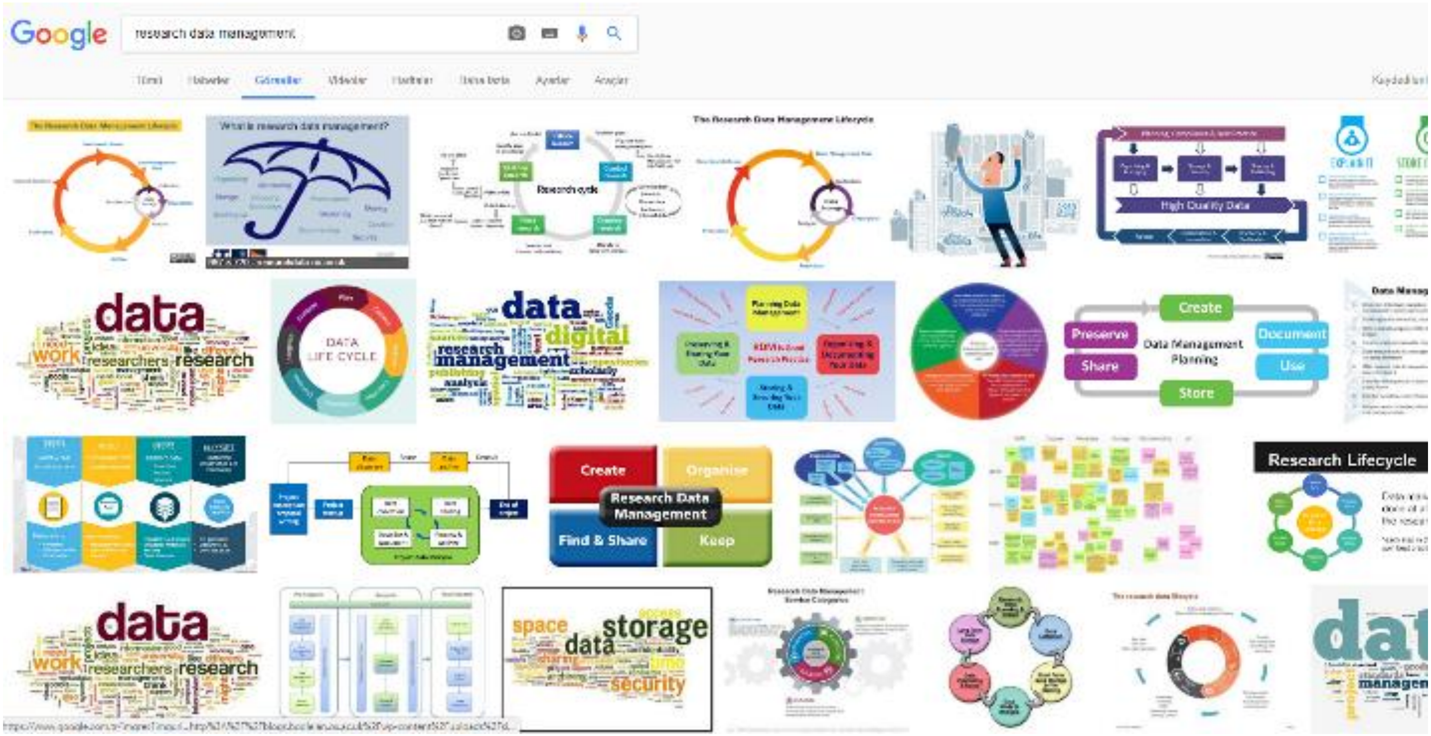
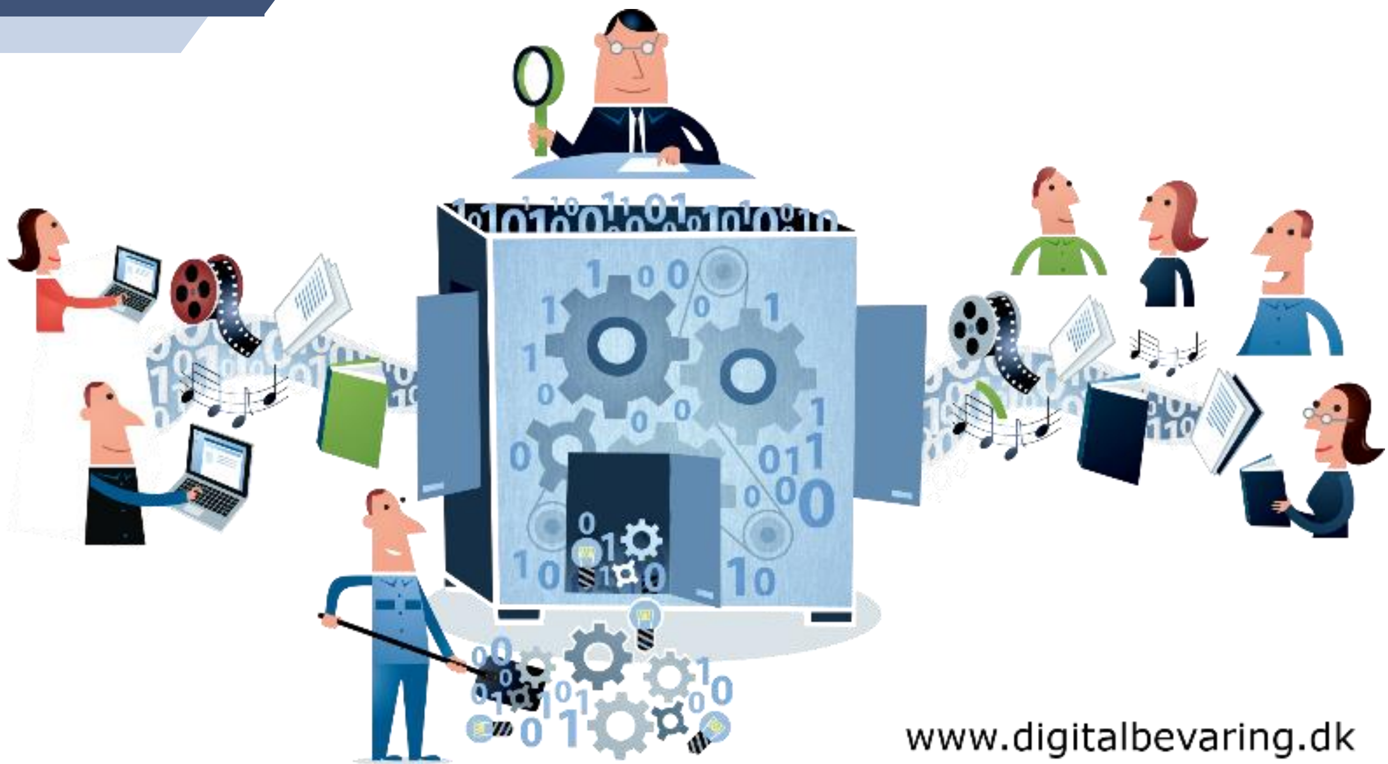
Veri Yönetiminde Başarıya Ulaşmak: Teknik Konular

Dr. Zehra Taşkın

Plan

- Araştırma verisi nedir?
- Her yönüyle «araştırma verileri»
  - Güvenilir araştırma verisi
  - Paylaşılabilir araştırma verisi
  - Saklanabilir araştırma verisi

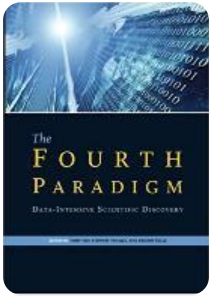

Google
research data management
Research cycle
High Quality Data
data
work
researchers
research
DATA LIFE CYCLE
digital
management
Create
Preserve
Share
Data Management Planning
Document
Use
Store
Organise
Research Data Management
Find & Share
Keep
Research Lifecycle
space
storage
security



www.digitalbevaring.dk

“Tüm bilim literatürünün ve bilimin tüm verisinin internet üzerinden erişilebilir olduğu; birbirleri ile konuşabildiği bir ortam

(Hey ve Diğerleri, 2009)



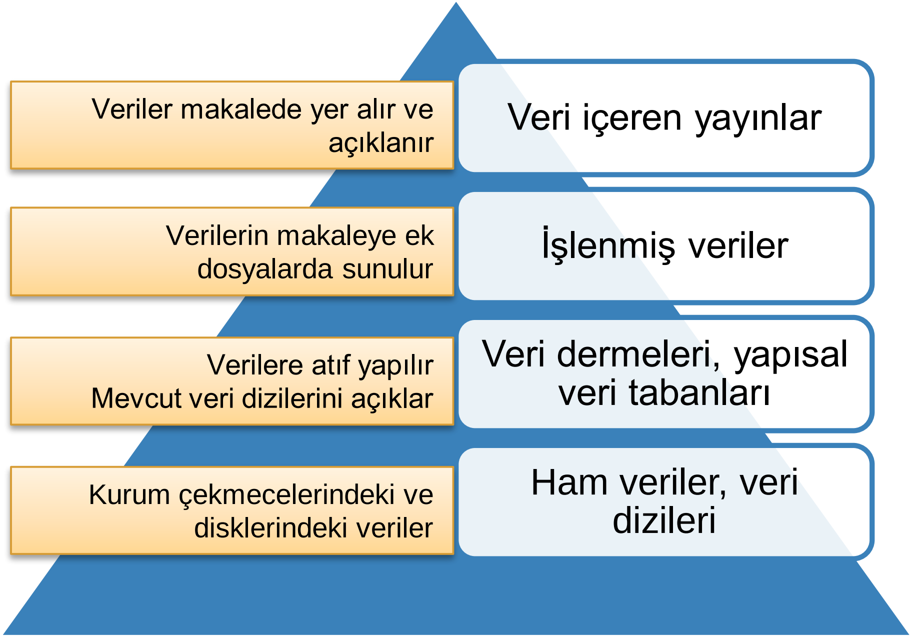
"Bilimsel çalışmalarda birincil kaynak olan ve araştırma sonuçlarını doğrulamak için kullanılan sayısal çıktılar, metinsel kayıtlar, görseller ya da sesler gibi maddi kayıtlar

«Lab notları, ilk analizler, makale taslakları, araştırma planları, hakem değerlendirmeleri, kişisel yazışmalar veya fiziksel nesneler (ör., lab numuneleri, test hayvanları) bu tanımın dışında»

(OECD, 2007)

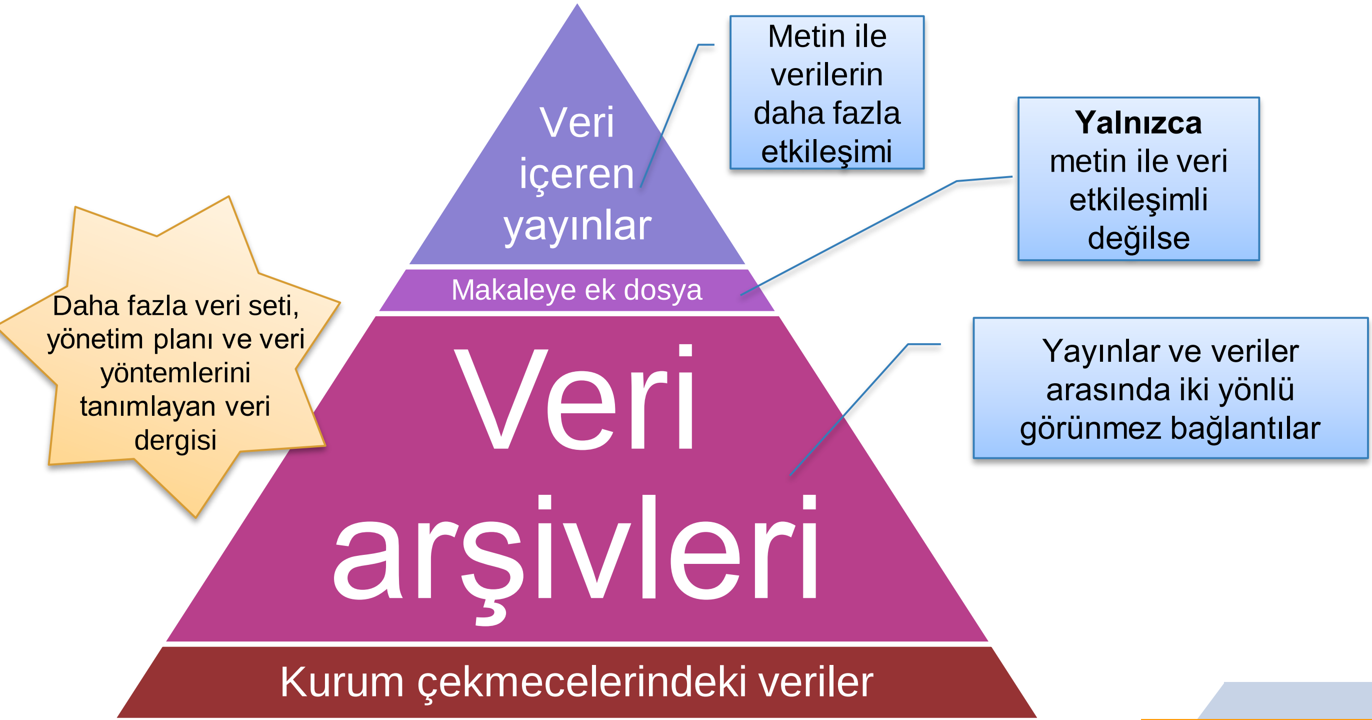
Erişilebilir?

Veri Yayın Piramidi



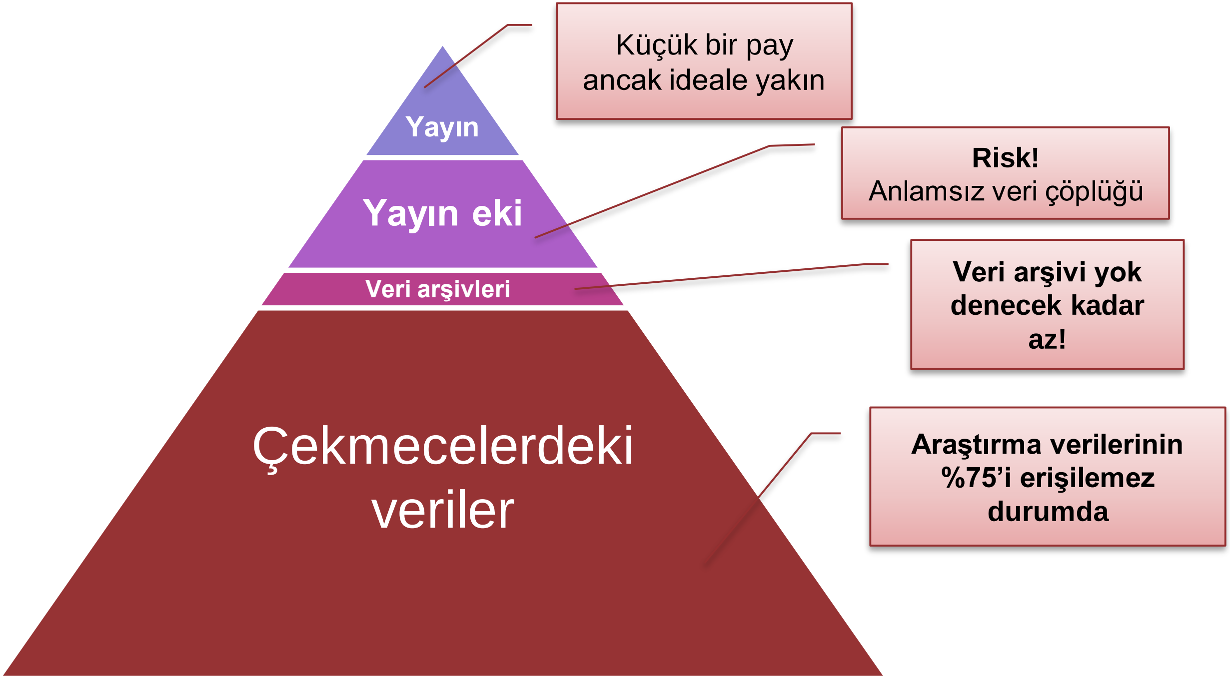
(LIBER Europe, 2012; çeviri: Tonta, 2014)

Veri Yayın Piramidi – İdeal Piramit



(LIBER Europe, 2012)

Veri Yayın Piramidi – Gerçekler



(LIBER Europe, 2012)

Verilerinizi paylaşır mısınız?

Ya verilerimi çalarlarsa?

Doğruluğu kanıtlanmamış veriyi paylaşmak ya birine zarar verirse?

Ya benden önce başkası benim verimle yayın yaparsa?

Benim verimi kim ne yapsın?

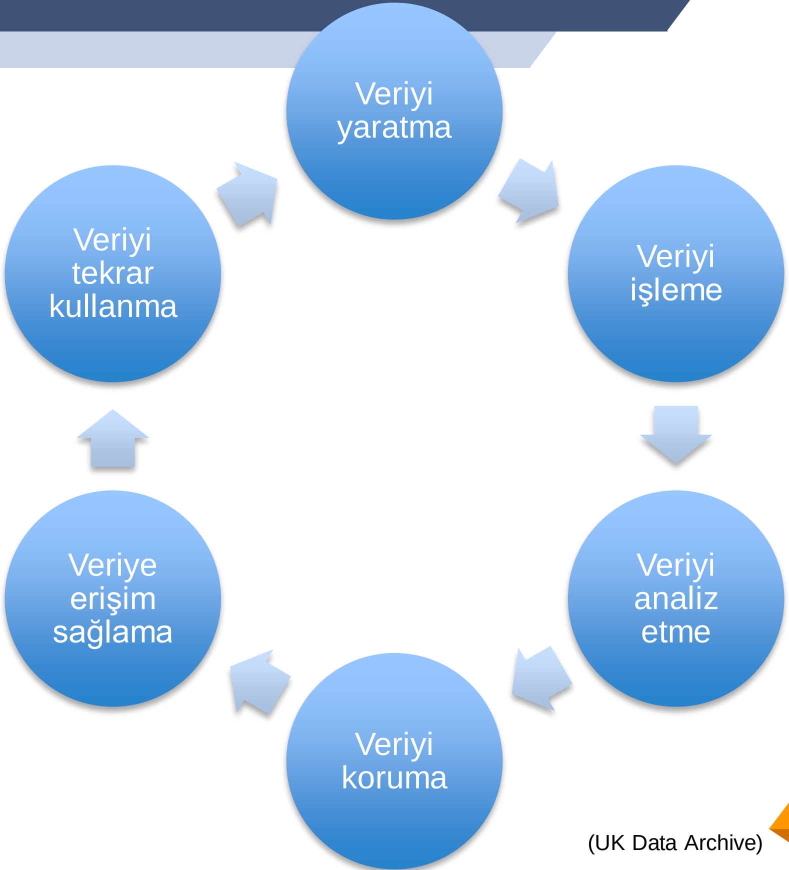
…

“ *Düzenli bir bilimsel yapı ile farklı alanlar aynı veri üzerinde birlikte çalışarak yeni bakış açıları geliştirebilir. Araştırmacılar verilerini tüm dünya ile verilerin bütünlüğünü veya sahipliğini kaybetmeden paylaşabilir. Verileri tekrar tekrar kullanabilir, birleştirebilir veya verimliliği artırabilir. Günümüz problemlerini çözmenin en kolay yolu bu…*

(High Level Expert Group on Scientific Data, 2010)

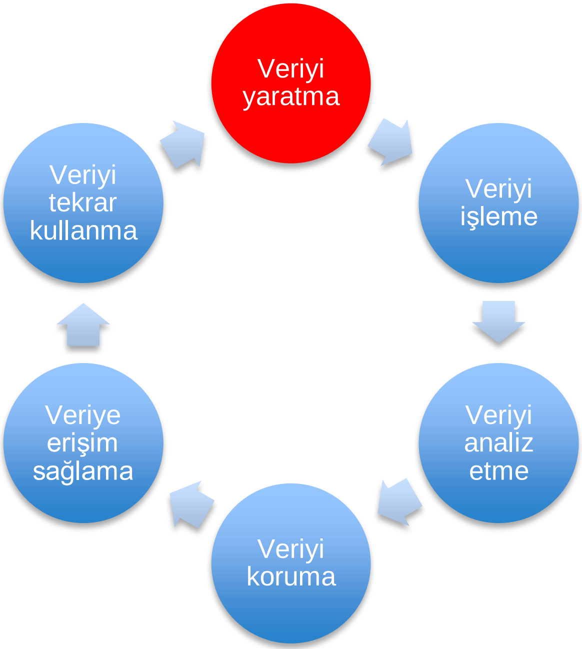
Nasıl?

Veri Yönetimi Döngüsü



(UK Data Archive)

Veriyi Yaratma



- Araştırmayı tasarla
- Veri yönetimini planla (biçim, saklama vb.)
- Verinin paylaşım izinlerini planla
- Mevcut verileri bul
- Veri topla (deney, gözlem, ölçüm, simülasyon vb.)

Veriyi Yarat: Araştırma Tasarımı

- Neyin, nasıl, hangi yöntemle araştırılacağının tasarlanması
- Araştırma için en iyi yol, yöntem ne?

Veriyi Yarat: Veri Yönetimini Planla

- Araştırmada kullanılacak veri hangi formatta saklanacak?
- Nerede saklanacak?
- «Yeni bir ayakkabı almadan önce ayakkabılıkta yer açmak»

Veriyi Yarat: Veri Paylaşım İzinlerini Planla

- Verinin sahibi kimler olacak?
- Veriye kimler erişebilecek?

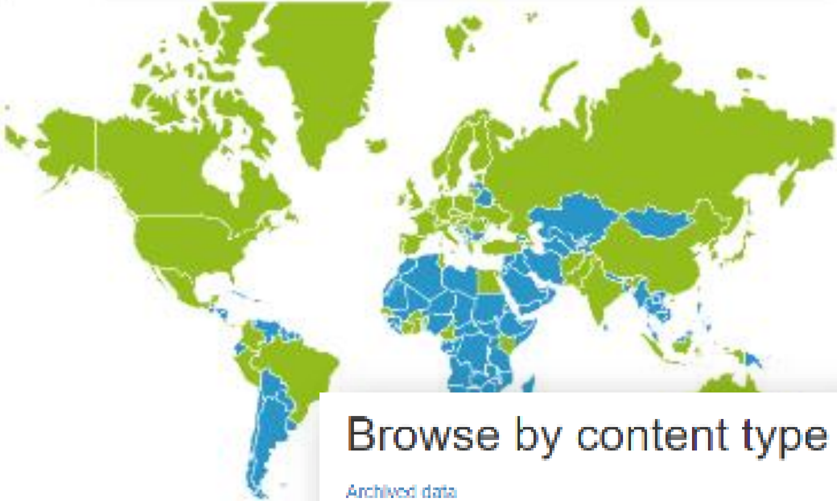
Veriyi Yarat: Mevcut Verileri Bul

- Daha önce ilgili veri toplanmış olabilir mi?
- Mevcut veriler nereden bulunabilir?

Veriyi Yarat: Mevcut Verileri Bul

Browse by country

Graphical | Text



Browse by content type

- Archived data
- Audiovisual data
- Configuration data
- Databases
- Images
- Networkbased data
- Plain text
- Raw data
- Scientific and statistical data formats
- Software applications
- Source code
- Standard office documents
- Structured graphics
- Structured text
- other

DataCite

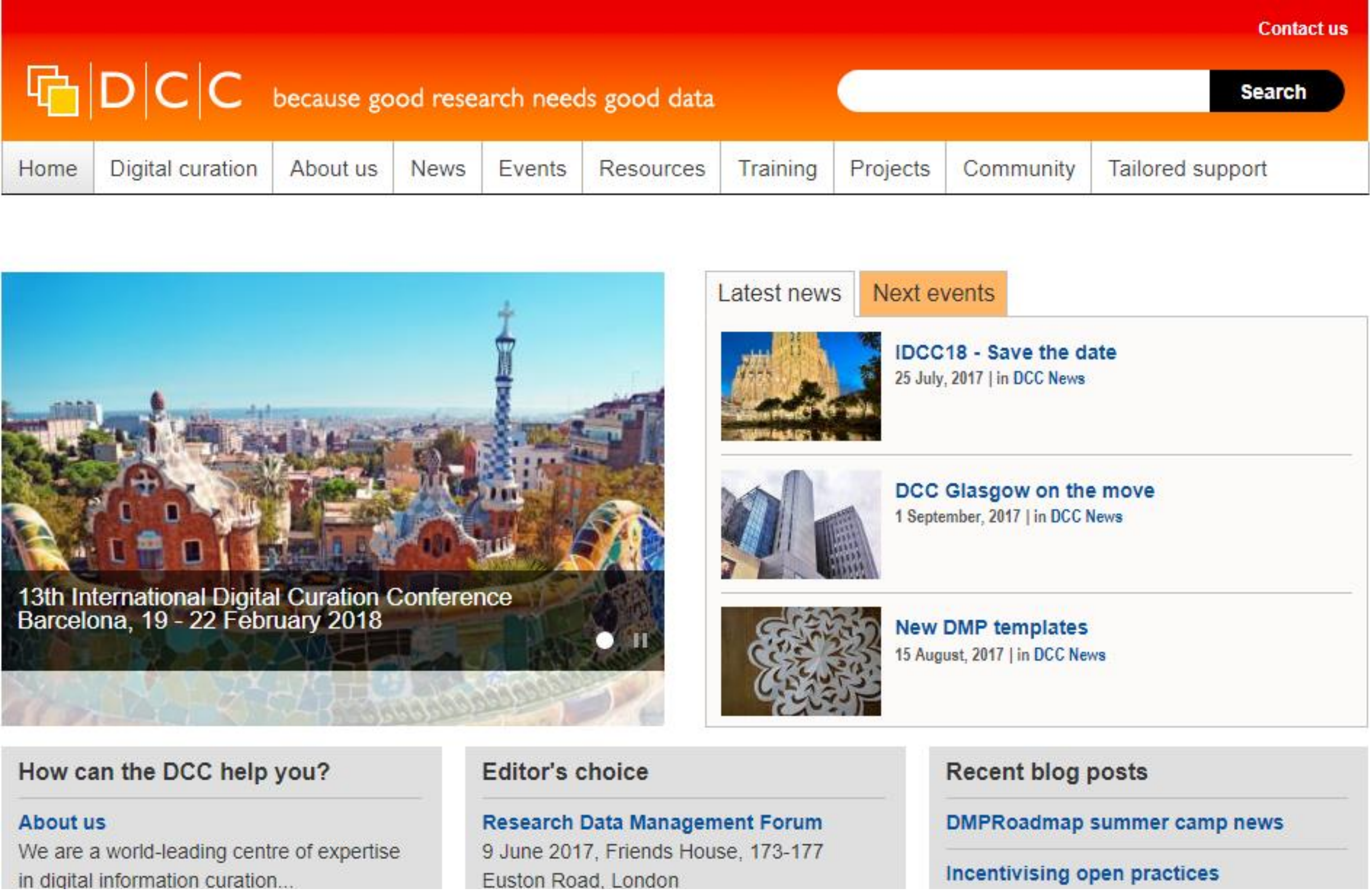
About us · Services · Resources · Community · Become a member

Find what you're looking for by searching millions of records with extensive, reliable metadata.

Share your data and reuse the data of others to create the highest impact in the research community.

Cite your research sources with confidence, and receive proper credit when your work is reused.

Connect your research – publications, datasets, software, authors, institutions, and funding data all in one place.


Contact us
DCC
because good research needs good data
Search
Home
Digital curation
About us
News
Events
Resources
Training
Projects
Community
Tailored support
13th International Digital Curation Conference
Barcelona, 19 - 22 February 2018
Latest news
Next events
IDCC18 - Save the date
25 July, 2017 | in DCC News
DCC Glasgow on the move
1 September, 2017 | in DCC News
New DMP templates
15 August, 2017 | in DCC News
How can the DCC help you?
About us
We are a world-leading centre of expertise
in digital information curation...
Editor's choice
Research Data Management Forum
9 June 2017, Friends House, 173-177
Euston Road, London
Recent blog posts
DMPRoadmap summer camp news
Incentivising open practices

About Zenodo

Passionate about Open Science!

Built and developed by researchers, to ensure that everyone can join in Open Science.

The OpenAIRE project, in the vanguard of the open access and open data movements in Europe was commissioned by the EC to support their nascent Open Data policy by providing a catch-all repository for EC funded research. CERN, an OpenAIRE partner and pioneer in open source, open access and open data, provided this capability and Zenodo was launched in May 2013.

In support of its research programme CERN has developed tools for Big Data management and extended Digital Library capabilities for Open Data. Through Zenodo these Big Science tools could be effectively shared with the long-tail of research.

Open Science knows no borders!

The need for a catch-all is not restricted to one funder, or one nation, so the concept caught on, and Zenodo rapidly started welcoming research from all over the world, and from every discipline.

The digital revolution has necessitated a retooling of the scholarly processes to handle data and software, but this is proceeding at varying speeds across different communities, disciplines, and nations. To ensure no one is left behind through lack of access to the necessary tools and resources, Zenodo makes the sharing, curation and publication of data and software a reality for all researchers.

zenodo | About

About Blog Help Developers

About Privacy Policy Terms of Use General Policies Infrastructure Principles Contact

Terms of Use v1.0

The Zenodo service ("Zenodo") is offered by CERN as part of its mission to make available the results of its work.

Use of Zenodo, both the uploading and downloading of data, denotes agreement with the following terms:

1. Zenodo is a research data repository for the preservation and making available of research, educational and informational content. Access to Zenodo's content is open to all, for non-military purposes only.
2. Content may be uploaded free of charge by those without ready access to an organized data centre.
3. The uploader shall ensure that their content is suitable for open dissemination, and that it complies with these terms and applicable laws, including, but not limited to, privacy, data protection and intellectual property rights.
4. All content is provided "as-is". Users of content ("Users") shall respect applicable license conditions. Download and use of content from Zenodo does not transfer any intellectual property rights in the content to the User.
5. Users are exclusively responsible for their use of content, and shall hold CERN free and harmless in connection with their download and/or use.
6. CERN reserves the right, without notice, at its sole discretion and without liability, (i) to alter or delete inappropriate content, and (ii) to restrict or remove User access where it considers that use of Zenodo interferes with its operations or violates these Terms of Use or applicable laws.
7. Unless specified otherwise, Zenodo metadata may be freely reused under the CC0 waiver. Bulk downloading of email addresses taken from Zenodo is not allowed.
8. These Terms of Use are subject to change by CERN at any time and without notice, other than through posting the updated Terms of Use on the Zenodo website.

If you have any questions or comments with respect to Zenodo, or if you are unsure whether your intended use is in line with these Terms of Use, or if you seek permission for a use that does not fall within these Terms of Use, please contact us.

wiei70-72157633022909105/

August 24, 2017 (v2) Dataset Open Access

View

Aligned ISNI and Ringgold identifiers for institutions

Delpeuch, Antonin

This dataset provides a correspondence between ISNI and Ringgold identifiers, by combining two datasets: Open ISNI for Institutions, available at http://isni.ringgold.com/ , which provides metadata for institutions identified by ISNI. The dataset of institutions used by ORCID for disambiguation,...

Uploaded on August 24, 2017

1 more version(s) exist for this record

August 22, 2017 (v2) Dataset Open Access

View

Supplementary Data: Status of the scalar singlet dark matter model (arXiv:1705.07931)

The GAMBIT Collaboration

Supplementary Data Status of the scalar singlet dark matter model arXiv:1705.07931 The files in this record contain data for the scalar singlet dark matter model considered in the GAMBIT "Round 1" scalar singlet paper. The files consist of Three YAML files, each corresponding to a different...

Uploaded on August 22, 2017

1 more version(s) exist for this record

More

August 22, 2017

Dataset Open Access

Supplementary Data: Status of the scalar singlet dark matter model (arXiv:1705.07931)

The GAMBIT Collaboration

Supplementary Data

Status of the scalar singlet dark matter model
arXiv:1705.07931

The files in this record contain data for the scalar singlet dark matter model considered in the GAMBIT "Round 1" scalar singlet paper.

The files consist of

- Three YAML files, each corresponding to a different parameter range
- StandardModel_SLHA2_SingletDM_scan_15.yaml, a universal YAML fragment included from the other three YAML files
- Three hdf5 files. SingletDM.hdf5 contains the combined results of all sampling runs, and is the basis for the profile likelihood plots in the paper. SingletDM_TW_full.hdf5 and SingletDM_TW_lowmass.hdf5 contain the results from T-Walk scans over the full and low-mass parameter ranges, respectively. These are the bases for the marginalised posterior plots in the paper.
- An example pip file corresponding to each hdf5 file, for producing plots using pippi
- A tarball best_fits_yaml.tar.gz containing YAML files of the best-fit point in each subregion of the fit.

The YAML files corresponding to different parameter ranges follow the naming scheme SingletDM_[slice].yaml, where slice may be full, lowmass or neck. Each of these YAML files contains entries in the Scanners node for running Diver, MultiNest, TWalk and GreAT.

A few caveats to keep in mind:

1. The YAML files that we give here are updated compared to the ones that we used when generating the hdf5 file, in order to match the set of available options in the release version of GAMBIT 1.0.0. The included physics and numerics are however identical.
2. The YAML files are designed to work with the tagged release of GAMBIT 1.0.0, and the pip file is tested with pippi 2.0, commit 2ab061a8. They may or may not work with later versions of either software (but you can of course always obtain the version that they do work with via the git history).
3. The pip file is an example only. Users wishing to reproduce the more advanced plots in any of the GAMBIT papers should contact us for tips or scripts, or experiment for themselves. Many of these scripts are in multiple parts and require undocumented manual interventions and steps in order to implement various plot-specific customisations, so please don't expect the same level of polish as for files provided here or in the GAMBIT repo.

Publication date:
August 22, 2017

DOI:
DOI 10.5281/zenodo.846860

Keyword(s):
dark matter | beyond standard model | global fits | particle physics phenomenology | Higgs physics

Related identifiers:
Supplement to:
arXiv:1705.07931

Communities:
GAMBIT (Collaboration)
INSPIRE HEP
Zenodo

Versions

Version 2	10.5281/zenodo.846860	Aug 22, 2017
Version 1	10.5281/zenodo.801511	Jun 1, 2017

Cite all versions? You can cite all versions by using the DOI 10.5281/zenodo.801510. This DOI represents all versions, and will always resolve to the latest one. Read more.

v2 adds YAML benchmark files for the best-fit points in each region of the model.

Preview

The files in this record contain data for the scalar singlet dark matter model considered in the GAMBIT "Round 1" scalar singlet paper.

The files consist of

- Three YAML files, each corresponding to a different parameter range
- `StandardModel_SLHA2_SingletDM_scan_15.yaml`, a universal YAML fragment included from the other three YAML files
- Three hdf5 files. `SingletDM.hdf5` contains the combined results of all sampling runs, and is the basis for the profile likelihood plots in the paper. `SingletDM_TW_full.hdf5` and `SingletDM_TW_lowmass.hdf5` contain the results from T-Walk scans over the full and low-mass parameter ranges, respectively. These are the bases for the marginalised posterior plots in the paper.
- An example pip file corresponding to each hdf5 file, for producing plots using pippi
- A tarball `best_fits_yaml.tar.gz` containing YAML files of the best-fit point in each subregion of the fit.

The YAML files corresponding to different parameter ranges follow the naming scheme `SingletDM_[slice].yaml`, where slice may be `full`, `lowmass` or `neck`. Each of these YAML files contains entries in the Scanners node for running Diver, MultiNest, TWalk and GreAT.

A few caveats to keep in mind:

1. The YAML files that we give here are updated compared to the ones that we used when generating the hdf5 file, in order to match the set of available options in the release version of GAMBIT 1.0.0. The

Share

Cite as

The GAMBIT Collaboration. (2017). Supplementary Data: Status of the scalar singlet dark matter model (arXiv:1705.07931) [Data set]. Zenodo. http://doi.org/10.5281/zenodo.846860

Start typing a citation style...

Export

BibTeX CSL DataCite Dublin Core JSON MARCXML Mendeley

Files (25.3 GB)

Name	Size	
best_fits_yaml.tar.gz md5:5e272d4baaef0eab86d692fbb0e44e2e	2.4 kB	Download
README_Singlet.md md5:b6c89c910063d7710b770f66e303fa17	2.5 kB	Preview Download
SingletDM.hdf5.tar.gz md5:e728ac709e7888a562e705635b0a3fd2	21.0 GB	Download
SingletDM.pip md5:e8c1173f02b1a963b4e0d70f8399db34	6.1 kB	Download
SingletDM_full.yaml md5:6d90cafb33dc17f6392034a7c671a9a3	4.9 kB	Download
SingletDM_lowmass.yaml md5:2eccc32e7eb58d350f30d75fd17ac2b8	4.9 kB	Download
SingletDM_neck.yaml md5:17c5e0f2e20deb19779c7e1dc2babd7f	4.6 kB	Download
SingletDM_TW_full.hdf5.tar.gz md5:b9f5229a0ae43574942c79ef7a2ed50d	2.5 GB	Download
SingletDM_TW_full.pip md5:03a17f095d1d8d186a3e20ca27a8c41c	6.3 kB	Download
SingletDM_TW_lowmass.hdf5.tar.gz md5:57188e42f87f5cc40fd5aac5412ae613	1.9 GB	Download
SingletDM_TW_lowmass.pip md5:45ffad7d04ccba4a0bbfd12020195078	5.9 kB	Download
StandardModel_SLHA2_SingletDM_scan_15.yaml md5:cfdf95a1e1b36056bd450102b5b94cf8	2.9 kB	Download

INVENIO

Showcase Getting Started Community

A showcase gallery illustrating some of the Invenio instances. Add yours!

All | Data | Library | Multimedia | Repository

CERN Document Server

CERN official repository for publications, articles, reports and multimedia content in HEP.

Library Multimedia Repository

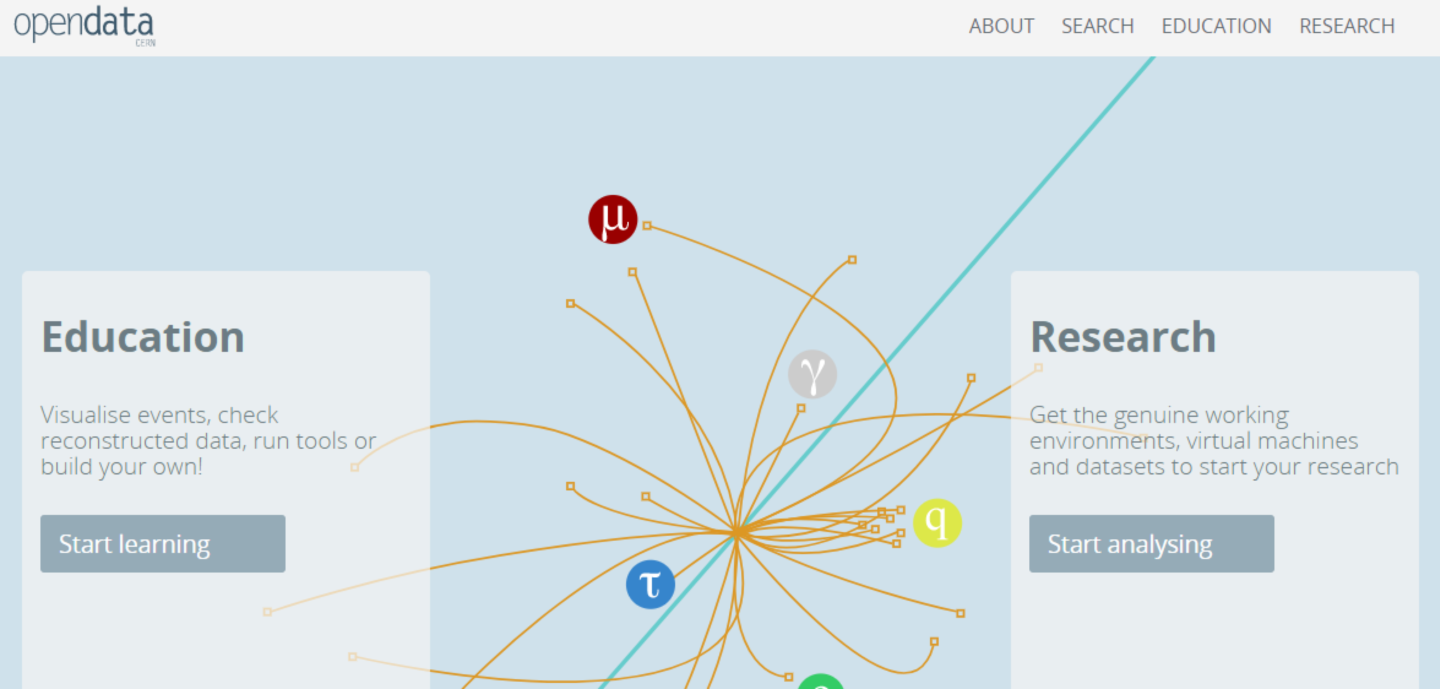
CERN Open Data

Discover open Research datasets and software of LHC experiments. Visualise events and run and your own analysis.

Data

HEPData

The Durham High Energy Physics Database (HEPData) has been built up over the past four decades as a unique open-access repository for scattering data from experimental particle physics. It currently comprises the data points from plots and


opendata
CERN
ABOUT
SEARCH
EDUCATION
RESEARCH
μ
γ
q
τ
Education
Visualise events, check reconstructed data, run tools or build your own!
Start learning
Research
Get the genuine working environments, virtual machines and datasets to start your research
Start analysing

Web of Science

Search

Select a database Data Citation Index | Learn More

Basic Search | Cited Reference Search | Advanced Search

Example: Crime Survey | Topic | Search

+ Add Another Field | Reset Form

Results: 580
(from Data Citation Index)

You searched for: ADDRESS: (turkey)
Timespan: All years. **Indexes:** DCI-S, DCI-SSH.
...Less

Results: 3,009
(from Data Citation Index)

You searched for: ADDRESS: (greece)
Timespan: All years. **Indexes:** DCI-S, DCI-SSH.
...Less

Results: 1,121,482
(from Data Citation Index)

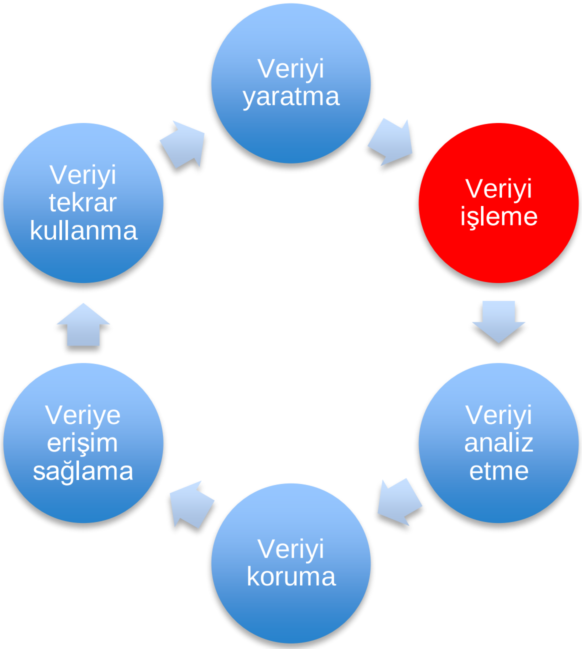
You searched for: ADDRESS: (usa)
Timespan: All years. **Indexes:** DCI-S, DCI-SSH.
...Less

Create Alert

Veriyi Yarat: Veri Topla

- Hangi veri toplama yöntemlerini kullanacağına karar ver!
- Deney, gözlem, anket, ölçüm vb.

Veriyi işleme



- Veriyi elde et, dijitalleştir, çevir
- Kontrol et, doğrula, temizle
- Verinin anonimleştirilmesi gerekiyorsa anonimleştir
- Veriyi tanımla
- Veriyi yönet ve sakla

Veriyi İşleme: Elde Et, Kodla, Dijitalleştir, Çevir

- En doğru veriyi, en doğru yöntemle ve en doğru formatla elde et

En Yaygın Veri Formatları

Veri türü	**Tercih edilen dosya formatı**
Kapsamlı niceliksel veriler	SPSS, R, Phyton, SAS, MS Access
Kapsamlı olmayan niceliksel veriler	.csv, .tab, .xls, .txt, .dbf
Niteliksel veriler	XML, .rtf, Unicode txt, Latex
Coğrafik veriler	ESRI, .dwg, .tif, .tfw, .mdb
Dijital görsel veriler	.tif, .jpeg, .png
Dijital ses verileri	.flac, .mpeg, .mp3, .wav
Dijital video verileri	.mp4, jp2, mj2
Diğer	.html, .odt, .doc, .pdf

Gelecekte Tercih Edilecek Veri Formatlarının Özellikleri

- Birine (kişi/kurum) ait olmayan,
- Açık ve belgelenmiş standarda sahip,
- Araştırma topluluğu tarafından ortak kullanılabilen,
- Standart karakter kodlamalarını kullanan,
- Sıkıştırılmamış

http://guides.library.oregonstate.edu/research-data-services/data-management-types-formats

Türkiye Akademisyenlerinin Veri Tipi ve Formatı Tercihi

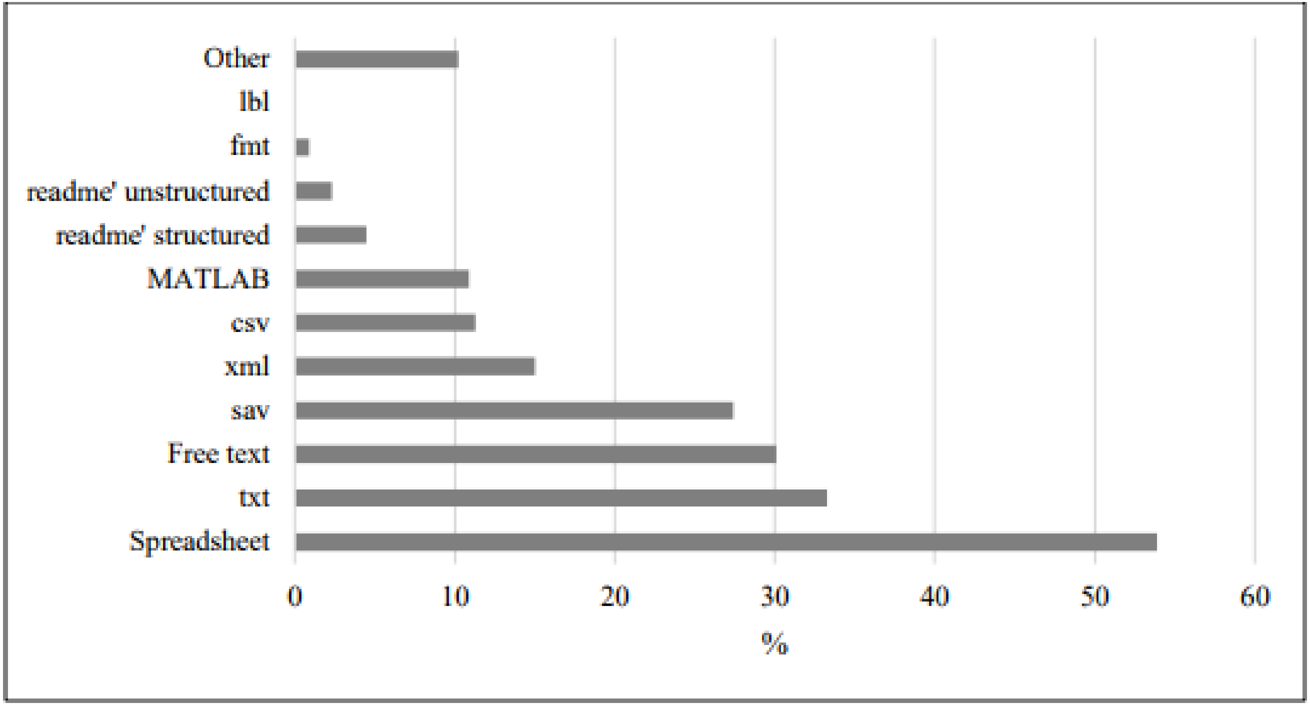


Figure 3. Research data formats.

Veriyi İşleme: Kontrol Et, Düzenle, Doğrula, Temizle

- Veriyi organize et!
- Dosya hiyerarşisinden dosya adlarına kontrol
- Nasıl bir araştırma aktivitesi sonrası nasıl bir veri toplandı?
- Versiyon bilgileri
- Sistematik kodlama
- Zaman damgası

Veriyi İşleme: Gerekiyorsa Veriyi Anonimleştir

DATA + FOLLOW THIS TOPIC

Why you can't really anonymize your data

It's time to accept and work within the limits of data anonymization.

By Pete Warden, May 17, 2011

One of the joys of the last few years has been the flood of real-world datasets being released by all sorts of organizations. These usually involve some record of individuals' activities, so to assuage privacy fears, the distributors will claim that any personally-identifying information (PII) has been stripped. The idea is that this makes it impossible to match any record with the person it's recording.

Something that my friend Arvind Narayanan has taught me, both with theoretical papers and repeated practical demonstrations, is that this anonymization process is an illusion. Precisely because there are now so many different public datasets to cross-reference, any set of records with a non-trivial amount of information on someone's actions has a good chance of matching identifiable public records. Arvind first demonstrated this when he and his fellow researcher took the "anonymous" dataset released as part of the first Netflix prize, and demonstrated how he could correlate the movie rentals listed with public IMDB reviews. That let them identify

Veriyi İşleme: Veriyi Tanımla

- Veri ne hakkında?
- Veri hakkında veri: Üst veri (metadata)
- Ne sağlar?
- Üst veri standartları

Innovation in metadata design, implementation & best practices

Dublin Core Metadata Initiative

home | News | DCMI Specifications | Community and Events | Join / Support | About

```
<title> </title>
<dc:creator> </dc:creator>
<dc:subject type="LCSH"> </dc:subject>
<dc:description> </dc:description>
<dc:publisher>Case Western Reserve University, digital content</dc:publisher>
<dc:contributor> </dc:contributor>
<dc:date.original> </dc:date.original>
<dc:date.modified scheme="W3CDTF"> </dc:date.modified>
<dc:type scheme="DCMIType"> </dc:type>
<dc:type> </dc:type>
<dc:relation> </dc:relation>
<dc:source> </dc:source>
<dc:format> </dc:format>
<dc:format.medium scheme="IMT"> </dc:format.medium>
<dc:identifier>LEAVE BLANK FOR NOW</dc:identifier>
<dc:language scheme="ISO 639-1"> </dc:language>
<dc:coverage> </dc:coverage>
<dc:rights></dc:rights>
```

Document, Discover and Interoperate


```xml
<codeBook version="1.2.2" ID="ZAF_2010_NIDS_v01_M" xml-lang="en" xmlns="http://www.icpsr.umich.edu/DDI" xml
  <docDscr>
    <citation>
      <titlStmt>
        <titl>
          ZAF_2010_NIDS_v01_M
        </titl>
        <IDNo>
          DDI_ZAF_2010_NIDS_v01_M
        </IDNo>
      </titlStmt>
      <rspStmt>
        <othId role="Primary Investigator" affiliation="University of Cape Town">
          <p>
            Southern Africa Labour and Development Research Unit
          </p>
        </othId>
      </rspStmt>
      <prodStmt>
        <producer abbr="SALDRU" affiliation="University of Cape Town" role="Documentation of the study">
          Southern Africa Labour and Development Research Unit
        </producer>
        <producer abbr="DataFirst" affiliation="University of Cape Town" role="Creation of the DDI">
          DataFirst
        </producer>
        <prodDate date="2012-05-09">
```

FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

HOME | WHO WE ARE | WHAT WE DO | INITIATIVES | ORGANIZATION | RESOURCES

GEOSPATIAL METADATA

Metadata is information abou... where, how, and why of a data... imagery, and other location b... GeoPlatform.gov) to National... Supplemental Guidance to su... discovery and access.

ABOUT GEOSPATIAL

The basics of metadata

- Geospatial Metadata Fa...
- Business Case for Metad...
- Value of Metadata (pdf, p...
- What is Metadata (pdf, p...

ISO ... serv... the quality, the spatial and tem... portrayal, distribution, and oth...

ISO 19115-1:2014 is applicabl...

-the cataloguing of all types of r... of datasets and services;

-geographic services, geograph... features and feature properties

ISO 19115-1:2014 defines:

-mandatory and conditional me...

FGED
Functional Genomics Data Society

Home | About Us | Mission | Projects | Conferences

MINSEQE

Minimum Information about a high-throughput SEQuencing Experiment

MINSEQE describes the **Minimum Information about a high-throughput nucleotide SEQuencing Experiment** that is needed to enable the unambiguous interpretation and facilitate reproduction of the results of the experiment. By analogy to the MIAME guidelines for microarray experiments, adherence to the MINSEQE guidelines will improve integration of multiple experiments across different modalities, thereby maximising the value of high-throughput research. [MINSEQE version 1.0 (pdf), FGED Society (June 2012)]

The five elements of experimental description considered essential when making data available supporting published high-throughput sequencing experiments are as follows:

1. **The description of the biological system, samples, and the experimental variables being studied.**
   - "compound" and "dose" in dose response experiments or "antibody" in ChIP-Seq experiments, the organism, tissue, and the treatment(s) applied.

... / Geospatial Metadata

...atial Metadata
...istry
...Standards
...m / Data.gov
...dations (pdf)
...eospatial Data
...A) Metadata
...(pdf)
...adata Working
...G)
...atial Metadata
...ation Forum

4

MIBBI - Minimum Information for Biological and Biomedical Investigations

A common portal to a group of nearly 40 checklists of Minimum Information for various biological disciplines. The MIBBI Foundry is developing a cross-analysis of these guidelines to create an intercompatible, extensible community of standards.

The concept was realized initially through the joint efforts of the Proteomics Standards Initiative, the Genomic Standards Consortium and the MGED RSBI Working Groups. The latest project to register with MIBBI is the MIABie guidelines for reporting biofilm research, as of January 2012.

Related Vocabularies	Ontology for Biomedical Investigations
Specification	http://mibbi.sourceforge.net/foundry.shtml
Standard's website	http://mibbi.sourceforge.net/portal.shtml

Extensions

MIBBI Portal
A list of nearly 40 Minimum Information standards projects registered with the MIBBI initiative.

Tools

ProteoRed Tools
Bioinformatics tools to create and extract metadata compliant with the MIBBI-registered MIAPE minimum requirements.

Use Cases

BioModels Database
A repository hosting computational models of biological systems, using the MIBBI-registered MIRIAM and MIASE minimal metadata requirements.

CARMEN
A a virtual laboratory for neurophysiology, enabling sharing and collaborative exploitation of data, analysis, code and expertise. Metadata must include the MIBBI-registered MINI recommendations.

FlowRepository
A database of flow cytometry experiments where you can query and download data collected and annotated according to the MIBBI-registered MIFlowCyt standard.

International Molecular Exchange Consortium
An international collaboration to provide access to a non-redundant set of protein-protein interaction data from a broad taxonomic range of organisms. IMEx partner databases require data to be MIMIx (a MIBBI-registered standard) compatible.

NEBC ISA Network BioInvestigationIndex
The NERC Environmental Bioinformatics Centre ISA Network's index of ISA-Tab and MIBBI-compliant environmental 'omics data.

PRIDE - PRoteomics IDEntifications database
A centralized, MIBBI standards compliant, public data repository for proteomics data, post-translational modifications and supporting spectral evidence.

Türkiye Akademisyenlerinin Üst Veri Standart Tercihi

- Katılımcıların %27'si üst verinin ne anlama geldiğini bilmiyor.
- Kendi geliştirdikleri üst veri standartlarını kullandıklarını söylüyorlar (%13).
- ISO %8,8, AWM, DwC, DIF, EML, NISO gibi diğer standartlar ise %1'in altında.

Veriyi İşleme: Veriyi Yönet ve Sakla

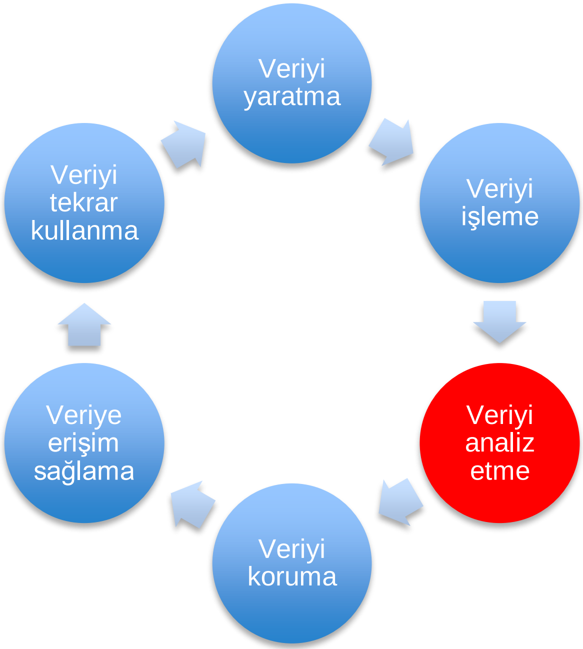
- Veriyi yedekleme ve koruma
- Şifreleme
- Güncelleme
- Güvenli internet protokolleri
- Sanal sunucuların güvenliği
- Uzaktan erişim

Türkiye Akademisyenlerinin Veri Saklama Tercihleri

Table III. Places to store data

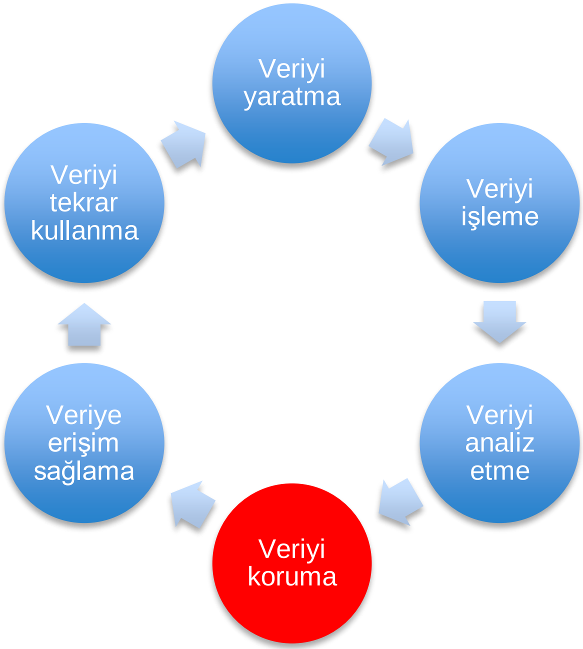
Medium	n	%
Local computers	381	71.6
Cloud	244	45.9
Open access data repository	44	8.3
Institutional open repository	17	3.2
Commercial data repository	2	0.4

Veriyi Analiz Etme



- Veriyi yorumla
- Araştırma çıktısı üret
- Yayın yap
- Veriyi koruma için hazırla

Veriyi Koruma



- Veriyi en iyi formata getir
- Veriyi en uygun ortama taşı
- Yedekle ve sakla
- Üst veriyi hazırla
- Veriyi arşivle
- Provenans

Veriyi Koruma: Provenans

Provenance Principles for Open Data

a)

Transport

Transport data covers the transport infrastructure and data about traffic flow. *Note: the data provided here dates from March 2010.*

Development

- National Public Transport Gazetteer
- National Public Transport Access Nodes

Search

- SPARQL Endpoint provided by Talis
- Search provided by Talis
- SPARQL Endpoint provided by TSO

Guides

- Guide to SPARQLing Transport Data

b)

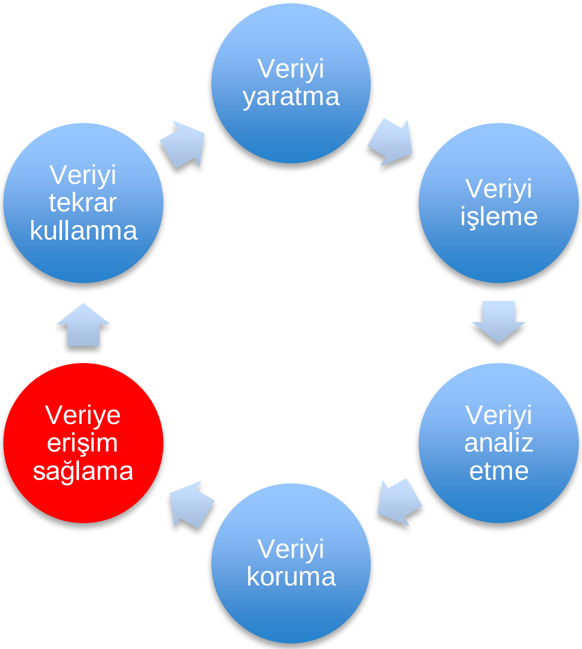

```
{  "name": "naptan",
   "author": "Transport Direct",
   "author_email": " naptandatamanager@landmark.co.uk",
   "maintainer": null, "maintainer_email": null,
   "extras": {
     "date_released": "2010-07-01",
     "date_updated": "2012-01-03",
     "temporal_coverage-from": "2011-04-01",
     "temporal_coverage-to": "2011-10-04",,
     "agency":"Transport Direct",
     "published_via":"Transport Direct [11663]",
     "published_by":"Department for Transport [11406]",
     "department":"Department for Transport",
     "import_source":"COSPREAD-cospread-2010-03-19-part1.csv",...}}
```


General Terms
Theory, Documentation, Management

Keywords
provenance, open data, linked data

The Provenance and Linked Open Data mini-theme[2] was an activity supported by the UK e-Science Institute[3] which was investigating provenance challenges in the context of Linked Open Data. As an outcome of a series of workshops organised under this activity we have identified and discussed a set of principles for publishing provenance of open data similar to the Linked Open Data rules discussed by Berners-Lee[1].

Erişim Sağlama



- Veriyi yay
- Veriyi paylaş
- Erişimi kontrol et
- Entelektüel mülkiyet haklarını al
- Veriyi duyur

Erişim Sağlama: Entelektüel Mülkiyet Hakları

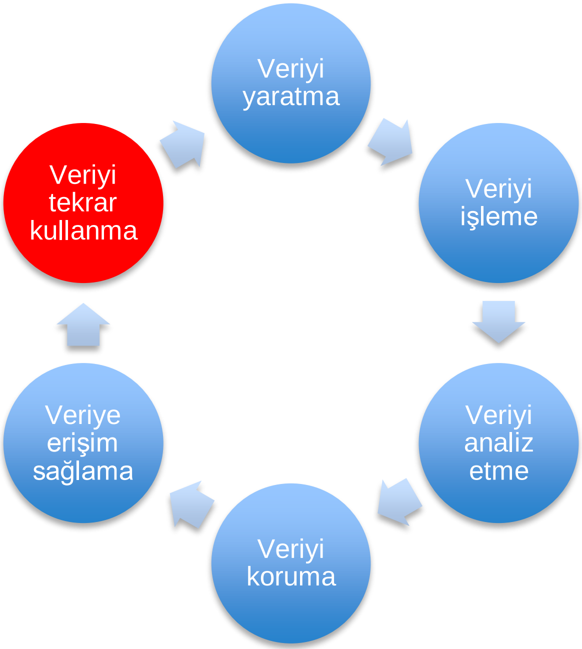
- Verinin telifi kime ait?
- Fonlayıcı veya üniversitelerin hakları
- Üçüncü parti telif sahipleri (sağlayıcı veya yayıncılar gibi)

- Sorumluluk paylaşımı

Erişim Sağlama: Lisanslama

- Creative Commons
- Open Data Commons
- Open Government License

Veriyi Tekrar Kullanma



- Yeni araştırmalar
- Bulguların değerlendirilmesi
- Öğretim ve öğrenme
- Deneyimlerin tekrar edilmesi

İYİ YÖNETİLEN ARAŞTIRMA VERİLERİ İÇİN YOL HARİTASI

(Waard, Cousijn ve Aalbersberg, 2015)

Kaynakça

- Aydinoglu, A.U., Doğan, G. ve Taşkın, Z. (2017). Research data management in Turkey: Perceptions and practices. Library Hi-Tech, 32(2), 271-289.
- Hey, T., Tansley, S. ve Tole, K. (2009). The Fourth Paradigm: Data-intensive Scientific Discovery. http://research.microsoft.com/enus/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf
- High Level Expert Group on Scientific Data. (2010). Riding the Wave: How Europe Can Gain from the Rising Tide of Scientific Data. https://www.fosteropenscience.eu/content/riding-wave-how-europe-can-gain-rising-tide-scientific-data
- LIBER Europe. (2012). Research Data Sharing LERU. https://www.slideshare.net/libereurope/research-data-sharing-leru
- OECD. (2007). OECD Principles and Guidelines for Access to Research Data from Public Funding. http://www.oecd.org/sti/sci-tech/38500813.pdf
- Tonta, Y. (2014). Açık Erişim, Açık Veri. http://yunus.hacettepe.edu.tr/~tonta/yayinlar/tonta-mersin-6Subat2014-acik-erisim-paneli-v3-10-Subat-2014.pdf
- UK Data Archive. Research Data Lifecycle. http://www.data-archive.ac.uk/create-manage/life-cycle
- Waard, A., Cousijn, H. ve Aalbersberg, I.J. (2015). 10 Aspects of Highly Effective Research Data. https://www.elsevier.com/connect/10-aspects-of-highly-effective-research-data

İletişim için:
@zehrataskin
ztaskin@hacettepe.edu.tr