



Manajemen Data

PELATIHAN PENGUATAN TEKNIK SUBSTANTIF JABATAN FUNGSIONAL
ANALIS DATA ILMIAH (ADI) BAGI PEGAWAI LEMBAGA KEBIJAKAN
PENGADAAN BARANG/JASA PEMERINTAH (LKPP)

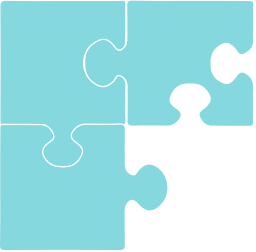
Kamis, 25 Juni 2026
08.00 – 10.15 WIB

Seno Yudhanto

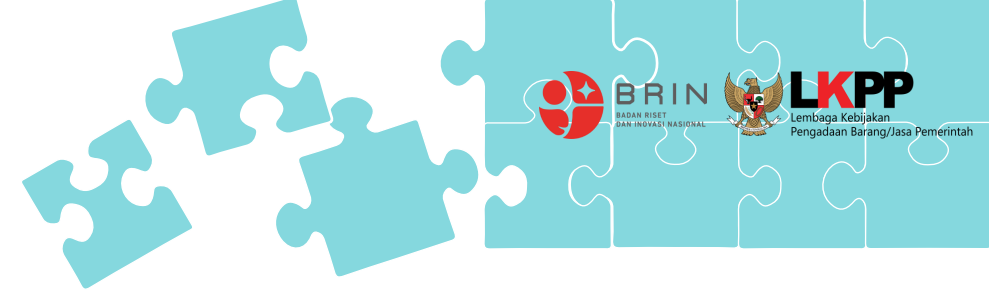
Seno Yudhanto



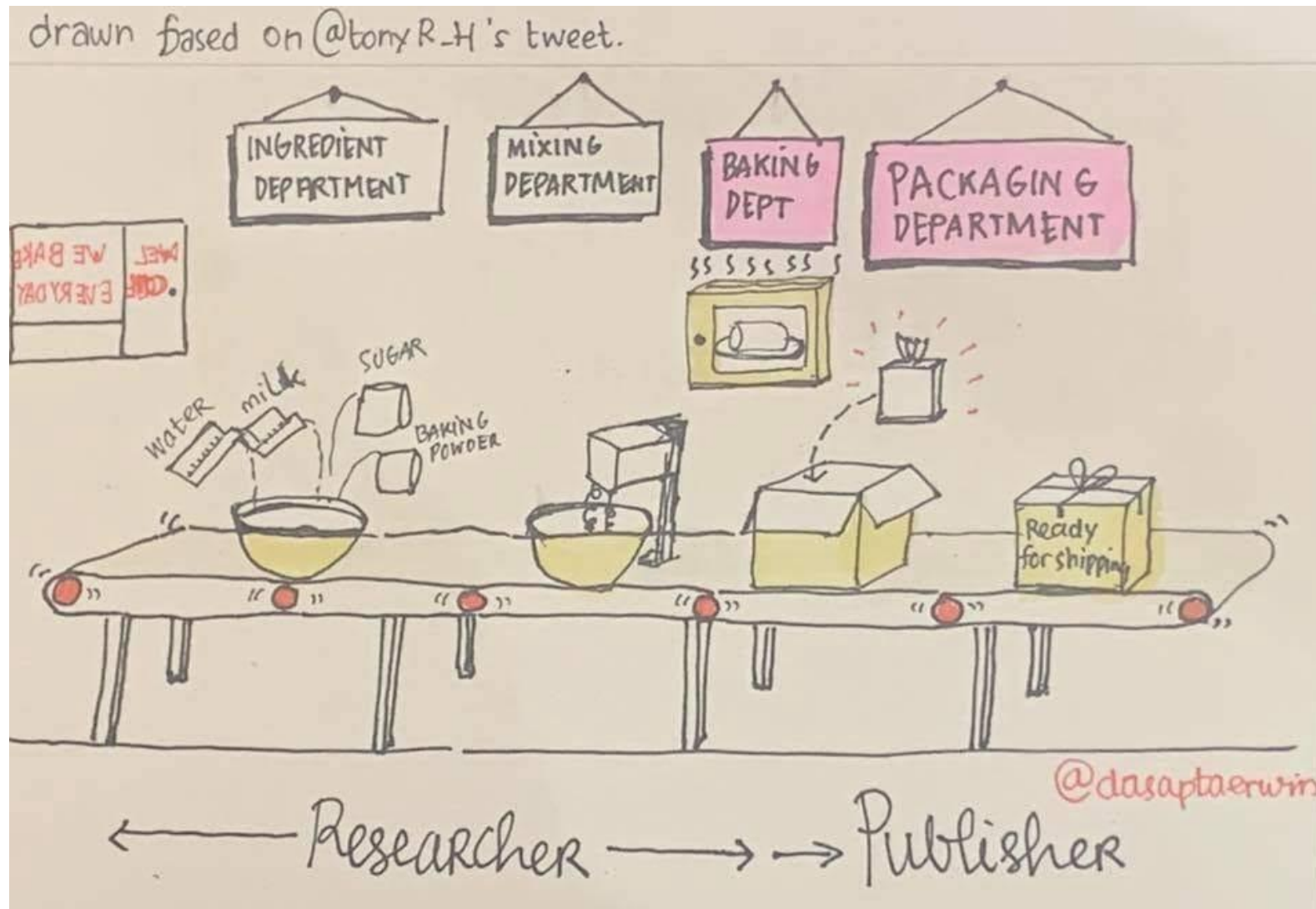
Pendidikan	S1 Ilmu Perpustakaan - Universitas Indonesia (2013) S2 Ilmu Perpustakaan - Universitas Indonesia (2022)
Jabatan	- Pustakawan Ahli Muda - Anggota Tim Pengelolaan Repositori Ilmiah Nasional - Editor BACA: Jurnal Dokumentasi dan Informasi
Unit Kerja	Direktorat Repositori, Multimedia, dan Penerbitan Ilmiah - Badan Riset dan Inovasi Nasional
Kontak	seno001@brin.go.id / yudhantoseno@gmail.com Gedung B.J. Habibie, Jl. M.H. Thamrin No. 8, Jakarta Pusat, DKI Jakarta 10340
E-profile	• https://orcid.org/0000-0003-2355-2281 • https://scholar.google.com/citations?user=3PTm_H0AAAAJ&hl=en • https://www.researchgate.net/profile/Seno-Yudhanto



Pendahuluan



- Deskripsi Mata Pelatihan → Menjelaskan tentang perencanaan manajemen data, siklus data dalam manajemen data, review konten data, pengenalan metadata, dokumentasi data, etika dan integritas data, serta sensitivitas data
- Tujuan Pembelajaran → Peserta pelatihan mampu memahami, mengaplikasikan, menganalisis, dan melakukan evaluasi laporan manajemen data.



Urgensi Manajemen Data

- Data yang dikelola dengan baik akan meningkatkan nilai tambah dari data tersebut. Semakin baik pengelolaannya, semakin berguna untuk analisis selanjutnya
- Komunikasi ilmiah tidak hanya melalui publikasi, melainkan juga data
- Sitasi/pengutipan tidak hanya output/luaran saja tetapi data primernya

DATA



SORTED



ARRANGED



PRESENTED
VISUALLY



EXPLAINED
WITH A STORY



ACTIONABLE
(USEFUL)



Risiko Penyimpanan Data

Laboratorium Influenza Unair Dimasuki Pencuri, 3 Laptop Peneliti Hilang

Imam Wahyudiarta - detikNews

Kamis, 07 Jan 2016 20:30 WIB

0 komentar

SHARE



Foto: Imam Wahyudiarta

Surabaya - Kampus C Universitas Airlangga (Unair) dibobol pencuri. Pelaku melakukannya di Gedung Institute of Tropical Disease (Lembaga Penyakit Tropis). Ada tiga laptop yang dibawa kabur pelaku.

Laptop yang dibawa kabur tersebut ada di ruang laboratorium influenza. Meski yang dicuri hanya laptop, namun kerugian yang dirasakan sungguh sangat besar, melebihi nilai mata uang.

"Laptop itu milik tiga profesor yang tengah melakukan riset penyakit infeksi dan virus. Satu diantara profesor itu adalah peneliti asal Jepang," ujar Rektor Unair Prof Muhammad Nasih, Kamis (7/1/2016).

Profesor asal Jepang tersebut adalah Profesor Shimizu. Sudah empat tahun Shimizu bekerjasama dengan Unair guna mengembangkan riset virus influenza. Satu temuan pentingnya adalah virus Avian Influenza pada unggas dan influenza pada manusia. Shimizu dan tim nya juga melakukan penelitian virus hepatitis.

"Itulah sebabnya kerugian ini sangat besar karena ini adalah hasil riset bertahun-tahun, yang jauh lebih mahal daripada laptop paling mahal," kata Nasih.

Nasih sendiri menaruh curiga pada kasus pencurian ini. Setidaknya ada dua faktor yang membuat kasus ini terasa aneh. Pertama, yang dicuri adalah laptop milik peneliti. Kedua, pelaku langsung masuk ke ruangan tempat laptop berada.

"Mengapa pelaku tidak memilih barang yang harganya lebih mahal, mikroskop misalnya, yang harganya ratusan juta rupiah. Dan lagi, tidak ada barang lain yang rusak," lanjut Nasih.

Sebenarnya ruangan laboratorium tersebut dilengkapi cctv. Sayangnya gerak-gerik pelaku tidak berhasil terekam cctv.

Photos: Fire at University of Southampton data recovery

Mountbatten building - University of Southampton fire data recovery

4 of 9



SOURCE: Harvey Rutt

Faisal Alani

Published: 1 Mar 2011

May 30, 2021
Please bantu RT !!

Tas temenku hilang dan didalam tasnya berisi data penelitian

PANDGLANG, BERANG, BANTEN

Punten mohon bantuannya abang dan kakak semuanya
Barangkali ada saudara di sekitar Jalan AMD LINTAS TIMUR sampai lampu merah Kadu Banen (Pandeglang, Banten), MELIHAT TAS JATUH di sepanjang jalan tersebut sekitar jam 1 Sabtu, 29 Mei 2021 kemarin. Tas yang jatuh ada 3 tas Gunung dan 1 tas ransel Arei hitam-merah berisi laptop, GPS, dan data beberapa penelitian. Mohon infonya jika menemukan tas tersebut. Terima kasih

CASH REWARD
for returning my lost backpack

Bad (or no) RDM leads to data loss

80 percent of scientific data from publicly-funded research is lost within two decades

Dude, where's my data?

Most Scientific Research Data From the 1990s Is Lost Forever

Can you still publish if you have lost the original data?

I have recently completed a PhD project and written it up as a paper. My computer has since been stolen and I have lost the PhD simulations (backup was taken too). Would the paper still be able to be published?

Years of Data Lost In Fire at University

CC-BY Peter Murray-Rust, [source](#)

May 20, 2017

Kostan kemalingan, laptop ke ambil, data penelitian hilang, di backup di fd, fdnya ikut di maling. Betapa sabarnya kau kawan

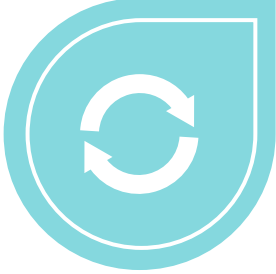
Manajemen Data: Materi Pokok



MATERI POKOK 1:

PERENCANAAN MANAJEMEN DATA

Deskripsi Data, Dokumentasi Data, Penyimpanan Data, dan Preservasi dan Berbagi Data



MATERI POKOK 2:

SIKLUS DATA DALAM MANAJEMEN DATA

Perencanaan dan Desain Data, Pengumpulan dan Dokumentasi Data, Kolaborasi dan Analisis Data, Pengelolaan, Penyimpanan dan Preservasi Data, Berbagi dan Publikasi Data, dan Penemuan, Penggunaan Kembali dan Sitasi Data



MATERI POKOK 3:

REVIU KONTEN DATA

Mereviu Konten Data untuk Preservasi dan Mereviu Konten Data untuk Analisis

Manajemen Data: Materi Pokok



MATERI POKOK 4: METADATA

Deskripsi Metadata, Ragam Metadata, dan Pemanfaatan Metadata untuk Proyek Penelitian

MATERI POKOK 5: DOKUMENTASI DATA

Deskripsi dan/atau informasi teknis dari suatu data

MATERI POKOK 6: ETIKA DAN INTEGRITAS DATA

Definisi etika data, Prinsip-prinsip etika data dalam pengelolaan dan pemanfaatan data, Regulasi terkait etika data, dan Ragam Lisensi Data

MATERI POKOK 7: SENSITIVITAS DATA

Pengklasifikasian Data, Klirens etik penelitian pada domain keilmuan, dan Anonimisasi atau pseudonimisasi

“Manajemen data bukan ‘pekerjaan tambahan’, melainkan fondasi dari analisis yang akan dilakukan di proses selanjutnya”



MP 1

PERENCANAAN

MANAJEMEN

DATA

*Deskripsi Data,
Dokumentasi Data,
Penyimpanan Data, dan
Preservasi dan Berbagi
Data*



Data Management Plan (DMP)

DMP merupakan suatu pernyataan yang bersifat formal yang disampaikan dalam rangka menggambarkan bagaimana data akan dikelola dan didokumentasi dalam suatu proyek riset, serta bagaimana data tersebut dapat terjamin untuk diakses jangka panjang sesuai prosedur preservasi data (JISC UK, n.d.)

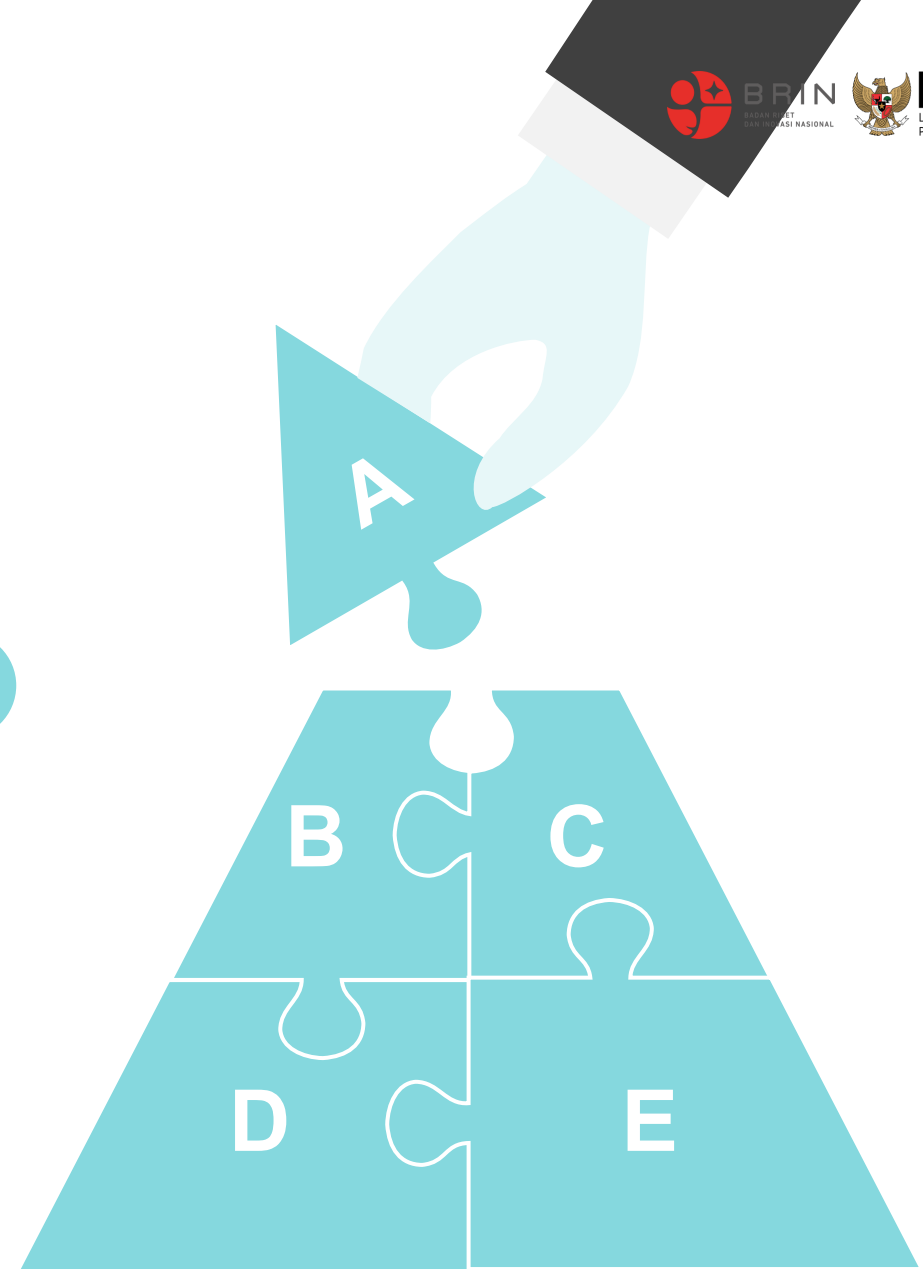
Struktur, panjang kata ataupun jumlah laman, dan format DMP akan sangat bervariasi menurut disiplin ilmu, penyandang dana, dan kebijakan terkait (Parham et al., 2016).

Dokumen DMP ini sudah mulai dibuat di awal kegiatan agar Anda dan tim dapat secara konsisten dalam memproduksi data dan mengelolanya, seperti sejak penamaan file, penentuan ekstensi file hingga penentuan hak akses, lisensi serta preservasi.

Perlu juga diperhatikan bahwa dokumen DMP ini adalah ‘dokumen hidup’ (*living document*), yang akan terus dimodifikasi sesuai dengan perkembangan kegiatan proyek yang dilakukan dan metodologi yang digunakan.

Manfaat DMP

- meningkatkan efisiensi dalam kegiatan/proyek yang dilakukan **A**
- memastikan data dan informasi yang dihasilkan terproteksi dengan baik **B**
- luaran atau hasil analisis dari kegiatan/proyek dapat dengan mudah diperiksa oleh tim **C**
- meningkatkan kesempatan untuk dapat berbagi data **D**
- memenuhi kebijakan penyandang dana **E**



Konten DMP

1. Deskripsi Data

Deskripsikan jenis data apa saja yang akan dihasilkan selama kegiatan berlangsung, termasuk metode pengumpulan data seperti melalui interview atau survei; serta bentuk data yang bisa berupa teks, gambar, audio, maupun video yang mungkin berupa format digital ataupun non-digital. Jelaskan juga perkiraan besaran data yang akan dihasilkan, baik dalam bentuk mb, gb, hingga tb atau lebih.

2. Dokumentasi Data

Deskripsikan teknis pendokumentasian data yang akan dan telah dihasilkan, termasuk konvensi penamaan file, penentuan persistent identifier, serta jika ada perubahan data/ versioning. Tahap ini dapat menyiapkan ReadMe file yang umumnya berbentuk teks (.txt) dengan mencantumkan setiap detail dokumentasi yang dilakukan.

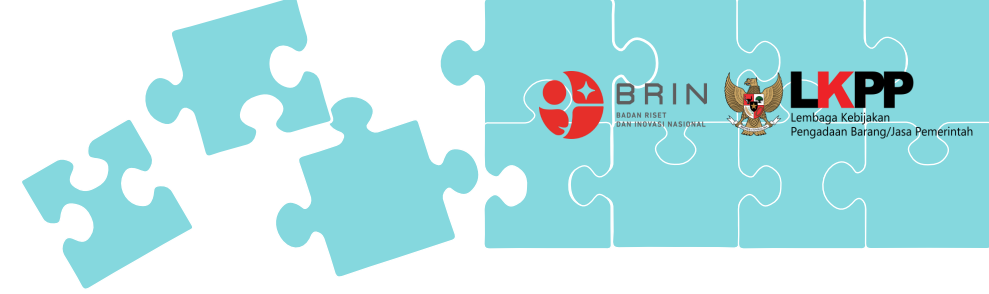
3. Penyimpanan Data

Deskripsikan penyimpanan data yang akan dan telah dihasilkan, dengan mempertimbangkan aspek keamanan, level akses serta sensitivitas data yang dimiliki. Penyimpanan data dapat memanfaatkan repositori yang terintegrasi atau memiliki interoperabilitas.

4. preservasi dan Berbagi Data

Dapat dijabarkan berdasar rentang waktu jangka pendek maupun jangka panjang. Hal ini dapat diakomodir apabila seluruh data tersimpan dalam repositori institusi karena keberlanjutan pemeliharaan yang dimiliki oleh repositori tersebut. Nyatakan juga apakah data akan dibagikan atau tidak.

Contoh DMP



1. Data Collection
 - What data will you collect or create?
 - How will the data be collected or created?
2. Documentation and Metadata
 - What documentation and metadata will accompany the data?
3. Ethics and Legal Compliance
 - How will you manage any ethical issues?
 - How will you manage copyright and Intellectual Property Rights (IPR) issues?
4. Storage and Backup
 - How will the data be stored and backed up during the research?
 - How will you manage access and security?
5. Selection and Preservation
 - Which data are of long-term value and should be retained, shared, and/or preserved?
 - What is the long-term preservation plan for the dataset?
6. Data Sharing
 - How will you share the data?
 - Are any restrictions on data sharing required?
7. Responsibilities and Resources
 - Who will be responsible for data management?
 - What resources will you require to deliver your plan?



<https://dmponline.dcc.ac.uk/>



<https://dmptool.org/>

Ilustrasi DMP di LKPP

“Analisis Keterlibatan UMKM dalam Pengadaan Barang/Jasa Pemerintah di Indonesia”

1. Data Collection

Penelitian menggunakan data paket pengadaan dari LKPP tahun 2021–2025 yang diperoleh dari master data e-katalog Inaproc. Data berupa file CSV dan XLSX dengan ukuran sekitar 100 MB. Data diproses menggunakan Python untuk menghasilkan dataset terintegrasi yang siap dianalisis.

2. Documentation and Metadata

Setiap dataset akan dilengkapi README.txt yang menjelaskan sumber data, struktur folder, definisi variabel, dan tahapan pengolahan data.

3. Ethics and Legal Compliance

Penelitian menggunakan data publik sehingga tidak melibatkan subjek manusia maupun data pribadi. Jika ditemukan data pribadi maka identitas responden akan dianonimkan sebelum analisis dilakukan. Dataset hasil penelitian akan dipublikasikan dengan lisensi Creative Commons Attribution (CC-BY).

4. Storage and Backup

Data disimpan pada server institusi dan komputer kerja peneliti. Backup otomatis dilakukan setiap minggu ke cloud storage institusi. Hanya anggota tim penelitian yang memiliki hak akses.

5. Selection and Preservation

Dataset final, data dictionary, kode analisis Python, dan laporan penelitian akan dipertahankan karena diperlukan untuk reproduksibilitas penelitian. Dataset akhir akan disimpan pada repositori institusi selama minimal 10 tahun dalam format CSV dan PDF/A untuk menjamin keterbacaan jangka panjang.

6. Data Sharing

Dataset akhir dan dokumentasi akan dipublikasikan melalui repositori institusi dan diberikan DOI agar dapat ditemukan dan disitasi. Tidak terdapat pembatasan karena seluruh data berasal dari sumber terbuka.

7. Responsibilities and Resources

Ketua peneliti bertanggung jawab terhadap kepatuhan DMP. Data manager bertanggung jawab atas dokumentasi, metadata, dan pengelolaan versi dataset. Penelitian akan membutuhkan ruang penyimpanan sebesar 50 GB, akses ke repositori institusi, perangkat lunak Python, dan dukungan teknis dari unit teknologi informasi.

MP 2

SIKLUS DATA DALAM MANAJEMEN DATA

*Perencanaan dan Desain Data,
Pengumpulan dan
Dokumentasi Data, Kolaborasi
dan Analisis Data,
Pengelolaan, Penyimpanan
dan Preservasi Data, Berbagi
dan Publikasi Data, dan
Penemuan, Penggunaan
Kembali dan Sitasi Data*

Tingkatan Proses Data



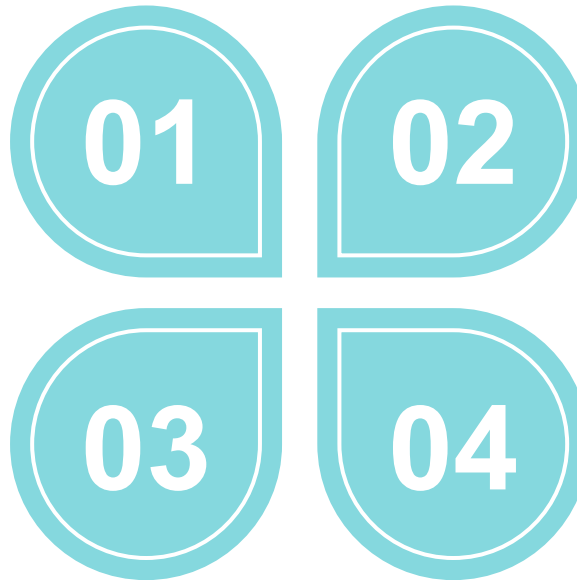
Data mentah (raw data)

Merupakan data yang dihasilkan langsung dari kegiatan pengumpulan data tanpa ada perubahan berarti. Data mentah umumnya dihasilkan pada tahap awal dari pembuatan data dari siklus hidup data (data creation).



Data hasil analisis (analyzed data)

Merupakan luaran data hasil analisis yang menunjukkan signifikansi atau temuan.



Data terolah (processed data)

Merupakan data yang telah didokumentasikan yang kemudian disiapkan sedemikian rupa sebagai bahan analisis. Pada tahapan ini, umumnya data tersebut dirapikan, dibersihkan, ataupun distandarkan agar seragam.

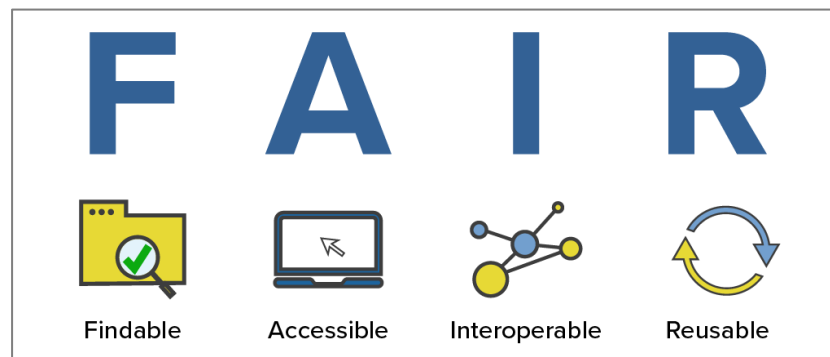


Data terpublikasi (finalized/ published data)

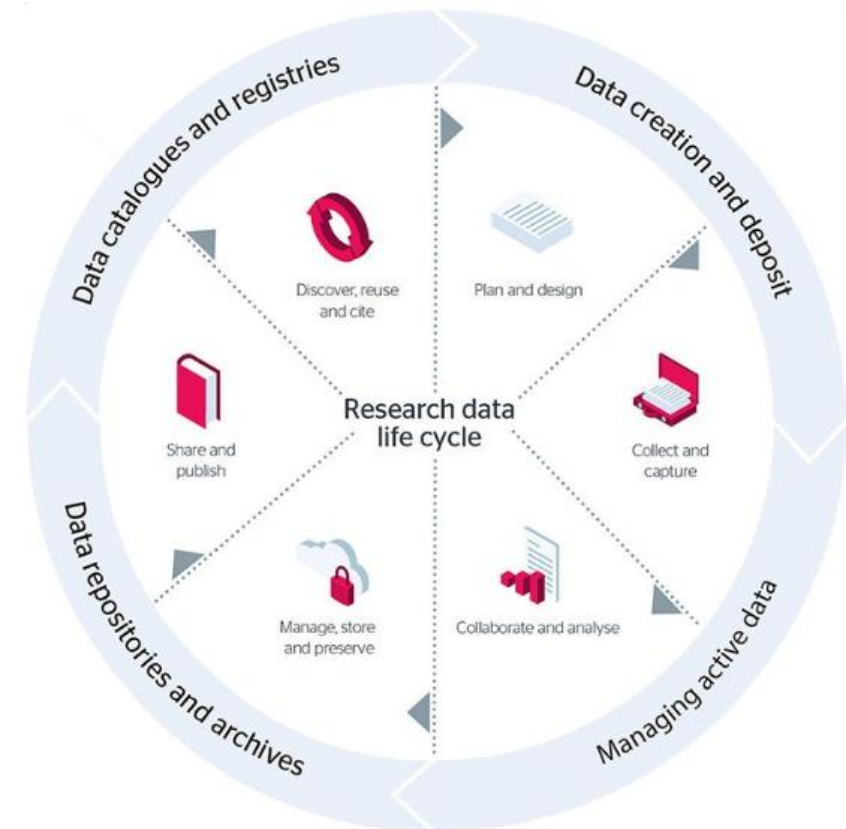
Merupakan data yang sudah siap untuk disebarluaskan kepada publik, dengan catatan sudah bebas dari data sensitif dan diberikan lisensi untuk ketentuan penggunaan kembali.

Siklus Hidup Data

- Perencanaan dan Desain Data
- Pengumpulan dan Dokumentasi Data
- Kolaborasi dan Analisis Data
- Pengelolaan, Penyimpanan dan Preservasi Data
- Berbagi dan Publikasi Data
- Penemuan, Penggunaan Kembali dan Sitasi Data

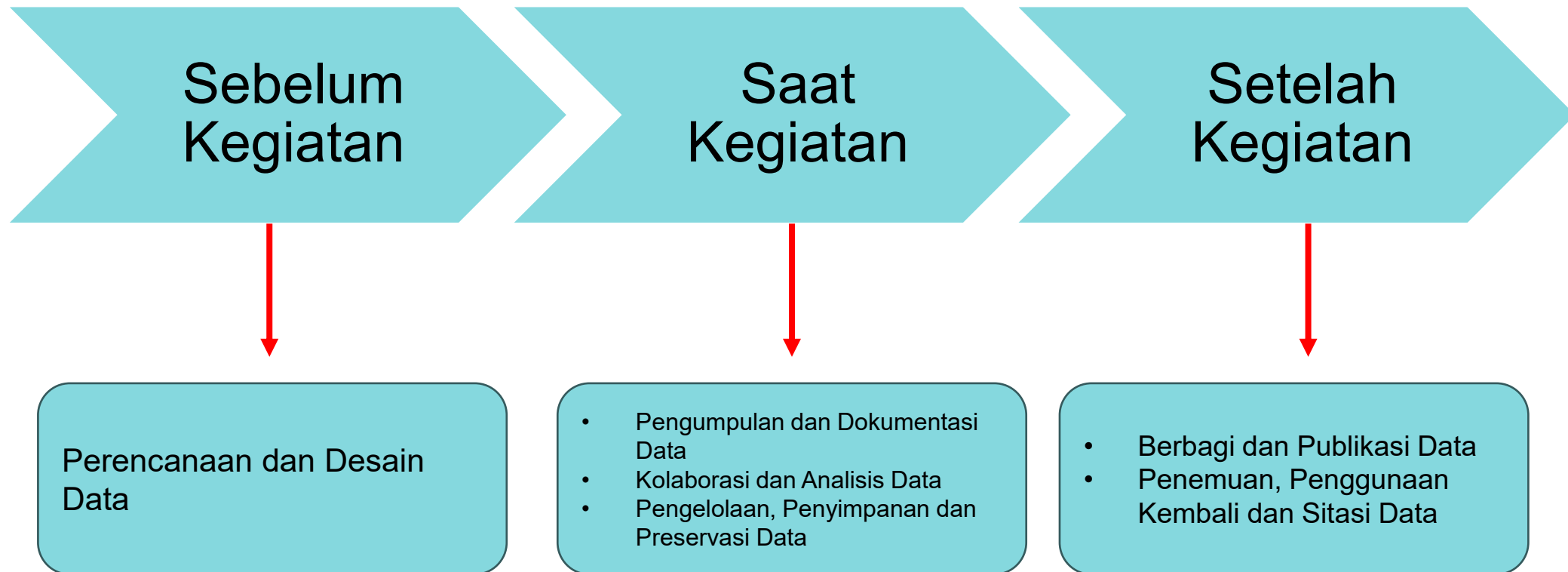


<https://www.nlm.nih.gov/oet/ed/cde/tutorial/02-200.html>



Willoughby, C., & Frey, J. G. (2022). Data management matters. *Digital Discovery*, 1(3), 183–194.
<https://doi.org/10.1039/d1dd00046b>

Implementasi Teknis

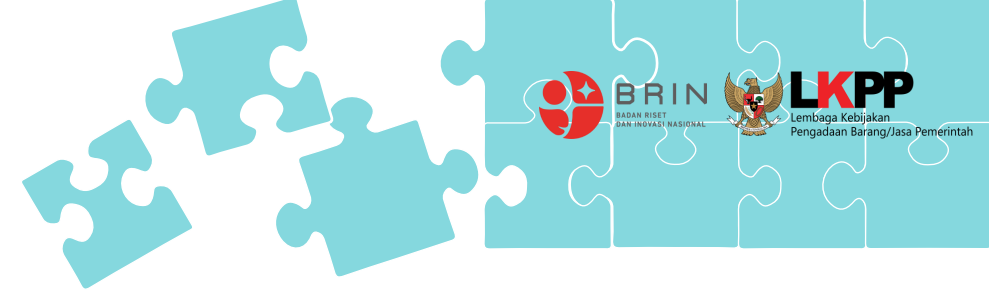


MP 3

REVIU KONTEN DATA

*Mereviu Konten Data
untuk preservasi dan
Mereviu Konten Data
untuk Analisis*

Reviu Konten Data



1. Proses mengevaluasi kualitas & kesiapan data untuk proses selanjutnya
2. Dilakukan sebelum:
 - Penyimpanan jangka panjang (preservasi)
 - Analisis data
3. Fokus:
 - Kualitas
 - Konsistensi
 - Kelengkapan
 - Keterbacaan
 - Dokumentasi pendukung
 - Standarisasi format

Reviu Konten Data

Untuk Preservasi

- Bertujuan untuk menjaga keamanan konten (data) dan konteks (metadata) untuk digunakan kembali di masa mendatang.
- Diperlukan informasi dari karakteristik format file untuk memastikan akses jangka panjang → terenkripsi/tidak; penggunaan format yg direkomendasikan/tidak

Untuk Analisis

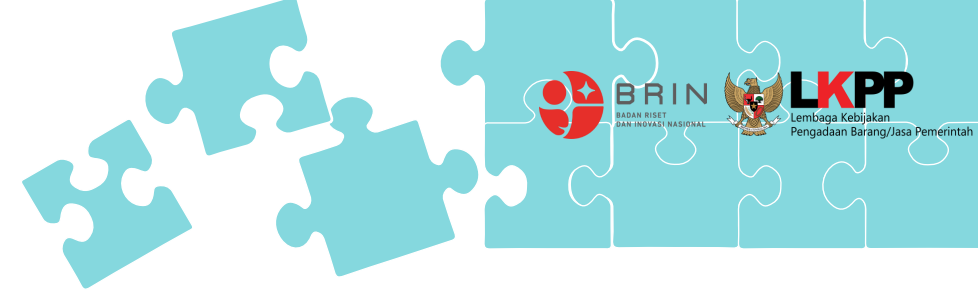
Bertujuan untuk mendapatkan data yang berkualitas agar menghasilkan analisis yang akurat

Disebut juga dengan *data preprocessing*, yang merupakan tahapan prapemrosesan data sebelum diolah menggunakan metode tertentu untuk dianalisis

Biasanya meliputi proses data integration, data cleaning, data reduction, dan data transformation



Contoh Format Data



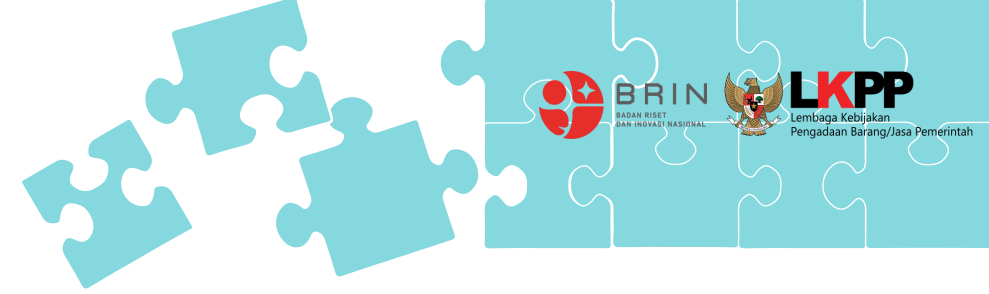
Data Type	Original Data Format	Preservation Friendly Formats (Open Standard, Uncompressed)
Text	Hand-written, docx, wpd, odt, rtf, txt, html, xml, pdf	xml, PDF/A, txt
Tabular Simple (minimal metadata)	csv, tsv, pipe-delimited, xls(x), ods, dif, xps	csv
Tabular Extensive (variable and value labels, missing data defined)	sav (SPSS), sas7bdat or xpt (SAS), dta (STATA)	csv, txt with setup file or associated script (r or m)
Database	mdb, dbf, sql, sqlite, db, db3, xml	xml, sqlite
Visual	static: pdf, jpeg, tiff, png, gif, bmp, moving: mpeg, mov, avi, mxf	PDF/A, tiff, JPEG2000 MPEG-4
Audio	wav(e), mp3, mp2, aiff, wma, aac, dct, flac, ogg,	wave, aiff

MP 4

METADATA

*Deskripsi Metadata,
Ragam Metadata, dan
Pemanfaatan Metadata
untuk Proyek Penelitian*

Definisi Metadata



NISO (2008) → informasi terstruktur yang terkait dengan objek tertentu dengan tujuan untuk penemuan, deskripsi, penggunaan, pengelolaan, dan pelestarian.

Farnel & Shiri (2014) → informasi terstruktur yang menyediakan konteks untuk semua jenis objek informasi, termasuk data

Metadata → Data tentang data

Jenis Metadata

Joudrey et al., (2018) menyebutkan bahwa metadata terdiri dari tiga jenis, yaitu metadata deskriptif, administratif, dan struktural.

Administratif

Berfokus pada hak penggunaan dan pengelolaan objek digital.
Contoh: Lisensi, alasan embargo, pelacakan pengguna, dsb.

Metadata

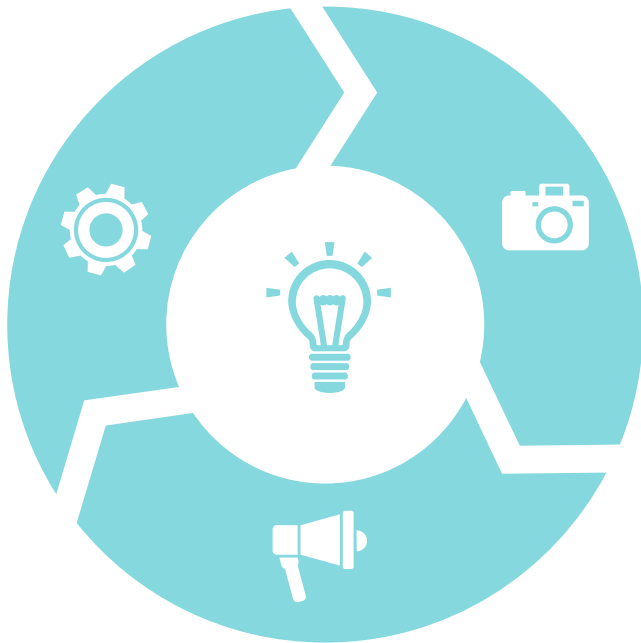
Deskriptif

Metadata minimum yang diperlukan untuk menemukan objek digital.
Contoh: Penulis, judul, abstrak/deskripsi, tanggal, dsb.

Struktural

Merekam hubungan antara objek digital individu/data yang membentuk satu unit.
Contoh: Link/relasi ke objek digital terkait, format file, nama file, dsb.

Contoh Ragam Metadata



Dublin Core: 15 elemen inti → contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title, type



DataCite: memiliki 3 tingkatan → a) **Wajib:** Identifier, Creator, Title, Publisher, PublicationYear, dan ResourceType; b) **Direkomendasikan:** Subject, Contributor, Date, RelatedIdentifier, Description, dan GeoLocation; c) **Opsional:** Language, AlternateIdentifier, Size, Format, Version, Rights, FundingReference, dan RelatedItem.



Darwin Core: Domain specific metadata (Ilmu Hayati) → Classes, Record level, Occurrence, Organism, Location, Taxon, dsb.

Catatan

Pemilihan dan penggunaan metadata sangat bergantung pada kebutuhan dan jenis data yang dimiliki organisasi

Prinsip FAIR Data

Findable

- Memiliki pengidentifikasi (PID) yang unik dan tetap secara global (misalnya, DOI)
- Memiliki metadata yang kaya dan dapat dibaca manusia dan mesin (misalnya, judul, pencipta, abstrak, kata kunci, dsb.)
- Terdaftar/terindeks dalam sumber yang dapat dicari (biasanya repositori/katalog) sehingga dapat ditemukan oleh manusia dan mesin.

Interoperable

- Format file terbuka dan terdokumentasi
- Skema/standar metadata yang disepakati komunitas
- Penggunaan kosakata terkontrol, tesaurus, dan ontologi yang konsisten
- Tautan yang terkait antarentitas melalui PID (misalnya, tautan DOI dataset ↔ DOI artikel ↔ ORCID iD)



Accessible

- Dapat diakses menggunakan protokol terbuka dan terstandarisasi (Misal API)
 - Mendukung otorisasi jika diperlukan (untuk data yg sensitif, dirahasiakan, atau dibatasi)
- Metadata tetap dapat diakses publik meskipun data dibatasi, tertutup, atau tidak tersedia.

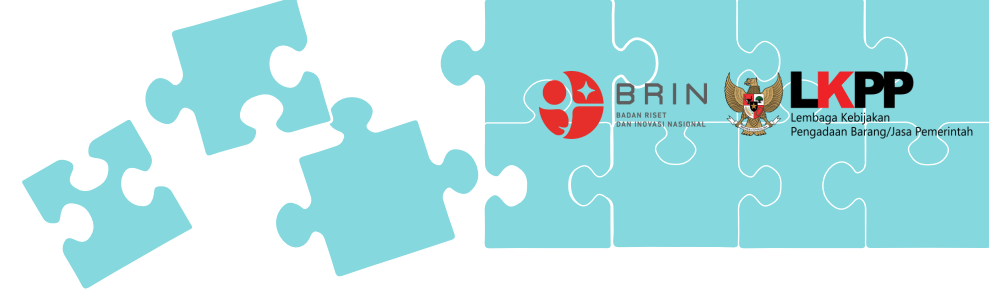
Reusable

- Terdokumentasi dengan baik sehingga di masa depan dapat digunakan dan ditafsirkan dengan benar
- Memiliki ketentuan penggunaan kembali yang jelas melalui lisensi, yang tercatat dalam metadata.
 - Mencakup asal usul dan konteks (bagaimana data dihasilkan/diproses, versi, dsb.)

MP 5 DOKUMENTASI DATA

*Deskripsi dan/atau
informasi teknis dari
suatu data*

Dokumentasi Data



- Berfungsi untuk memperkenalkan data Anda, memberikan deskripsi yang detail tentang atribut utamanya, dan memberi konteks agar orang lain memahami data Anda.
- Dalam dokumentasi, perlu menyertakan informasi yang menjelaskan tentang metode apa yang dilakukan (pengambilan, pemrosesan, analisis, dsb.), bagaimana hal itu dilakukan, mengapa hal itu dilakukan, kapan dilakukan, di mana dilakukan, dan siapa yang melakukan pekerjaan tersebut.
- Dokumentasi data dapat berupa kamus data, file readme, dan metadata.

Kamus Data / Data Dictionary

Kamus data penting untuk membuat penelitian Anda lebih mudah direproduksi karena memungkinkan orang lain untuk memahami data Anda. Tujuan kamus data adalah untuk menjelaskan arti sebenarnya dari semua nama variabel dan nilai dalam spreadsheet Anda.

data/

OSF Storage (United States)

0.1 datadictionary.csv

0.2 raw.pdf

0.3 clean.csv

Tags

Sheet_1				
Show rows with cells including:				
Variable	Variable name	Mesaurement unit	Allowed values	Description
Participant ID number	ID	Numeric	001-999	ID number assigned to participant in sequential order
Group number	GROUP	Numeric	1-30	Group assigned to participant based on ID number
Age in years	AGE	Numeric	18.0-65.0	Age of participant in years
Date of birth	DOB	mm/dd/yyyy	1-12/1-31/1951-1998	Participant's date of birth
Gender	SEX	Numeric	1 = male 2 = female	Participant's gender
Date of survey	SURVEY	mm/dd/yyyy	01/01/2015 – 01/01/2016	When the participant completed the survey
Self-reported consumer spending	SPEND	Numeric	0-100,000,000	Self-reported average yearly expenditure
Market sentiment	SENTIMENT	Numeric	1 = negative 2 = neutral 3 = positive	Sentiment towards US domestic economy
Actual GDP growth	GDP	Numeric	-5.0-5.0	Average US yearly GDP growth

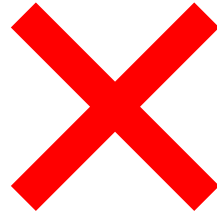
File ReadMe

- Format dokumentasi data yang umum digunakan
- Berisi informasi penting tentang file data, termasuk mengenai informasi kutipan, struktur organisasi file, informasi pengumpulan data, software yang digunakan dan versinya, informasi lisensi, dsb.
- Seringkali dibuat dalam format file .txt

Informasi Umum	• Judul dataset • Nama dan Informasi Kontak (misal kontributor, narahubung/contact person) • Tanggal/Rentang tanggal pengumpulan data • Lokasi geografis pengumpulan data • Kata kunci • Bahasa • Informasi pendanaan
Gambaran Umum dari Data dan File	• Deskripsi struktur file dan hubungan antara file data • Deskripsi singkat dari setiap file data dan hubungannya dengan konten (misal tabel, gambar) dari publikasi terkait • Tanggal file dibuat/diperbarui (jika ada) • Format file • Informasi khusus untuk file data tertentu
Informasi tentang Akses dan Berbagi Data	• Informasi Lisensi/Pembatasan • Publikasi/dataset terkait (URL) • Pengutipan yang direkomendasikan (misal APA, MLA, IEEE, dsb.)
Informasi Metodologi	• Metode pengumpulan/pembuatan data • Langkah-langkah pengolahan data • Informasi khusus instrumen apa saja yang diperlukan untuk memahami atau menginterpretasikan data, misalnya, detail sistem operasi/perangkat lunak tertentu yang diperlukan untuk memanfaatkan data

Struktur Folder & File Naming

Umkm.csv
Umkmfix.csv
Umkmfix2.csv
Umkmfix3.csv
Umkmfixbanget.csv



00_Admin
01_Raw_Data
02_Processed_Data
03_Analyzed_Data
04_Published_Data

20260625_raw_umkm.csv
20260625_processed_umkm.csv
20260625_analyzed_umkm.csv
20260625_published_umkm.csv
raw_umkm_v01.csv
processed_umkm_v01.csv
analyzed_umkm_v01.csv
published_umkm_v01.csv



01_Raw_Data
↳ 20260625_raw_inaproc.csv

02_Processed_Data
↳ 20260725_clean_inaproc.csv

03_Analyzed_Data
↳ 20260815_graph_inaproc.png

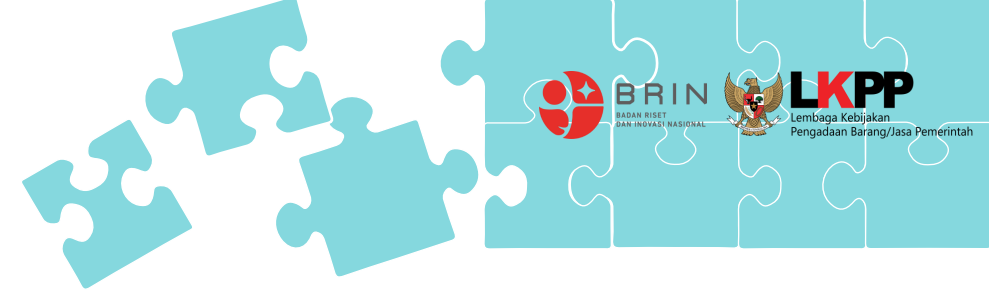
04_Published_Data
↳ 20261221_pub_inaproc.xls

MP 6

ETIKA DAN INTEGRITAS DATA

*Definisi etika data,
Prinsip-prinsip etika data
dalam pengelolaan dan
pemanfaatan data,
Regulasi terkait etika
data, dan Ragam
Lisensi Data*

Etika Data

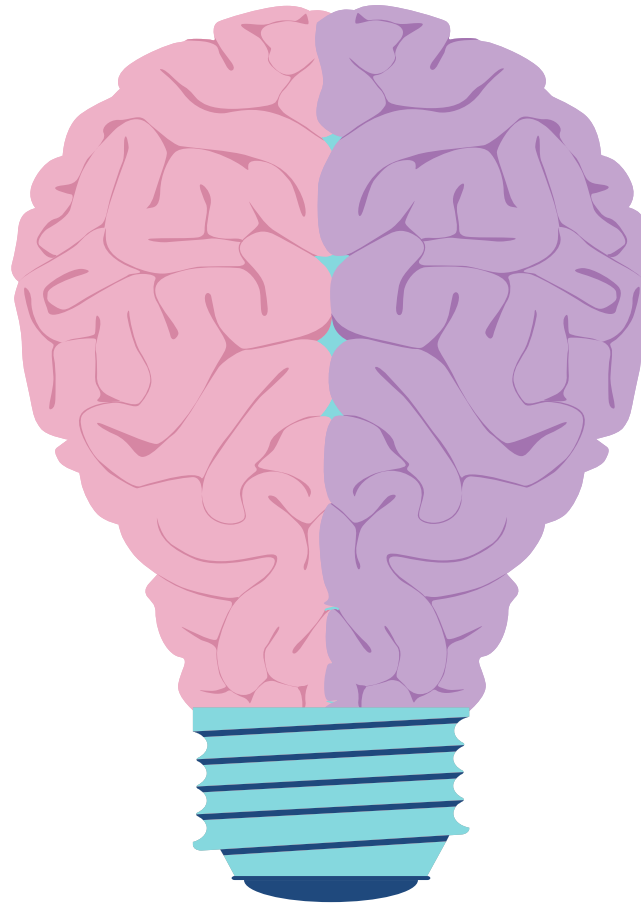


- Etika data → segala hal tentang penggunaan data yang bertanggung jawab dan berkelanjutan.
- Hasselbalch (2021) menyebutkan etika data harus menjadi lebih dari sekadar kewajiban moral, seperangkat aturan yang sudah terstruktur, namun juga harus 'manusiawi'. Hal ini berarti bahwa etika data harus berwujud budaya, menjadi proses budaya, dihayati dan dipraktikkan sebagai cara pandang di dunia.
- Prinsip etika data → mengutamakan manusia, pengawasan data individu, transparansi, akuntabel/dapat dipertanggungjawabkan, dan kesetaraan (Tranberg et al., 2018)

Regulasi terkait Etika Data

Undang-undang Nomor 27 tahun 2022 tentang Perlindungan Data Pribadi

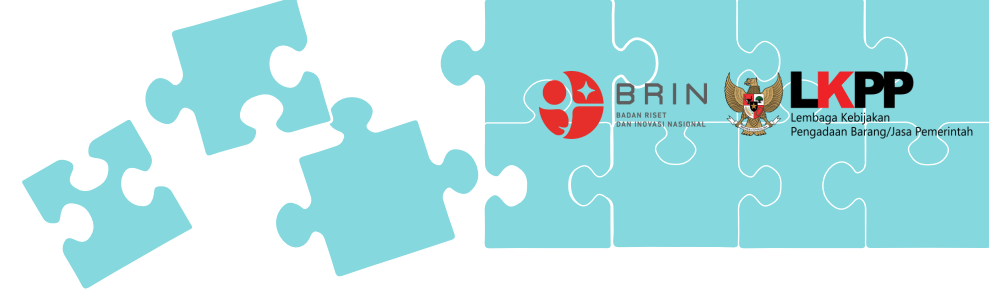
- pasal 16 ayat 2 menyatakan pemrosesan data pribadi setidaknya mengikuti 7 prinsip
- pasal 47 menambahkan bahwa pengendali data pribadi harus dapat mempertanggungjawabkan dalam setiap pemrosesan data dalam hal memenuhi kewajiban pelaksanaan prinsip perlindungan data pribadi.



General Data Protection Regulation (GDPR)

- berkaitan dengan perlindungan orang perseorangan yang berhubungan dengan pemrosesan data pribadi
- Regulasi ini melindungi hak dasar dan kebebasan individu dan khususnya hak mereka atas perlindungan data pribadi

Lisensi Data



Lisensi data adalah pengaturan hukum/perjanjian antara pemilik/pembuat data dan pengguna atau tempat penyimpanan data untuk menentukan apa saja yang dapat dilakukan pengguna dengan data tersebut (Deutz, et al., 2020)

Contoh penyedia lisensi yaitu Open Data Commons (Database) dan Creative Commons (Objek Digital, termasuk data)

CREATIVE COMMONS LICENSES		COPY & PUBLISH	ATTRIBUTION REQUIRED	COMMERCIAL USE	MODIFY & ADAPT	CHANGE LICENSE
	PUBLIC DOMAIN	✓	✗	✓	✓	✓
	CC BY	✓	✓	✓	✓	✓
	CC BY-SA	✓	✓	✓	✓	✗
	CC BY-ND	✓	✓	✓	✗	✓
	CC BY-NC	✓	✓	✗	✓	✓
	CC BY-NC-SA	✓	✓	✗	✓	✗
	CC BY-NC-ND	✓	✓	✗	✗	✓

 You can redistribute
(copy, publish, display,
communicate, etc.)
  You have to attribute
the original work
  You can use the work
commercially
  You can modify and
adapt the original work
  You can choose license
type for your adaptations
of the work.

<https://www.wur.nl/en/article/What-are-Creative-Commons-licenses.htm>

<https://creativecommons.org/chooser/>



MP 7

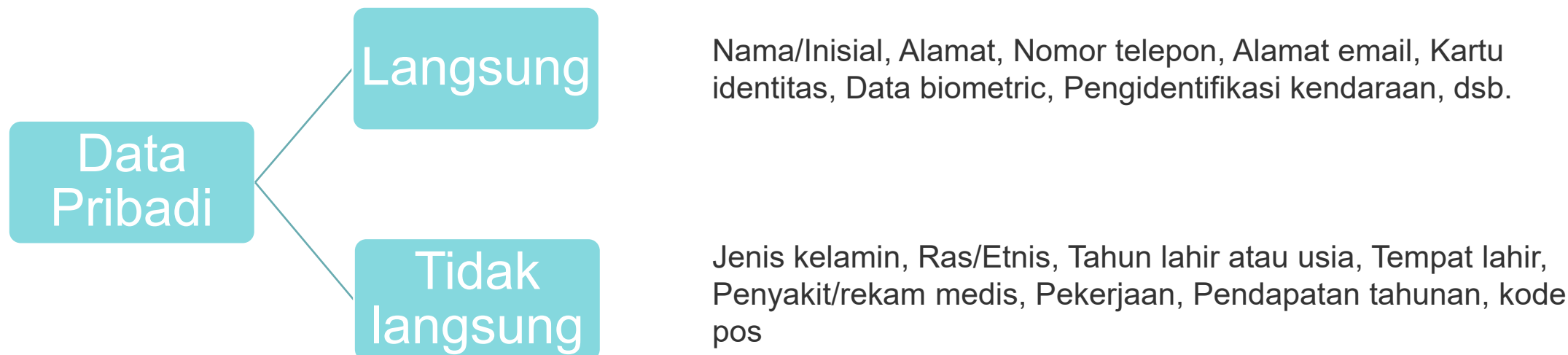
SENSITIVITAS

DATA

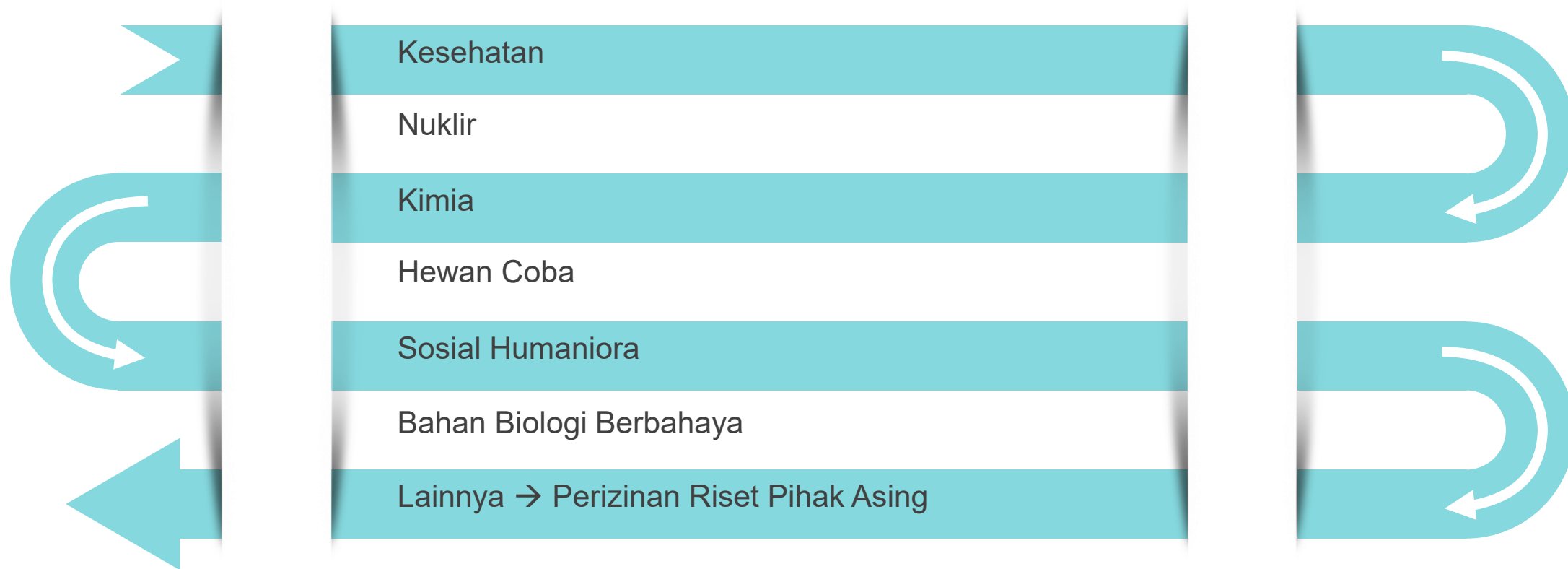
*Pengklasifikasian Data,
Klirens etik penelitian
pada domain keilmuan,
dan Anonimisasi atau
pseudonimisasi*

Data Sensitif

Australian Research Data Commons (ARDC) → "data yang dapat digunakan untuk mengidentifikasi individu, spesies, objek, proses, atau lokasi yang berpotensi menimbulkan risiko diskriminasi, kerugian, atau dampak lainnya yang tidak diinginkan".



Klirens Etik



Review oleh Komisi Etik dari Direktorat Tata Kelola Perijinan Riset dan Inovasi, dan Otoritas Ilmiah BRIN

<https://klirensetik.brin.go.id/prosedur>



Kategori Hijau → Tidak berisiko / rendah untuk Nuklir
Kategori Kuning → Risiko sedang
Kategori Merah → Risiko tinggi

Penanganan Data Sensitif

Anonimisasi

Hapus atau ubah informasi identitas pribadi (misalnya, nama, alamat) untuk mencegah teridentifikasinya individu.

Persetujuan Tertulis / Consent Form

Mendapatkan persetujuan secara tegas dari individu sebelum mengumpulkan, menggunakan, atau membagikan data mereka, serta memastikan bahwa mereka telah sepenuhnya memahami bagaimana data mereka akan digunakan.

Pseudonimisasi

Ganti identitas pribadi dengan nama samaran atau kode yang hanya dapat dihubungkan dengan data aslinya melalui kunci terpisah yang aman.

Generalisasi Informasi

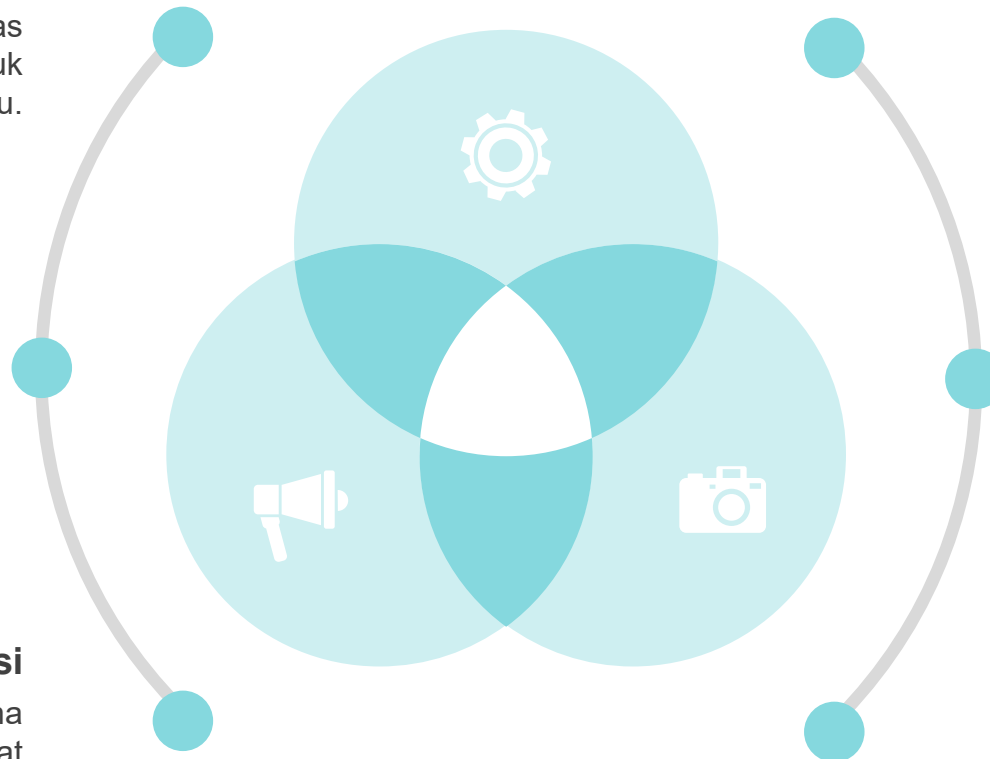
Gabungkan data atau gunakan kategori yang lebih luas untuk mengurangi tingkat spesifik informasi (misalnya, gunakan rentang usia dibanding usia pasti).

Minimalisir Data

Kumpulkanlah hanya data yang diperlukan untuk tujuan penelitian, dan hindari pengumpulan informasi yang berlebihan atau tidak relevan

Perjanjian Kerahasiaan

Mewajibkan seluruh personel yang menangani data sensitif untuk menandatangani perjanjian kerahasiaan, yang berisi komitmen untuk melindungi privasi data tersebut.





Silakan akses <https://kahoot.it/>

Output Praktis Pelatihan

- ✓ Membuat perencanaan manajemen data (DMP)
- ✓ Membuat struktur folder dan file naming
- ✓ Membuat pendokumentasian data (readme/data dictionary/metadata)
- ✓ Menentukan lisensi data



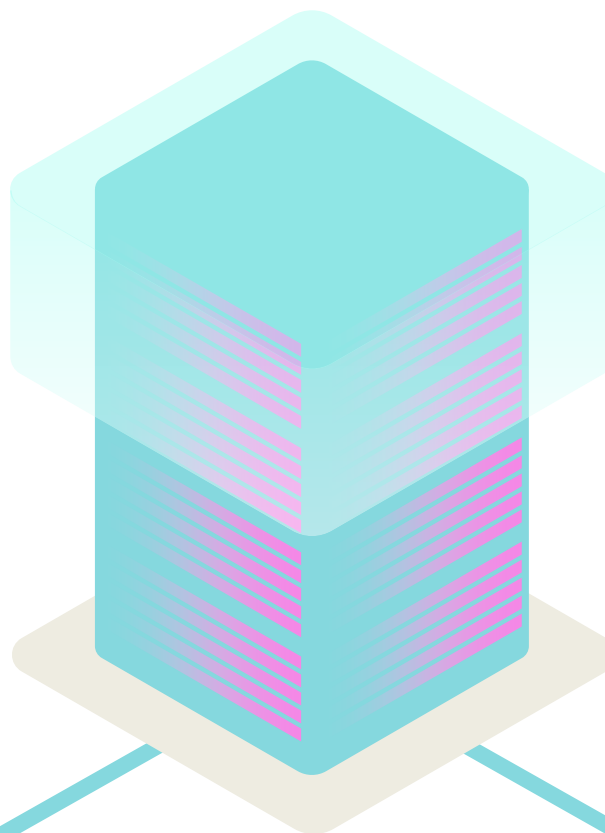
STUDI KASUS

*Manajemen
Data dalam
Repositori
Ilmiah Nasional*

Repositori Ilmiah Nasional



Sarana untuk
menyimpan,
melestarikan, mengutip,
menganalisis dan
berbagi data penelitian



Menggunakan platform
Dataverse yang
dikembangkan oleh
IQSS-Harvard University
untuk pengelolaan
repositori data

Komponen di dalam RIN

Organisasi / Kelompok Riset
/ Laboratorium / Program
Riset / Departemen /
Fakultas

Judul/nama kegiatan
riset

☒  **Dataverses (35)**

Folder, Directory

☒  **Datasets (3,522)**

Research Project, Publications

☐  **Files (8,665)**

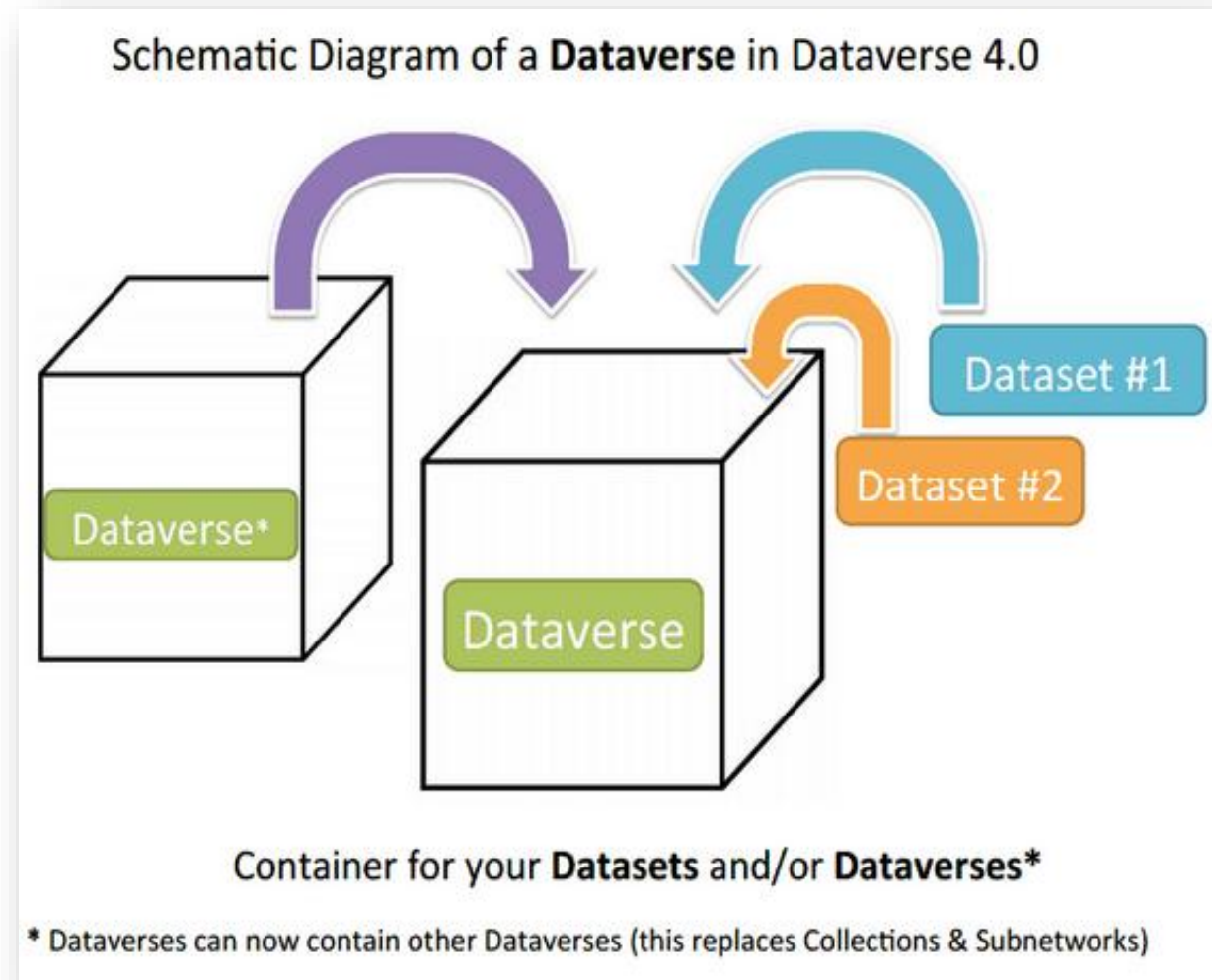
*Documents (text, audio, video,
images, spatial, streaming)*

File data



Dataverse

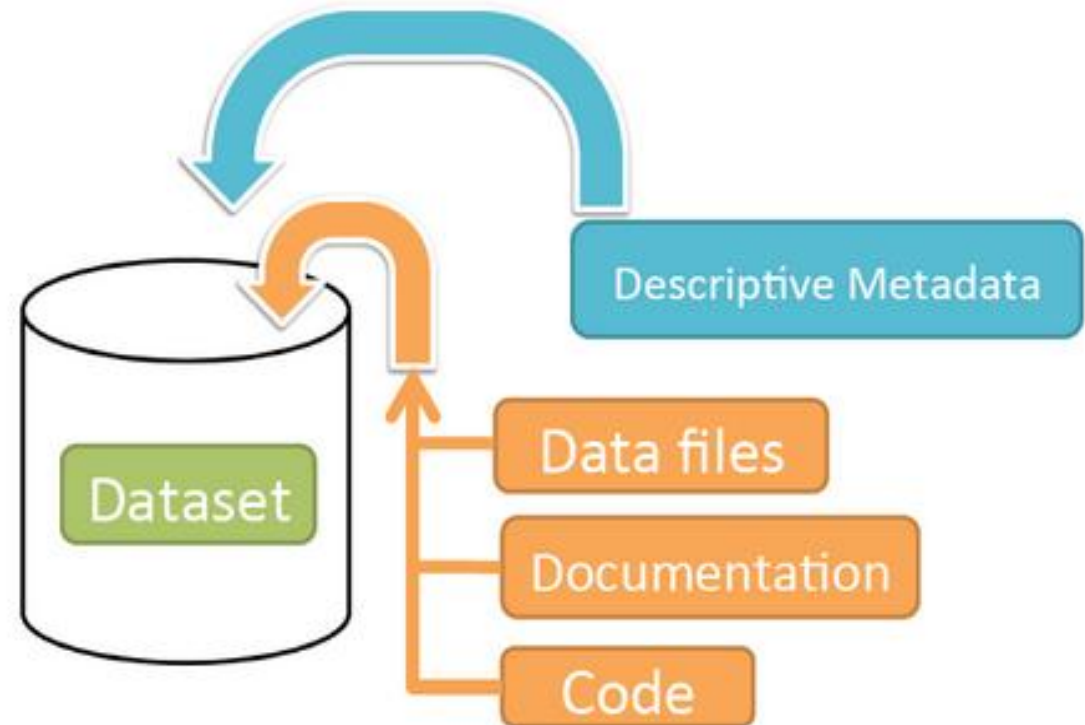
Dataverse adalah wadah utama untuk kumpulan dari beberapa dataset dan/atau kumpulan dataverse yang dapat diatur sesuai dengan kebutuhan dari masing-masing peneliti, departemen, jurnal dan organisasi.



Dataset

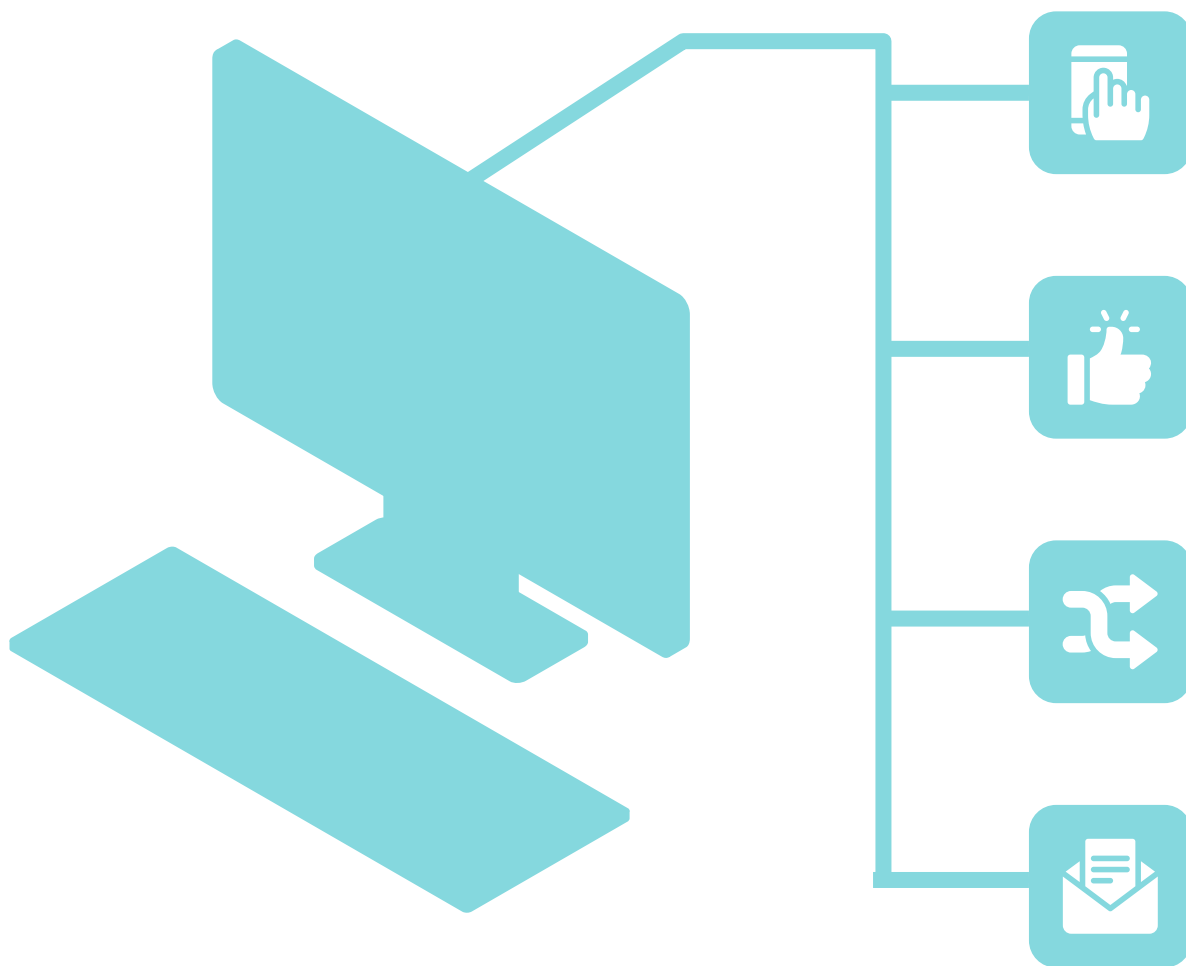
Dataset adalah wadah untuk menyimpan dan mengelola data, dokumentasi, kode, dan metadata yang mendeskripsikan dataset.

Schematic Diagram of a **Dataset** in Dataverse 4.0



Container for your data, documentation, and code.

Files



Jenis data

Semua jenis data

Jenis format

Semua jenis format files

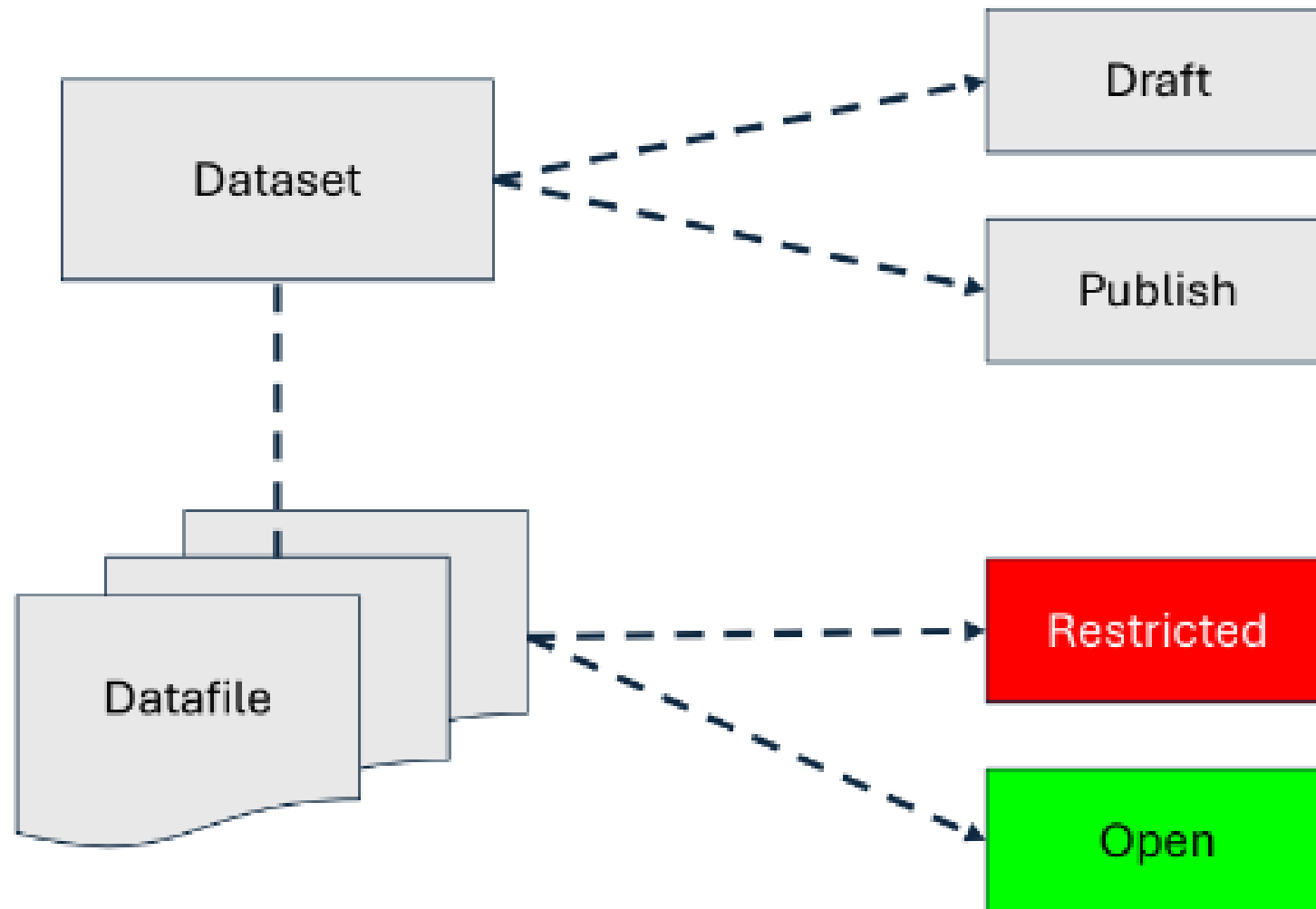
Kuota dan Size

Tidak ada batasan kuota files; up to 60gb/files

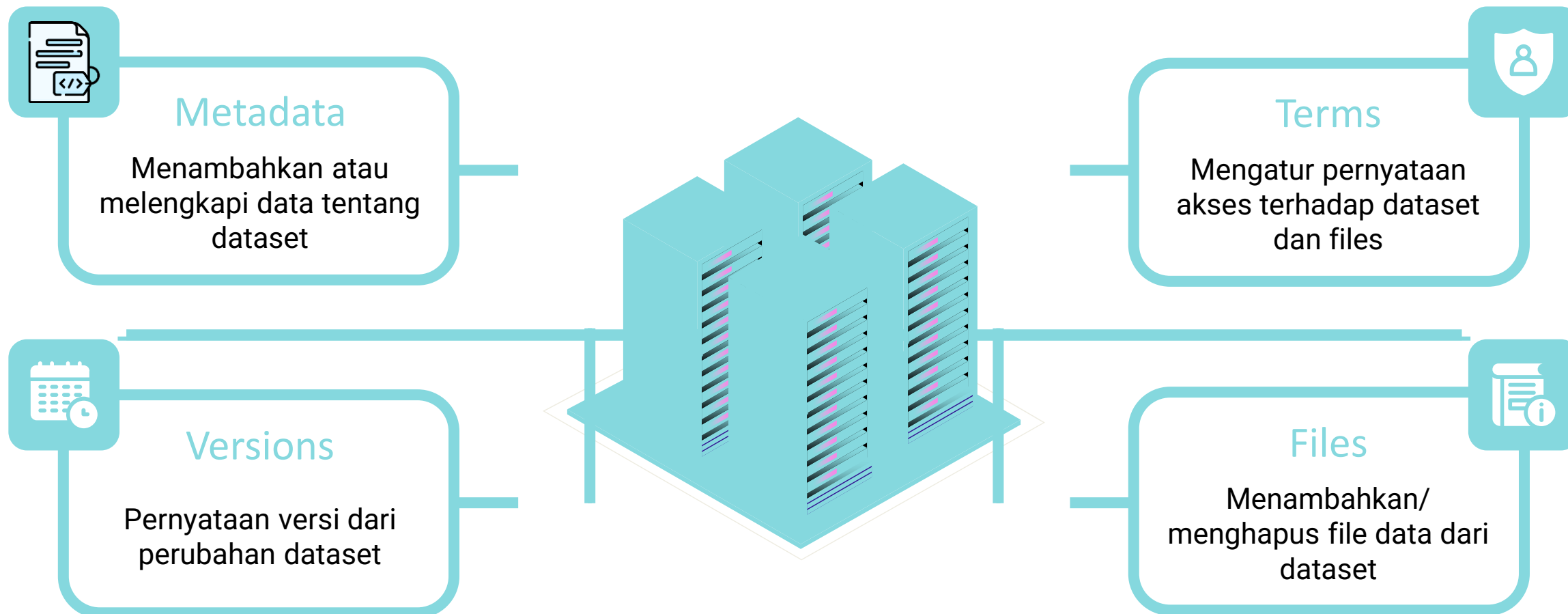
Akses

Pemilik data yang mengatur akses terhadap file data

Akses di RIN



Manajemen Dataset



Terima Kasih

“Jumlah data akan terus bertambah, tetapi tanpa pengelolaan yang baik, nilainya akan hilang. Tugas kita bukan hanya menyimpan data, tetapi memastikan data tetap bermakna.”

Seno Yudhanto

seno.yudhanto@brin.go.id

Direktorat Repositori, Multimedia, dan Penerbitan Ilmiah
DFRI – BRIN