



Erasmus+



# Üst Veri Kalitesi

**Ayşen Binen**

İzmir Yüksek Teknoloji Enstitüsü



17 Kasım 2023  
izmir





# Yeni Nesil Akademik Arşivler

# Yeni Nesil Kurumsal Akademik Arşivler

## Kurumsal Akademik Arşiv,

bir kurum tarafından doğrudan ve dolaylı olarak yayınlanan; makale, bildiri, kitap, tez, rapor, araştırma verisi gibi tüm araştırma çıktıları için uluslararası standartlarda içerik arşivleme, yönetim, yayın performans değerlendirme, açık erişim altyapı, uzun süreli saklama ve içeriğe telif haklarına uygun erişim sağlayan sistemdir.



# Yeni Nesil Akademik Arşivler;



Kurumsal akademik arşivler, açık erişimin alt yapısı olmasının yanı sıra yayınların görünürlüklerini ve erişilebilirliğini artırır. Bu yolla araştırmanın etki faktörünün daha da yükseltilmesine katkı sağlar.



En yaygın kurumsal akademik arşiv yazılımları DSpace, EPrints, Fedora ve Islandora'dır ve bu arşivler Dublin Core üst verilerini destekler.



Yeni nesil kurumsal akademik arşivler bir kurumun tüm araştırma ekosistemini dünya ile entegre edebilmesi, doğru yönetebilmesi ve isabetli kararlar alabilmesine yardımcı olmayı amaçlar.



Kurum içi ve kurum dışı bilgi kaynaklarından gelen araştırma bilgilerini toplar ve stratejik kararlara yön veren raporların güvenilir, kapsamlı ve gerçek zamanlı olarak erişilebilir olmasını sağlar.



Çok çeşitli raporlarla, kurumun performans değerlendirmeleri yapmasına, araştırmacıların profillerini yönetmesine, araştırma ağları kurmasına ve beraber çalışabilecekleri uzmanları bulmasına ve çok daha fazlasına imkan verir.

# Yeni Nesil Akademik Arşivler

✓ Görünürlük

✓ Araştırma Desteği

✓ Performans  
İzleme

✓ Karar Desteği

✓ Dünya ile  
Entegrasyon

✓ Güvenilir  
Raporlama



CRIS systems



NEXT GENERATION  
REPOSITORIES

**Confederation of Open Access Repositories (COAR, 2020), farklı türdeki (yayın, kurumsal, veri vb.) açık arşivler tarafından benimsenebilecek ve kullanılabilir en iyi uygulamaları değerlendirmek ve küresel açık arşiv kriterlerini belirlemek için bir çalışma yapmıştır.**

**Raporun Erişim Linki:**

**<https://zenodo.org/records/8077381>**

The screenshot shows the Zenodo record page for the report. The page is titled "Next Generation Repositories: Behaviours and Technical Recommendations of the COAR Next Generation Repositories Working Group". It features a blue header with the Zenodo logo and navigation links. The main content area includes a summary of the report, a list of authors, and a list of keywords. The report is available in PDF format and has 1,489 views and 1,103 downloads. The page also includes a "Share" section with social media icons and a "Cite as" section with the report's DOI and URL. The footer contains links for "About", "Blog", "Help", "Developers", and "Contribute", along with logos for CERN, OpenAIRE, and the European Union.

# COAR'a göre Yeni Nesil Akademik Arşivlerin Özellikleri;

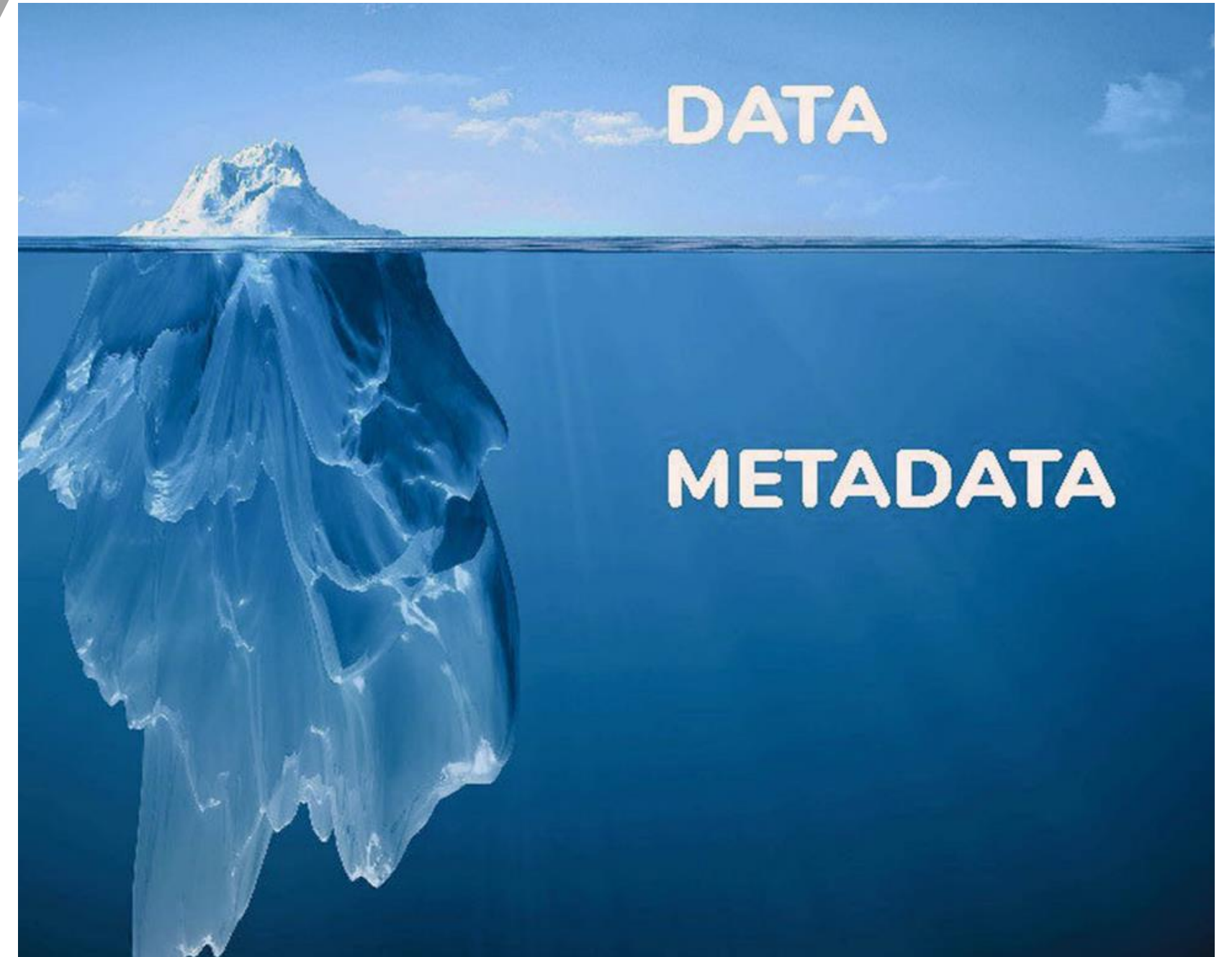
- Farklı türlerde kaynaklar içermeli.
- Makaleler, veriler ve yazılım gibi ilgili kaynaklar arasında bağlantı kurulmasını desteklemeli.
- İnsan ve makine tarafından okunabilir formatlarda **kaliteli üst veri içermeli.**
- Hassas araştırma verisi toplanan durumlarda, verilere farklı düzeylerdeki güvenlik gereksinimlerine dayalı olarak katmanlı erişim sağlamalı.
- Diğer sistemlerle birlikte çalışabilmeli.
- Araştırma bilgi sistemi özelliği taşımalı.
- FAIR Prensiplerine uyumlu olmalı.
- Kurumun araştırma ekosisteminin bir parçası olmalı.



# Üst Veri Ne Değildir?

"Metadata is data that describes data, but it isn't the data itself."

Merle Beljaev







**Üst veri; belge, görüntü, veri kümesi gibi kavramları anlamlandırır.**



**Kaliteli üst veri olmadan bilgiye erişim ve raporlama ancak tesadüf eseri olur**



**Kaliteli üst veri, aranabilirliği ve yeniden kullanımını doğrudan etkiler.**



Uluslararası standartlardaki araştırma bilgi sistemleri ve akademik arşivlerde;

- Sistemlerin dünyadaki diğer sistemlerle sorunsuz bir şekilde birlikte çalışabilmesi için

**STANDARTLARA UYGUN ve ZENGİNLEŞTİRİLMİŞ ÜST VERİ** gereklidir.

# Üst veriniz kaliteli ve standartlara uygun değilse, bilgiye erişmek tamamen tesadüflere kalmıştır.

## FAIR Prensipleri

### Veriniz Olmalı:

- **F**indable (Bulunabilir)
- **A**ccessible (Erişilebilir)
- **I**nteroperable (Birlikte Çalışabilir)
- **R**eusable (Yeniden Kullanılabilir)

#### Bulunabilir

- Tanımlayıcı üstveri
- Kalıcı Tanımlayıcılar

#### Erişilebilir

- Neyi paylaşacağını saptamak
- Katılımcı izni ve risk yönetimi
- Erişim durumu

#### Birlikte Çalışabilir

- XML standartları
- Veri Dokümantasyonu Girişimi
- CDISC (Clinical Data Interchange Standards Consortium)

#### Yeniden Kullanılabilir

- Haklar ve lisans modelleri
- İzin verilen ve izin verilmeyen kullanım

<https://www.openaire.eu/>

Türkiye Yardım Masası:

Gültekin Gürdal

İzmir Yüksek Teknoloji Enstitüsü

[openaire@iyte.edu.tr](mailto:openaire@iyte.edu.tr)

<http://libguides.iyte.edu.tr/OpenAIRE>

İndirmek için:

<http://libguides.iyte.edu.tr/OpenAIRE>



<http://datafairport.org/>



**Standartlara uygun zenginleştirilmiş bir üst veri,  
birlikte çalışabilirliğin ve üst veri sağlayıcısı  
olmanın temel anahtarıdır.**

# Üst Veri Kategorileri



## Tanımlayıcı Üst Veriler

İçeriği tanımlar.

- Title,
- Creator,
- Contributor,
- Subject,
- Document type,
- Description,
- Language,
- Source, ...



## Yapısal Üst Veriler

Dosyanın diğer dosyalarla nasıl çalıştığı bilgisini içerir.

- Relation
- Format



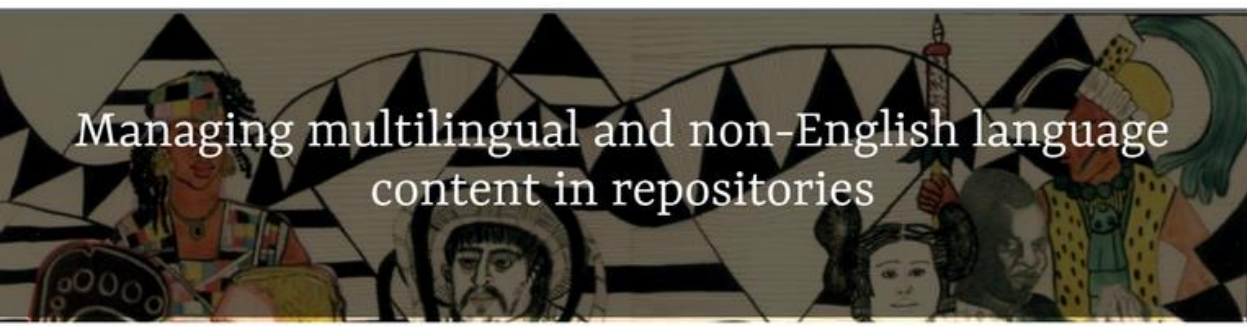
## Yönetimsel Üst Veriler

Dosyanın nasıl kullanılabileceği veya kimler tarafından görülebileceği hakkında bilgiyi içerir.

- Data created
- File type
- Identifier
- Rights

(Hanlon, 2001)





## Managing multilingual and non-English language content in repositories

Multilingualism is critical for a healthy, inclusive, and diverse research communications landscape. After decades of policies that have directed researchers to publish in English, we are starting to see a reversal of this trend. The UNESCO Recommendation on Open Science, for example, calls on member states to encourage "multilingualism in the practice of science, in scientific publications and in academic communications". In China, Europe, and other jurisdictions, policy makers are introducing new measures that encourage researchers to publish in local languages.

In August 2022, COAR launched the COAR Task Force on Supporting Multilingualism and non-English Content in Repositories to develop and promote good practices for repositories in managing multilingual and non-English content. Based on 17 use cases contributed from different stakeholders communities (repository managers and users, authors and translators, aggregators and discovery systems), the Task Force identified three relevant areas for their work:

1. Enhancing discoverability of non-English content
2. Curating multilingual content in a repository
3. Supporting translations

In June 2023, the Task Force published an initial set of draft recommendations for community review. The consultation resulted in a wide range of input, which was reviewed by the Task Force and incorporated into the recommendations document. This document presents the updated recommendations based on this community input. The recommendations identify good practices for repository managers and repository software developers, and focus on the topics of metadata, multilingual keywords, user interfaces, formats, and licences that will improve the visibility, discovery and reuse of repository content in a variety of languages.

Visit our webpage on managing multilingualism and non-English content in repositories to see the detailed recommendations

### Summary of Recommendations

#### Creating and Curating Metadata

1. Declare the language of the resource at the item level
2. Declare the language of the metadata (e.g. xml:lang attribute)
3. Use standard (two-letter or three-letter) language codes (ISO 639)
4. Enable UTF-8 support in your repository and use the original alphabet / the writing system whenever possible. If it is necessary to transliterate metadata, use recognized standards (e.g. ISO)
5. If the repository software supports multiple interface languages, set up the user interface in the native language(s) of the target group, along with that in English
6. Write personal name/s using the writing system used in the deposited document and provide a persistent identifier enabling unambiguous identification, such as ORCID
7. Include keywords in many languages, use multilingual vocabularies and thesauri if possible
8. Recommendations for repository managers on translated content

#### Repository Software / Platform Developers

1. Ensure that language codes can consistently be used across the repository collections
2. Expose the language of metadata via metadata exchange protocol, e.g. OAI-PMH, GraphQL API, etc.
3. Improve support for ISO language codes, e.g. three-letter codes needed for some languages.
4. Ensure that persistent identifiers are exposed via OAI-PMH. PIDs in Dublin Core™ Working Group has developed recommendations to make it possible to expose persistent identifiers including ORCID, via OAI-PMH.
5. Provide support for multilingual keywords to increase the discoverability of multilingual repository content. For example, enable a real-time integration of Wikidata – e.g. when a user starts typing in the appropriate metadata field, relevant Wikidata terms appear in a drop-down list for the user to select
6. Enable automatic assignment of controlled terms based on the existing metadata

# COAR Görev Gücü tarafından 30 Ekim'de Kurumsal Arşivlerde Çok Dilli ve İngilizce Olmayan İçeriklerin Yönetimi için İyi Uygulama Önerilerini listelediği bir rapor yayınlandı ve bu raporun Türkçe'ye çevirisi ekibimiz tarafından gerçekleştirildi.

Rapor için; <https://www.coar-repositories.org/news-updates/managing-multilingual-and-non-english-language-content-in-repositories/>

Çeviri için; <https://hdl.handle.net/11147/13916>

## Good Practice Advice for Managing Multilingual and non-English Language Content in Repositories



Image credit | Tommaso D'Incalci | Ikon Images [CC BY-NC](#)

### Contributing Task Force Members

- Iryna Kuchma, EIFL (chair), Ukraine
- Jagadish Aryal, Social Science Baha, Nepal
- Andreas Czerniak, Bielefeld University – Library, Germany
- Christophe Dony, ULiège Library, Belgium
- Joe Cera, Berkeley Law Library, University of California, USA



# Üst Veri Standartları

Açık Arşiv Girişimleri OAI-PMH protokolünün (Open Archives Initiative Protocol for Metadata Harvesting - Açık Arşivler Girişimi Üst Veri Harmanlama Protokolü) desteklediği üst veri standartları;

- |            |         |
|------------|---------|
| ✓ oai_dc   | ✓ dim   |
| ✓ xoad     | ✓ etdms |
| ✓ marc     | ✓ qdc   |
| ✓ rdf      | ✓ ore   |
| ✓ uketd_dc | ✓ mods  |
| ✓ mets     | ✓ didl  |



## **Üst Veri Standartları: Dublin Core**

**Dublin Core (DC), konu ve kaynak tipi ayrımı gözetmemesinden ötürü ön plana çıkmıştır.**

**DC üst veri element seti; Mart 1995'te OCLC ve İleri Hesaplama Uygulamaları Ulusal Merkezi (NCSA) tarafından sponsorluğu yapılan bir çalıştay sonucunda ortaya çıkmıştır.  
(Hwei and Smeltzer, 2000)**



# GCRIS'te Yer Alan Üst Verilerin Temeli



## Dublin Core Standardı

Yayınlardaki üst verilerin temeli Dublin Core standardına dayanmaktadır.



## euroCRIS CERIF Standardı

Araştırmacı ve bölüm profilleri, projeler ve ekipmanlar gibi bileşenlerdeki üst verilerin temeli euroCRIS CERIF standardına dayanmaktadır.

# Niçin Üst Verimiz Kaliteli Olmalı?

- Erişim ve görünürlük
- Atıf
- Kullanımı artırma
- Uygulamaya hızlı geçiş
- Araştırma ve keşif



# Kaliteli üst veri oluştururken;

Tüm üst veriler toplanmalı

Çok dilli yayınlarda üst veri alanları dile göre tekrarlanmalı.

Üst verinin doğru ve eksiksiz olduğundan emin olunmalı.

Üst veri mutlaka bir editör tarafından gözden geçirilmeli.

Üst veride jargon kullanılmamalı.

Üst veride kullanılan teknik terimler ve kısaltmalar mutlaka tanımlanmalı.

Üst veride olmayan ve bilinmeyen veriler mutlaka açıklanmalı.

Kaynak ile ilgili sınırlamalar varsa açıkça belirtilmedi.

# Kaliteli Üst Verinin Çıktıları;

- Analiz edilebilir
- Ölçülebilir
- Öngörü sağlayabilir
- Harmanlanabilir
- Başka sistemlerle konuşabilir
- Planlama yapmaya katkı sunabilir
- Raporlanabilir



# DSpace ve GCRIS'te Üst Veri Oluřturma Süreci

Kütüphaneciler tarafından oluşturulan veriler ile,

Farklı platformlardan derlenen verilerden oluşur:

- WoS,
- Scopus,
- TR-Dizin,
- YÖK Tez Merkezi,
- PubMed
- ORCID
- Open Journal Systems (OJS) ve
- Kütüphane Otomasyon Sistemleri



# DSpace ve GCRIS'te Üst Veri Sayıları



DSpace ve GCRIS'e zenginleştirme yoluyla veri giriş seçeneği kullanan arşivlerdeki üst veri alanı sayısı:

**96**

# Özet Olarak Zenginleştirilmiş Veri

Kaynakla ilgili en geniş bilgileri sunmayı sağlar.

Zenginleştirilmiş ve standartlara uygun veri, güvenilir rapor üretmeyi sağlar.

Temsil edilen kaynakları ön plana çıkarır.

Araştırmaların tekrarını önler.

Kişi ve kurumları görünür yapar.

Öngörü ve doğru raporlama için olmazsa olmazdır

Araştırmacılara zaman kazandırır.

# Zenginleştirilen ve standartlara uygun hale getirilen her bir üst veri sizi keşfedilebilir kılar...

Clarivate

Web of Science™

Search

English Products

Sign In Register

Search > Results for Izmir Institute of... > Results for Izmir Institute of Technology (Affiliation) and Open Access

4,268 results from Web of Science Core Collection for:

Izmir Institute of Technology (Affiliation)

Analyze Results Citation Report Create Alert

Refined By: Open Access X Clear all

Copy query link

Publications You may also like...

Refine results

0/4,268 Add To Marked List Export Sort by: Relevance 1 of 86

Search within results...

Filter by Marked List

Quick Filters

- Review Article 89
- Early Access 19
- Open Access 4,268
- Enriched Cited References 171

1 Efficient Search in a Panoramic Image Database for Long-term Visual Localization

Orhan, S and Bastanlar, Y

18th IEEE/CVF International Conference on Computer Vision (ICCV) 2021 | 2021 IEEE/CVF INTERNATIONAL CONFERENCE ON COMPUTER VISION WORKSHOPS (ICCVW 2021), pp.1727-1734

In this work, we focus on a localization technique that is based on image retrieval. In this technique, database images are kept with GPS coordinates and the geographic location of the retrieved database image serves as an approximate position of the query image. In our scenari... Show more

Free Published Article From Repository Full Text at Publisher

Click to view final published versions of

gcris.iyte.edu.tr/bitstream/11147/12340/1/Efficient\_Search\_in\_a\_Panoramic.pdf

Efficient Search in a Panoramic Image Database for Long-term ... 1 / 8 100%

2021 IEEE/CVF International Conference on Computer Vision Workshops

## Efficient Search in a Panoramic Image Database for Long-term Visual Localization

Semih Orhan  
Department of Computer Engineering  
Izmir Institute of Technology  
semihorhan@iyte.edu.tr

Yalın Bastanlar  
Department of Computer Engineering  
Izmir Institute of Technology  
yalinbastanlar@iyte.edu.tr

### Abstract

In this work, we focus on a localization technique that is based on image retrieval. In this technique, database images are kept with GPS coordinates and the geographic location of the retrieved database image serves as an approximate position of the query image. In our scenario, database consists of panoramic images (e.g. Google Street View) and query images are collected with a standard field-of-view camera in a different time. While searching the match of a perspective query image in a panoramic image database, unlike previous studies, we do not generate a number of perspective images from the panoramic image. Instead, taking advantage of CNNs, we slide a search window in the last convolutional layer belonging to the panoramic image and compute the similarity with the descriptor extracted from the query image. In this way, more locations are visited in less amount of time. We conducted experiments with state-of-the-art descriptors and results reveal that the proposed sliding window approach reaches higher accuracy

long-term localization (LTL) is a challenging task due to various challenges such as illumination changes, seasonal changes, and object changes in the scene. Any localization problem should address these changes.

In our setting, the environment changes over time. Images acquired at different times than a metric localization (also called as location-based localization [10, 2, 7, 14]) that could be used as agents. In our work, we focus on perspective images in particular. The main contribution of this paper is to exploit their wide field-of-view. With this advantage, recall is improved. This is achieved in some scenarios where other methods fail due to their non-overlapping fields-of-view.

Our main contribution is to propose a localization technique for perspective (standard FOV) images in a panoramic image database.



# TEŞEKKÜRLER



**Ayşen BİNEN**

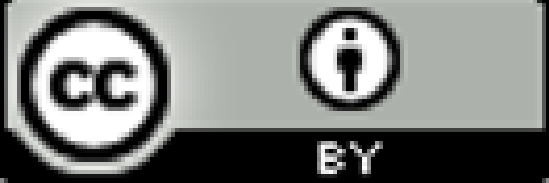
**Kurumsal Akademik Arşiv Yöneticisi**



**+90-232-750-6343**



**aysenbinen@iyte.edu.tr**



**Aksi belirtilmediği sürece, bu sunumun içeriği Creative Commons Atıf 4.0 Uluslararası Lisansı ile lisanslanmıştır.**

