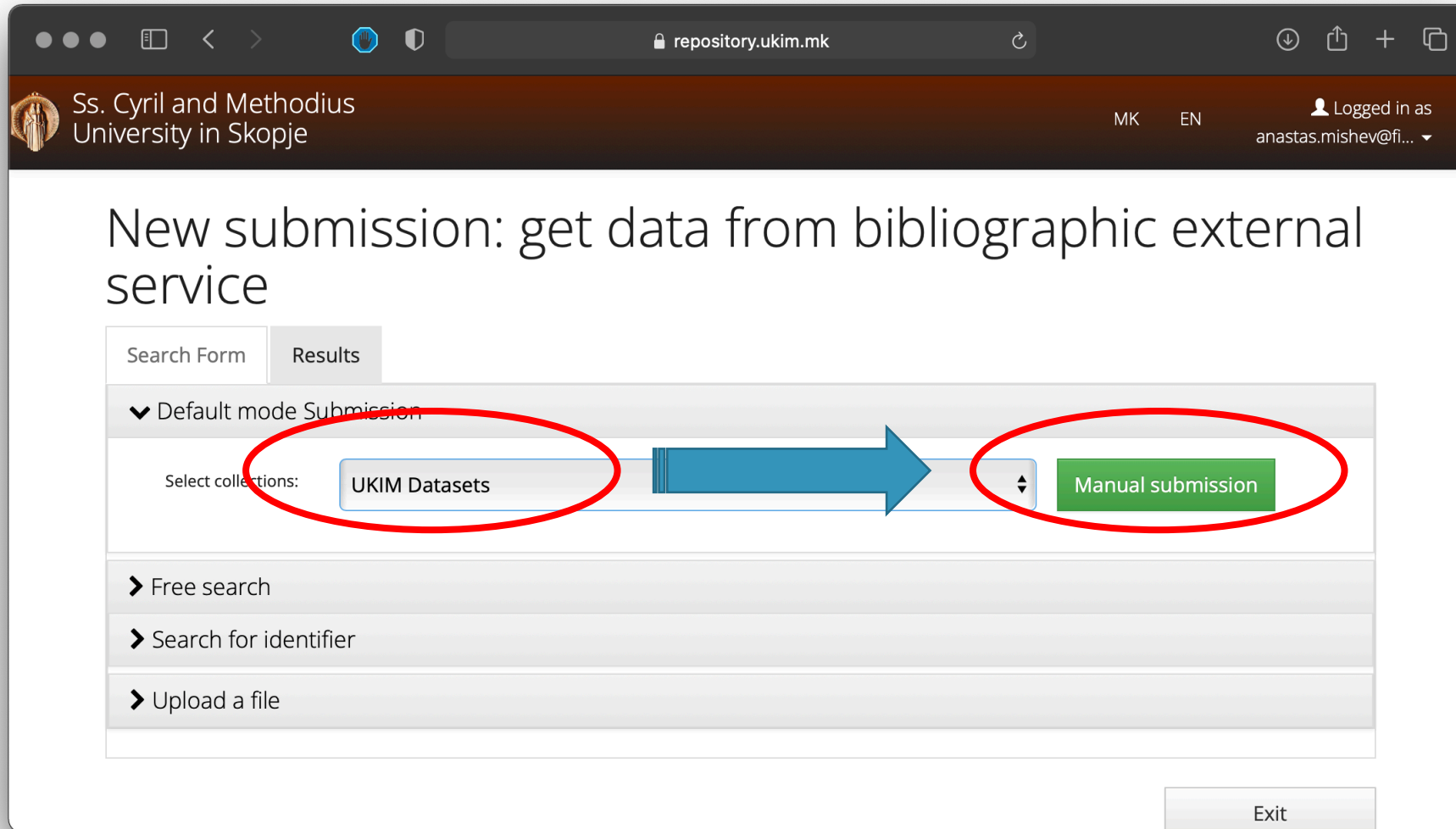
Researcher profile status: public ✎

**Start a New Submission** View Accepted Submissions

Tasks in the Pool

Below are tasks in the task pool that have been assigned to you.

Task	Item	Submitted To	Submitted By
------	------	--------------	--------------



repository.ukim.mk

Ss. Cyril and Methodius University in Skopje

MK EN

Logged in as anastas.mishev@fi...

## New submission: get data from bibliographic external service

Search Form Results

▼ Default mode Submission

Select collections: UKIM Datasets

Manual submission

► Free search

► Search for identifier

► Upload a file

Exit

Please fill in the requested information about this submission below. In most browsers, you can use the tab key to move the cursor to the next input box or button, to save you having to use the mouse each time.

Enter the names of the authors of this item in the form Surname, Firstname [i.e. Smith, Josh or Smith, J].

Authors

Filiposka, Sonja	✓	Q	Remove Entry
Mishev, Anastas	✓	Q	
Gilly, Katja		Q	

+ Add More

Enter the main title of the item.

Title \*

Simulation dataset for community-based allocation and migration strategies for fc

If the item has any alternative titles, please enter them here.

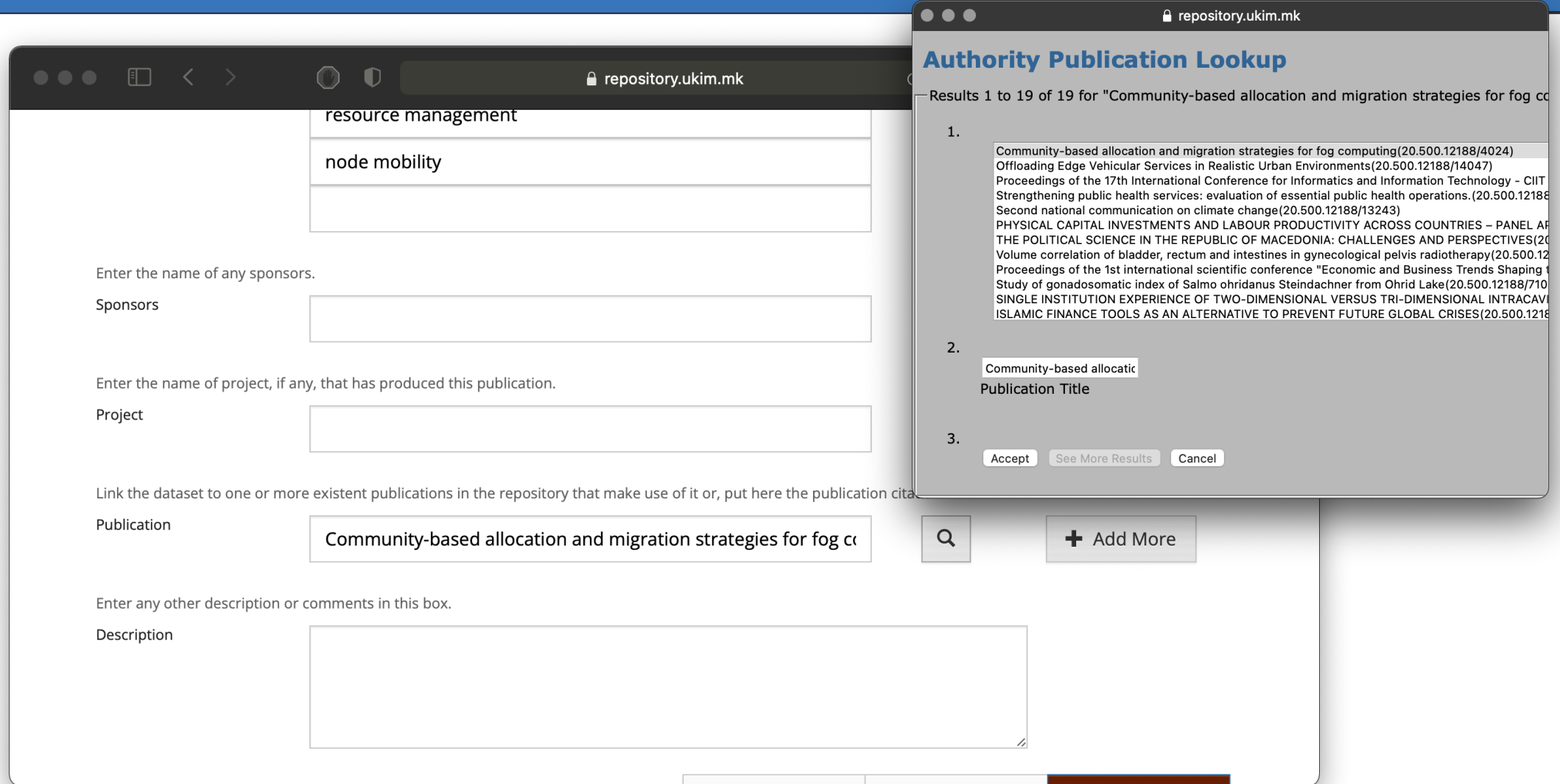
Other Titles

+ Add More

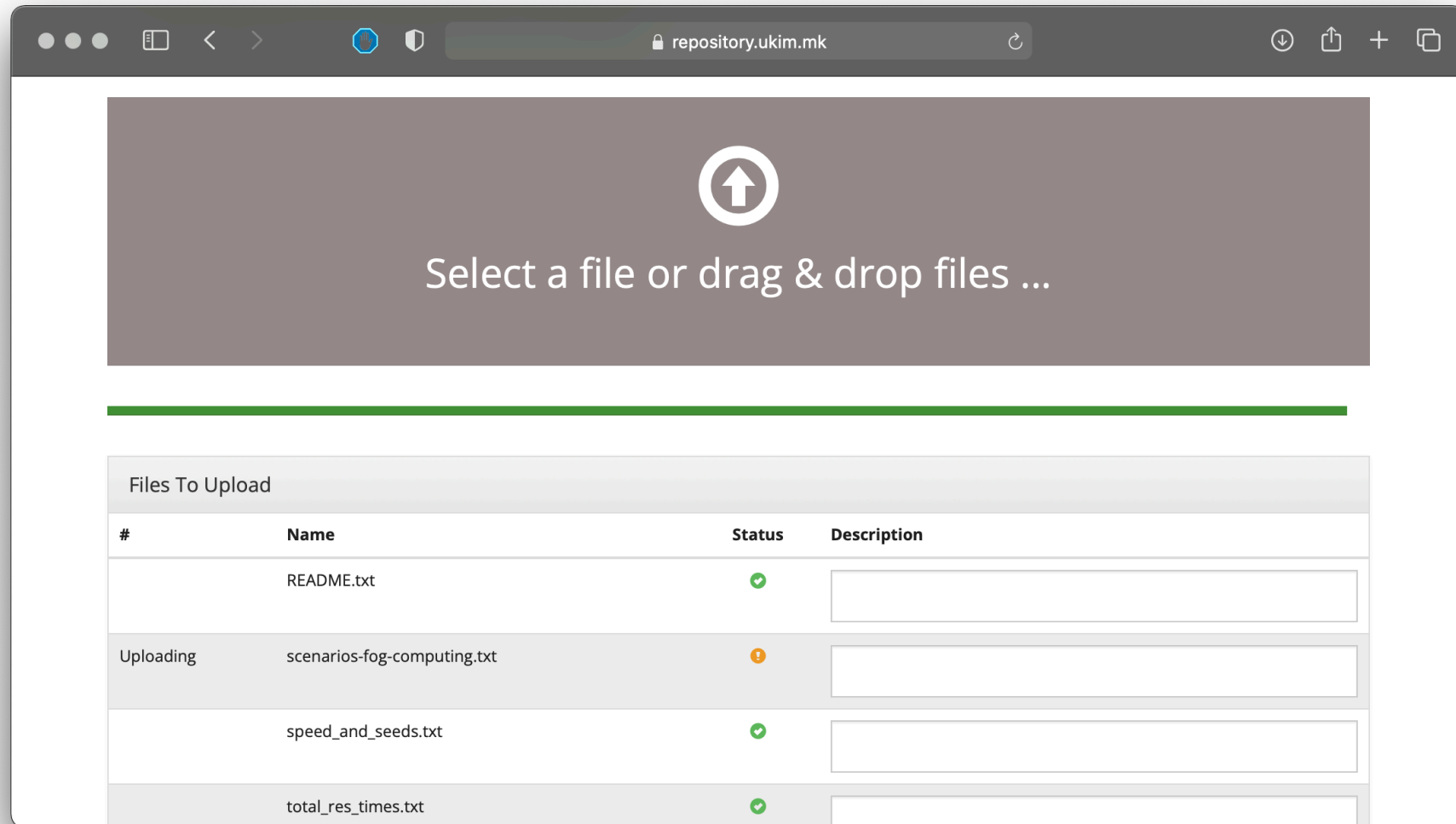
Please give the date of previous publication or public distribution. You can leave out the day and/or month if they aren't applicable.

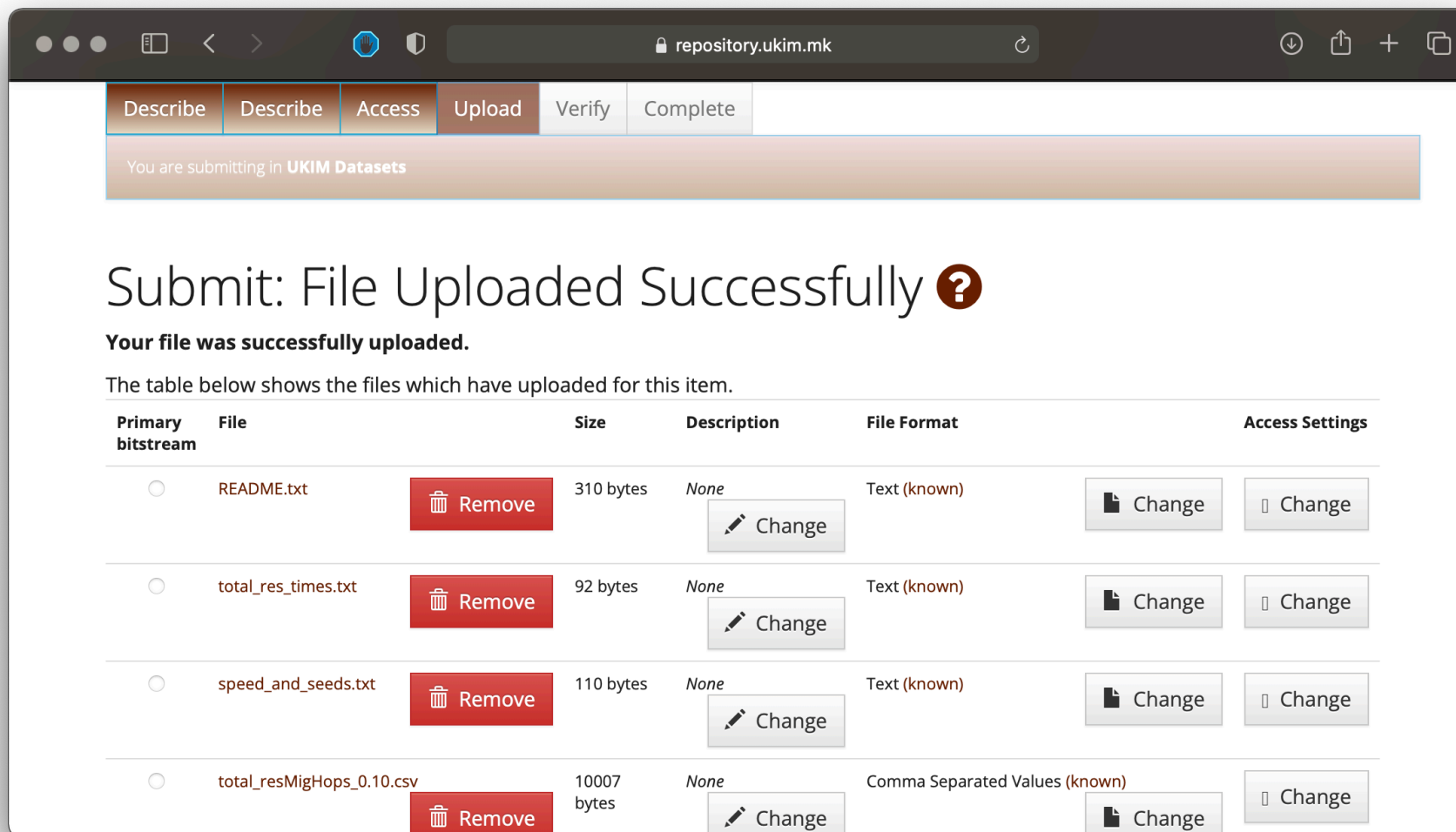
Date of Issue \*

Month:	април
Day:	
Year:	2018



The screenshot displays the repository.ukim.mk interface. The main form includes fields for 'resource management', 'node mobility', 'Sponsors', 'Project', 'Publication' (with the value 'Community-based allocation and migration strategies for fog c'), and 'Description'. An 'Authority Publication Lookup' modal window is open, showing search results for the query 'Community-based allocation and migration strategies for fog c'. The modal lists 19 results, with the first result being 'Community-based allocation and migration strategies for fog computing(20.500.12188/4024)'. The modal also features 'Accept', 'See More Results', and 'Cancel' buttons.








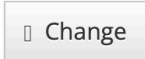



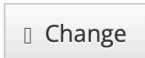
Describe Describe Access Upload Verify Complete

You are submitting in UKIM Datasets

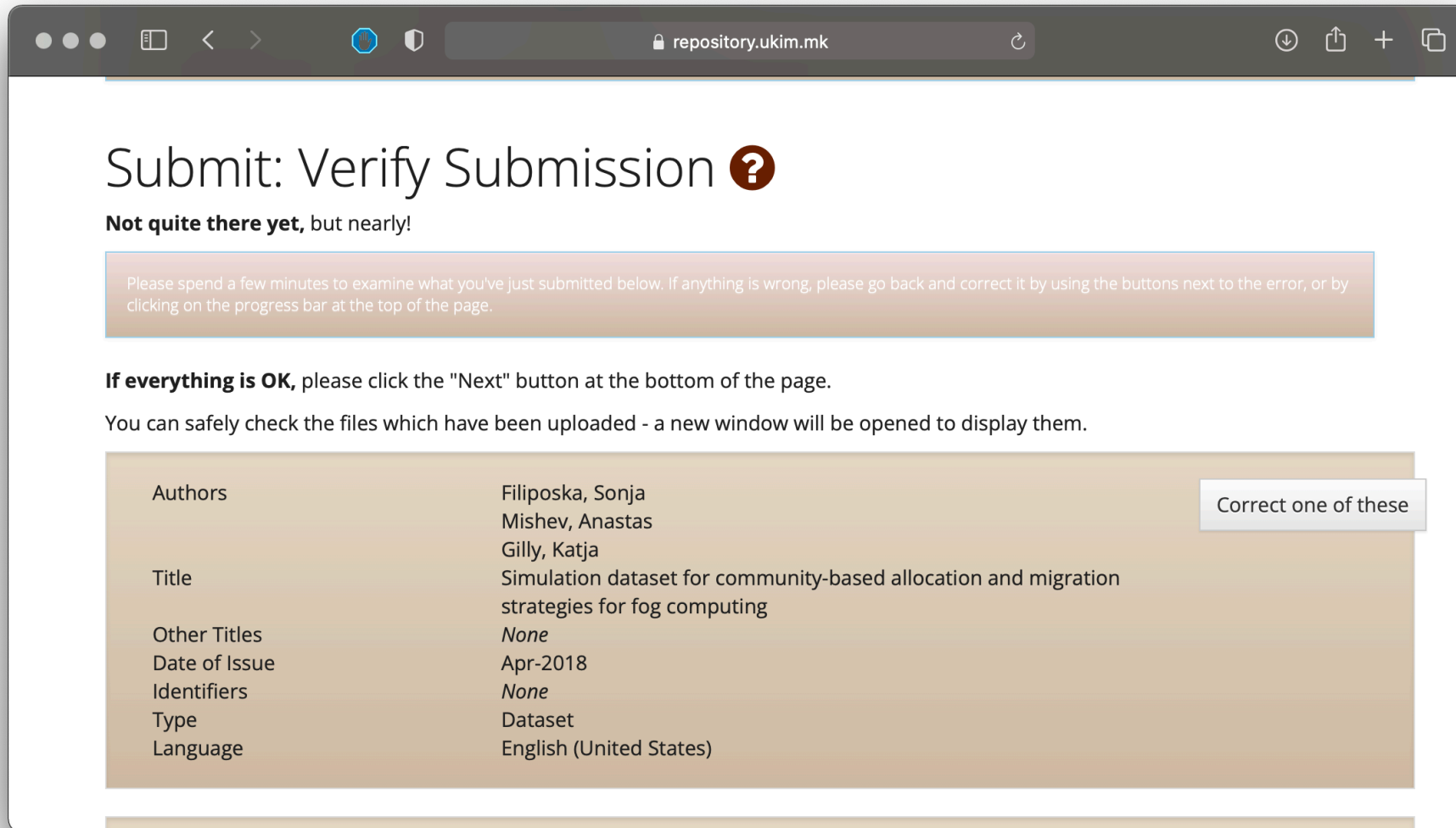
## Submit: File Uploaded Successfully ?

**Your file was successfully uploaded.**

The table below shows the files which have uploaded for this item.

Primary bitstream	File	Size	Description	File Format	Access Settings
<input type="radio"/>	README.txt	310 bytes	None	Text (known)	 Change  Change
<input type="radio"/>	total_res_times.txt	92 bytes	None	Text (known)	 Change  Change
<input type="radio"/>	speed_and_seeds.txt	110 bytes	None	Text (known)	 Change  Change
<input type="radio"/>	total_resMigHops_0.10.csv	10007 bytes	None	Comma Separated Values (known)	 Change  Change





Submit: Verify Submission ?

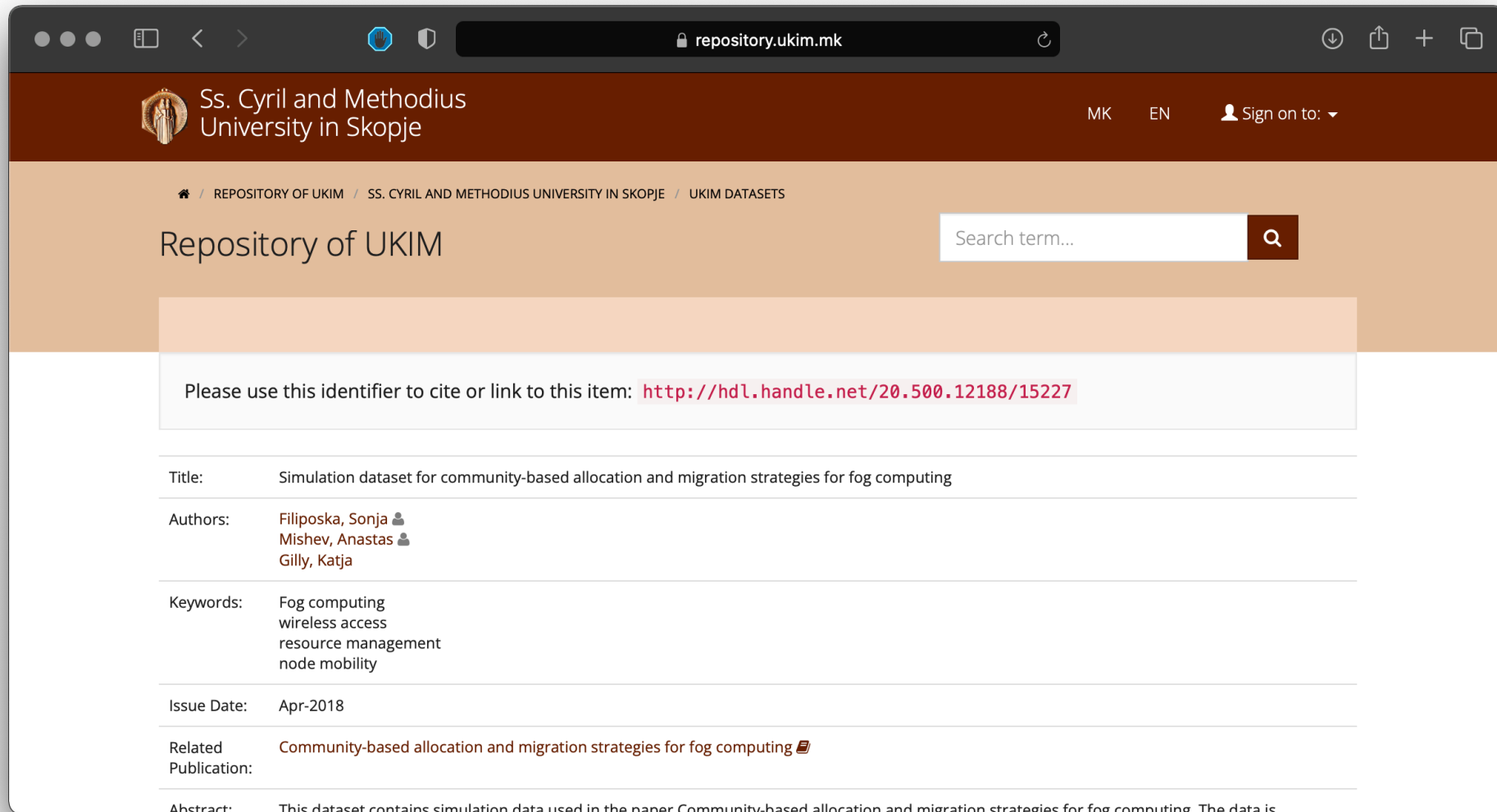
**Not quite there yet, but nearly!**

Please spend a few minutes to examine what you've just submitted below. If anything is wrong, please go back and correct it by using the buttons next to the error, or by clicking on the progress bar at the top of the page.

**If everything is OK,** please click the "Next" button at the bottom of the page.

You can safely check the files which have been uploaded - a new window will be opened to display them.

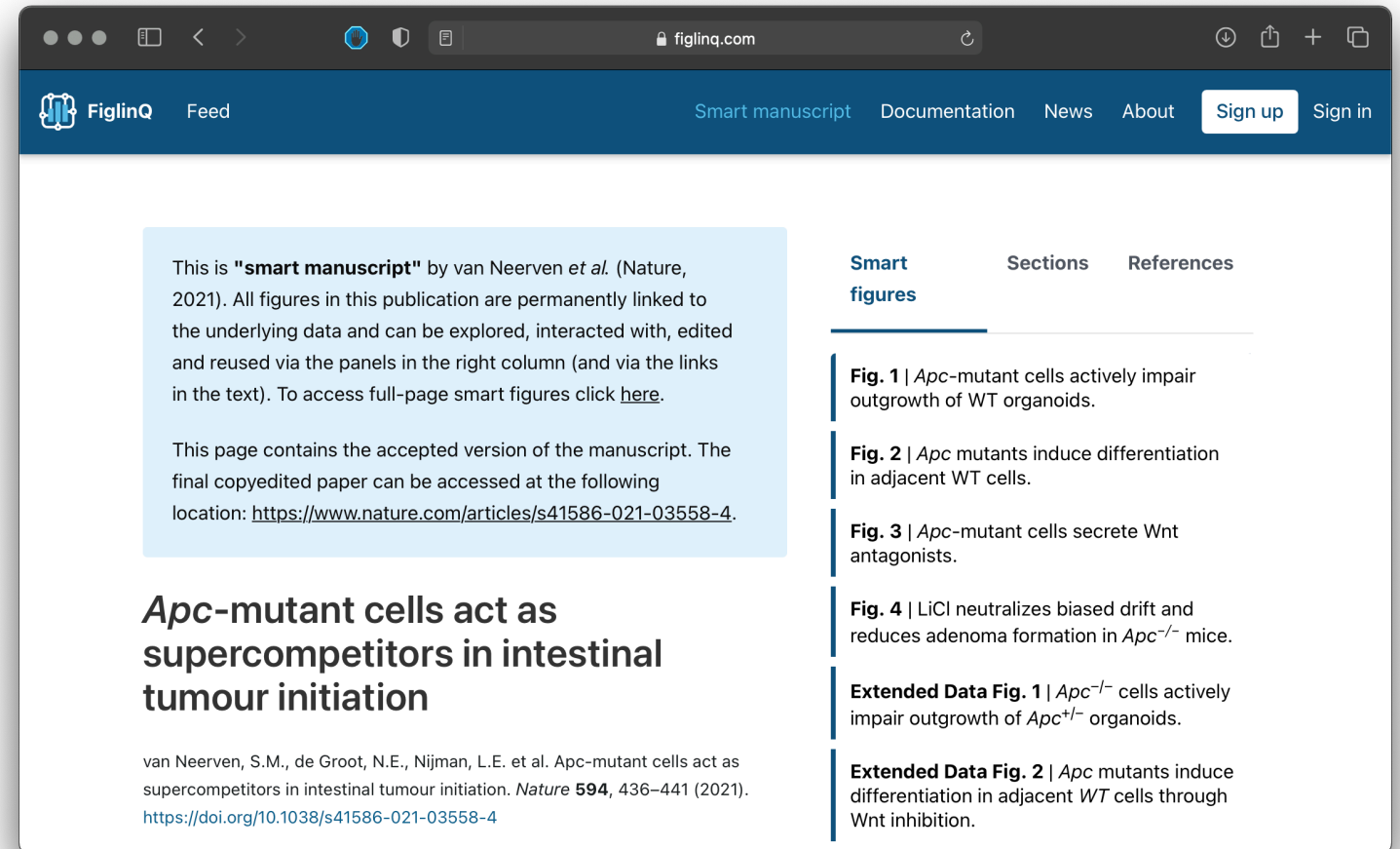
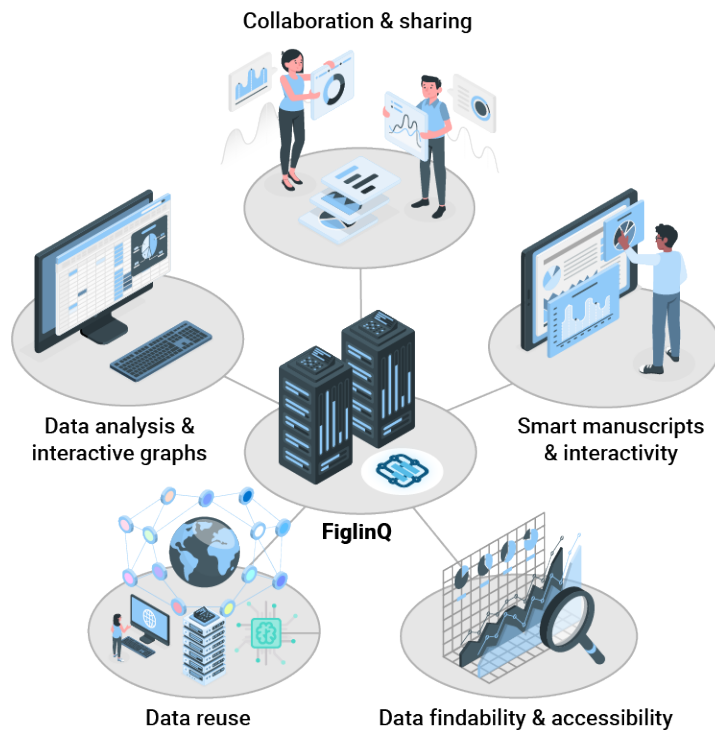
Authors	Filiposka, Sonja Mishev, Anastas Gilly, Katja	Correct one of these
Title	Simulation dataset for community-based allocation and migration strategies for fog computing	
Other Titles	None	
Date of Issue	Apr-2018	
Identifiers	None	
Type	Dataset	
Language	English (United States)	



The screenshot shows a web browser window with the URL `repository.ukim.mk`. The page header includes the logo of Ss. Cyril and Methodius University in Skopje and navigation links for MK, EN, and a sign-on option. The breadcrumb trail reads: `REPOSITORY OF UKIM / SS. CYRIL AND METHODIUS UNIVERSITY IN SKOPJE / UKIM DATASETS`. The main heading is "Repository of UKIM" with a search bar containing "Search term...". A highlighted box contains the text: "Please use this identifier to cite or link to this item: <http://hdl.handle.net/20.500.12188/15227>". Below this, the following metadata is displayed:

Title:	Simulation dataset for community-based allocation and migration strategies for fog computing
Authors:	Filiposka, Sonja Mishev, Anastas Gilly, Katja
Keywords:	Fog computing wireless access resource management node mobility
Issue Date:	Apr-2018
Related Publication:	<a href="#">Community-based allocation and migration strategies for fog computing</a>
Abstract:	This dataset contains simulation data used in the paper <a href="#">Community-based allocation and migration strategies for fog computing</a> . The data is

<https://figlinq.com/smart>



FiglinQ Feed Smart manuscript Documentation News About Sign up Sign in

This is "**smart manuscript**" by van Neerven *et al.* (Nature, 2021). All figures in this publication are permanently linked to the underlying data and can be explored, interacted with, edited and reused via the panels in the right column (and via the links in the text). To access full-page smart figures click [here](#).

This page contains the accepted version of the manuscript. The final copyedited paper can be accessed at the following location: <https://www.nature.com/articles/s41586-021-03558-4>.

## **Apc**-mutant cells act as supercompetitors in intestinal tumour initiation

van Neerven, S.M., de Groot, N.E., Nijman, L.E. et al. *Apc*-mutant cells act as supercompetitors in intestinal tumour initiation. *Nature* **594**, 436–441 (2021). <https://doi.org/10.1038/s41586-021-03558-4>

**Smart figures** Sections References

- Fig. 1** | *Apc*-mutant cells actively impair outgrowth of WT organoids.
- Fig. 2** | *Apc* mutants induce differentiation in adjacent WT cells.
- Fig. 3** | *Apc*-mutant cells secrete Wnt antagonists.
- Fig. 4** | LiCl neutralizes biased drift and reduces adenoma formation in *Apc*<sup>-/-</sup> mice.
- Extended Data Fig. 1** | *Apc*<sup>-/-</sup> cells actively impair outgrowth of *Apc*<sup>+/-</sup> organoids.
- Extended Data Fig. 2** | *Apc* mutants induce differentiation in adjacent WT cells through Wnt inhibition.

- ❑ <https://repository.ukim.mk/handle/20.500.12188/11971>
- ❑ <https://figlinq.com/smart>
- ❑ <https://isps.yale.edu/news/blog/2013/07/the-role-of-data-repositories-in-reproducible-research>
- ❑ <https://digitalguardian.com/blog/what-data-repository>
- ❑ <https://www.astera.com/type/blog/data-repository/>
- ❑ <https://authorservices.taylorandfrancis.com/data-sharing/share-your-data/repositories/>
- ❑ <https://stealthbits.com/blog/what-is-a-data-repository-and-what-is-it-used-for/>
- ❑ <https://journals.asm.org/list-data-repositories>
- ❑ <https://www.nature.com/sdata/policies/repositories>
- ❑ [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en)
- ❑ [https://like-itn-digitalization.readthedocs.io/en/latest/2\\_open\\_science/](https://like-itn-digitalization.readthedocs.io/en/latest/2_open_science/)

