

# eduTDM : Text-Mining Scientific Literature Without Barriers

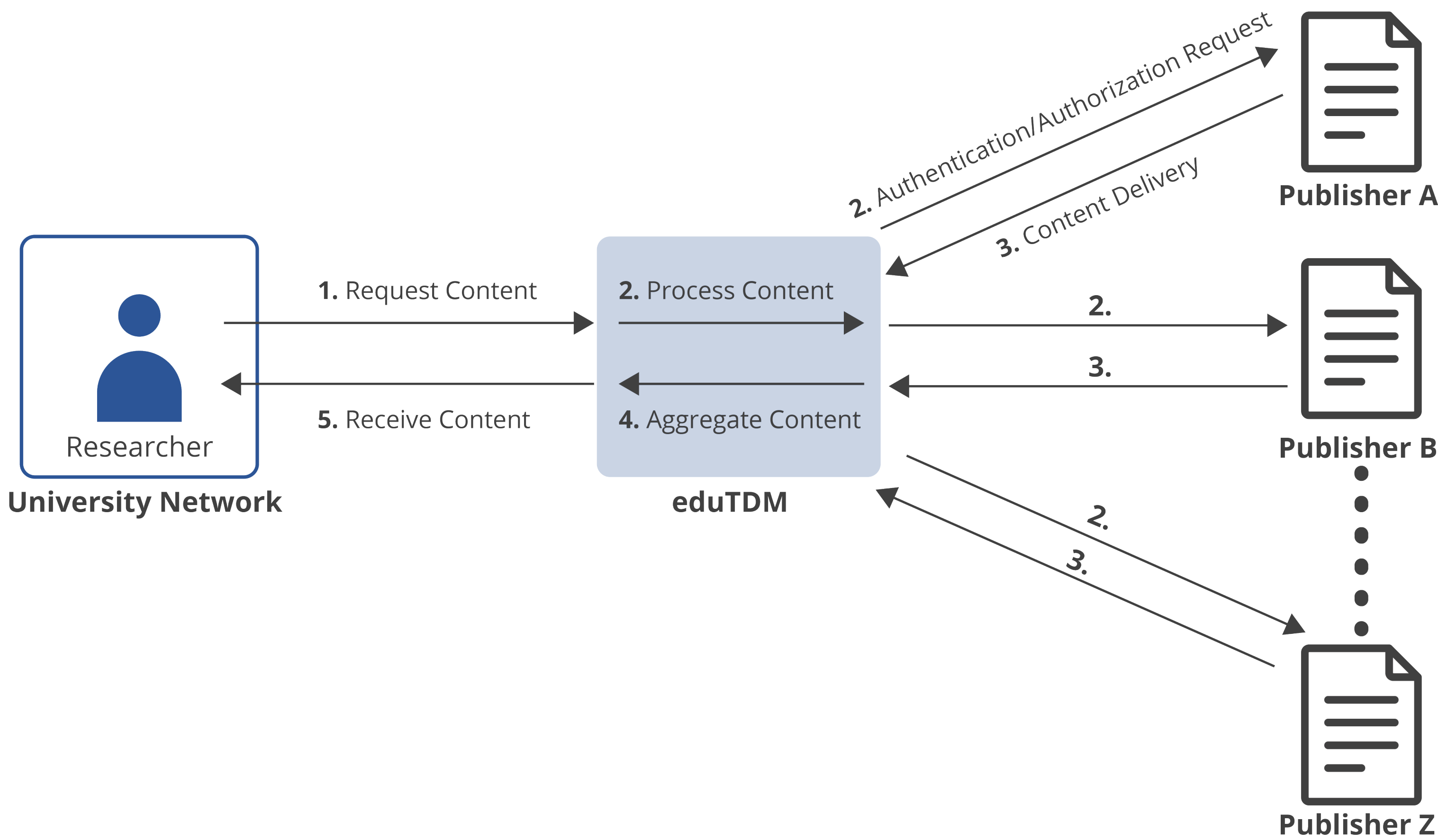


Petr Knoth, Bikash Gyawali and Nancy Pontika  
Knowledge Media Institute, The Open University, Milton Keynes, UK



## Problem Statement

Staff and students affiliated to a university can access and download all the research papers the institution subscribes to, provided that they are logged in to the institution’s network. However, it is not possible for text and data miners to machine access research literature their university subscribes to effectively and at scale.

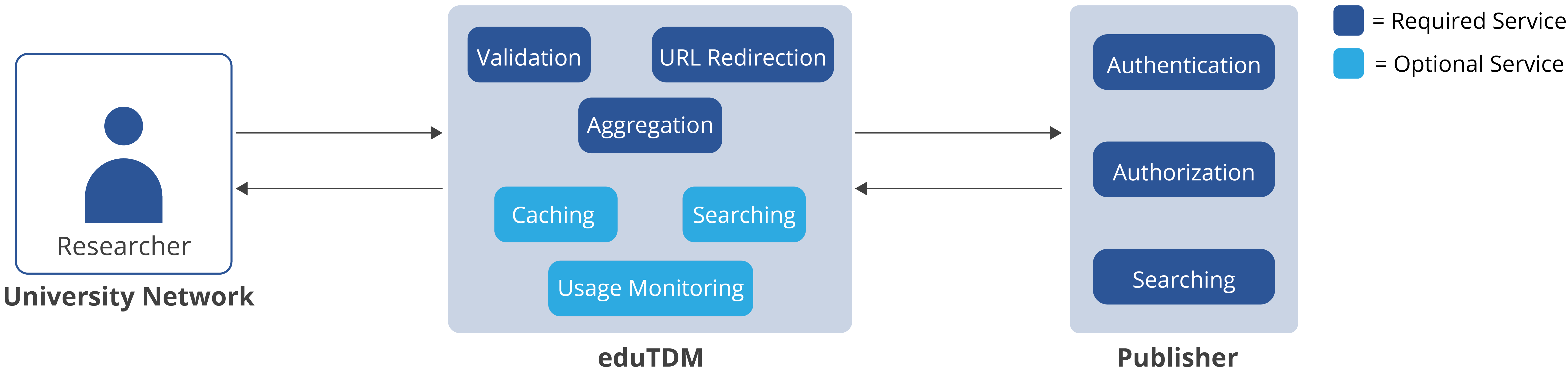


## eduTDM Vision

**eduTDM** aims to provide a central platform to access research articles from multiple repositories at once. This is useful for researchers who want to text/data mine content hosted on multiple different repositories without the burden of issuing and maintaining separate calls to each repository.

### eduTDM Group Mission

- Aims to :
  - Initiate and establish a collaboration and a communication route between a range of stakeholders
  - Define, discuss and disseminate the principles and views of the stakeholders
  - Understand the position of the stakeholders on eduTDM
  - Establish a development plan for putting eduTDM into practice.
- Stakeholders Meeting
  - October 2018
  - February 2019
- White Paper Proposal on implementing eduTDM – December 2018



## WHY use eduTDM?

Comparison Criteria	Publisher’s API	Crossref	eduTDM
Content Across Publishers	NO	Yes	Yes
API Output	Metadata + Full Text	Metadata + Link to Full Text.	Metadata + Full Text
Subscription based content	Separate authentication tokens (API Keys) per publisher.	Two-step process -- 1) get fulltext link from CrossRef 2) issue request per publisher.	Automatic identification of access rights across publishers.
Inter-operability	None	Interoperability by Centralization	True Interoperability
TDM Code Maintainability	Users need to update their code to changes in publishers’ end.	Users need to update their workflow to adapt to the changes.	eduTDM takes care of updates in publishers’ end.
Query support	Mainly DOI matching queries.	Mainly DOI matching queries.	Supports queries for matching different metadata values.

### eduTDM Working Group Members

Commercial publishers:



WILEY

Open Access publishers:



Industrial Partners:



Academic Partners:



Policy Makers:

