

Modul MANAJEMEN DATA

Pelatihan Jabatan Fungsional
Analisis Data Ilmiah



Fungsi Layanan Pengembangan Kompetensi Kedinasan
Direktorat Pengembangan Kompetensi
Deputi Sumber Daya Manusia
2022

Penanggung Jawab:

1. Edy Giri Racman Putra, Ph.D.
2. Nining Setyowati Dwi Andayani, S.E., M.M.
3. Raden Arthur Ario Lelono, Ph.D.
4. Alpha Fadila Juliana Rahman, S.Pd., M.Pd.

Tim Penyusun Modul:

1. Wagiyah, M.Sc.
2. Seno Yudhanto S.Hum
3. Madiareni Sulaiman S.Hum., M.Sc.
4. Naily Kamaliah, M.Si.

Diterbitkan oleh:

Direktorat Pengembangan Kompetensi - BRIN
Gedung B.J. Habibie, Jalan M.H. Thamrin Nomor 8,
Jakarta Pusat 10340

Diterbitkan pertama kali tahun 2022

KATA PENGANTAR

Pengembangan kompetensi Aparatur Sipil Negara (ASN) khususnya Pegawai Negeri Sipil (PNS) dalam mengembangkan karier jabatan fungsionalnya menjadi suatu tuntutan sehingga mampu menjalankan tugas dan fungsinya dengan sebaik – baiknya sehingga mampu memberikan kontribusi yang nyata terhadap bangsa dan Negara Kesatuan Republik Indonesia (NKRI).

Badan Riset dan Inovasi Nasional (BRIN) sebagai Instansi pembina 11 (sebelas) jabatan fungsional yang meliputi peneliti, perekayasa, pengembangan teknologi nuklir, analis perkebunrayaan, analis pemanfaatan iptek, analis data ilmiah, kurator koleksi hayati, penata penerbitan ilmiah, teknisi perkebunrayaan, teknisi penelitian dan perekayasaan, dan pranata nuklir. BRIN melalui kedeputian Sumber Daya Manusia Ilmu Pengetahuan dan Teknologi (SDMI) bertanggung jawab dalam penyelenggaraan pengembangan kompetensi 11 (sebelas) jabatan fungsional tersebut.

Kedeputian SDMI – BRIN melalui Direktorat Pengembangan Kompetensi sebagai penyelenggara pengembangan kompetensi jabatan fungsional bertanggung jawab dalam menyiapkan kebutuhan tersebut baik berupa pengelolaan pembelajaran, fasilitator, modul, bahan ajar dan sebagainya, yang merujuk kepada regulasi peraturan BRIN nomor 28 tahun 2022 tentang pedoman pelatihan pembentukan jabatan fungsional peneliti, peraturan BRIN nomor 29 tahun 2022 tentang pedoman pelatihan jabatan fungsional kurator koleksi hayati, peraturan BRIN nomor 30 tahun 2022 tentang pedoman pelatihan jabatan fungsional analis pemanfaatan iptek, peraturan BRIN nomor 31 tahun 2022 tentang pedoman pelatihan jabatan fungsional analis data ilmiah, peraturan BRIN nomor 32 tahun 2022 tentang pedoman pelatihan jabatan fungsional analis perkebunrayaan, peraturan BRIN nomor 33 tahun 2022 tentang pedoman pelatihan jabatan fungsional teknisi perkebunrayaan, dan peraturan BRIN nomor 34 tahun 2022 tentang pedoman pelatihan jabatan fungsional penata penerbitan ilmiah.

Kami mengucapkan syukur ke hadirat Tuhan Yang Maha Esa, atas berkat rahmat-Nya, modul pelatihan jabatan fungsional analis data ilmiah “MANAJEMEN DATA”

dapat diselesaikan tepat waktu. Modul ini digunakan dalam pelatihan jabatan fungsional yang dibina oleh BRIN yang diselenggarakan oleh Kedeputian SDMI - BRIN melalui Direktorat Pengembangan Kompetensi. Kami berharap modul ini dapat memberikan manfaat dan kontribusi dalam meningkatkan dan mengembangkan kompetensi jabatan fungsional yang dibina BRIN.

Jakarta, Desember 2022

Plt. Deputi Sumber Daya Manusia
Ilmu Pengetahuan dan Teknologi
Badan Riset dan Inovasi Nasional

(Tanda tangan)

Edy Giri Rachman Putra, Ph.D.

DAFTAR ISI

KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR GAMBAR	vii
DAFTAR TABEL	viii
PENDAHULUAN	1
A. Deskripsi Singkat	1
B. Alokasi Waktu	1
C. Tujuan Pembelajaran	1
D. Materi Pokok.....	2
MATERI POKOK 1: PERENCANAAN MANAJEMEN DATA	4
A. Deskripsi Data.....	7
B. Dokumentasi Data	7
C. Penyimpanan Data	7
D. Preservasi dan Berbagi Data	8
E. Rangkuman.....	8
F. Evaluasi	9
MATERI POKOK 2: SIKLUS DATA DALAM MANAJEMEN DATA	10
A. Perencanaan dan Desain Data	11
B. Pengumpulan dan Dokumentasi Data	13
C. Kolaborasi dan Analisis Data	13
D. Pengelolaan, Penyimpanan dan Preservasi Data	13
E. Berbagi dan Publikasi Data.....	14
F. Penemuan, Penggunaan Kembali dan Sitasi Data	15
G. Rangkuman.....	16
H. Evaluasi	16
MATERI POKOK 3: REVIU KONTEN DATA	17
A. Mereviu Konten Data untuk Preservasi.....	17
B. Mereviu Konten Data untuk Analisis	20

C. Rangkuman.....	21
D. Evaluasi	21
MATERI POKOK 4: PENGENALAN METADATA	22
A. Deskripsi Metadata	22
B. Ragam Metadata	26
C. Pemanfaatan Metadata untuk Proyek Penelitian	28
D. Rangkuman.....	32
E. Evaluasi	32
MATERI POKOK 5: DOKUMENTASI DATA.....	33
A. Deskripsi dan atau informasi teknis dari suatu data.....	33
B. Rangkuman.....	36
C. Evaluasi	36
MATERI POKOK 6: ETIKA DAN INTEGRITAS DATA	37
A. Definisi etika data.....	37
B. Prinsip-prinsip etika data dalam pengelolaan dan pemanfaatan data	39
C. Regulasi terkait etika data.....	39
D. Ragam Lisensi Data.....	40
E. Rangkuman.....	47
F. Evaluasi	47
MATERI POKOK 7: SENSITIVITAS DATA	48
A. Pengklasifikasian data berdasarkan tingkat sensitivitasnya	48
B. Klirens etik penelitian pada domain keilmuan	50
C. Anonimisasi atau pseudonimisasi	62
D. Rangkuman.....	64
E. Evaluasi	64
DAFTAR PUSTAKA	65

DAFTAR GAMBAR

Gambar 1. Siklus data riset menurut JISC UK.....	10
---	----

DAFTAR TABEL

Tabel 1. Jenis Metadata.....	24
Tabel 2. Elemen-elemen inti dari DCMES.....	26
Tabel 3. File README	34
Tabel 4. Jenis Lisensi	41
Tabel 5. Klasifikasi data menurut UKAN	49
Tabel 6. Spektrum Sensitivitas Data	50
Tabel 7. Biaya tarif PNBK Klirens Etik untuk mengajukan izin riset	62

PENDAHULUAN

A. Deskripsi Singkat

Mata pelatihan ini menjelaskan tentang perencanaan manajemen data, siklus data dalam manajemen data, reviu konten data, pengenalan metadata, dokumentasi data, etika dan integritas data, serta sensitivitas data.

B. Alokasi Waktu

Pelatihan Jabatan Fungsional Analisis Data Ilmiah dapat diselenggarakan dengan tiga metode. Setiap metode memiliki alokasi waktu yang berbeda. Berikut adalah alokasi waktu pembelajaran untuk mata pelatihan Manajemen Data.

Metode	On Kampus	Asinkronus	Sinkronus	Total
Klasikal	8 JP	-	-	8 JP
Bauran	5 JP	3 JP	3 JP	11 JP
Jarak Jauh	-	3 JP	6 JP	9 JP

C. Tujuan Pembelajaran

1. Hasil Belajar

Hasil belajar yang ingin dicapai melalui modul pelatihan ini adalah peserta pelatihan mampu memahami, mengaplikasikan, menganalisis, dan melakukan evaluasi laporan manajemen data dengan benar.

2. Indikator Hasil Belajar

Ketercapaian hasil belajar dapat diukur melalui indikator hasil belajar, yaitu peserta pelatihan mampu:

- a. memahami konsep dasar perencanaan manajemen data dan mengembangkan naskah dokumen perencanaan manajemen data

- b. memproses data dari pengumpulan data, integrasi, hingga diseminasi data;
- c. membuat revid konten untuk preservasi;
- d. memahami konsep dasar metadata dan mampu menentukan standar metadata yang akan digunakan;
- e. merevid konten untuk analisis;
- f. menerapkan etika data: lisensi, hak cipta/copyright, kode etik pengambilan data, data security; dan
- g. memahami prinsip perlindungan data.

D. Materi Pokok

Mata pelatihan ini terdiri dari 7 (tujuh) Materi Pokok, yaitu:

1. Perencanaan Manajemen Data
 - a. Deskripsi Data
 - b. Dokumentasi Data
 - c. Penyimpanan Data
 - d. Preservasi dan Berbagi Data
2. Siklus Data dalam Manajemen Data
 - a. Perencanaan dan Desain Data
 - b. Pengumpulan dan Dokumentasi Data
 - c. Kolaborasi dan Analisis Data
 - d. Pengelolaan, Penyimpanan dan Preservasi Data
 - e. Berbagi dan Publikasi Data
 - f. Penemuan, Penggunaan Kembali dan Sitasi Data
3. Revid Konten Data
 - a. Merevid Konten Data untuk Preservasi
 - b. Merevid Konten Data untuk Analisis
4. Pengenalan Metadata
 - a. Deskripsi Metadata
 - b. Ragam Metadata
 - c. Pemanfaatan Metadata untuk Proyek Penelitian
5. Dokumentasi Data
 - a. Deskripsi dan atau informasi teknis dari suatu data

6. Etika dan Integritas Data

- a. Definisi etika data
- b. Prinsip-prinsip etika data dalam pengelolaan dan pemanfaatan data
- c. Regulasi terkait etika data
- d. Ragam Lisensi Data

7. Sensitivitas Data

- a. Pengklasifikasian Data
- b. Klirens etik penelitian pada domain keilmuan
- c. Anonimisasi atau pseudonimisasi

MATERI POKOK 1:

PERENCANAAN MANAJEMEN DATA

Indikator Hasil Belajar

Peserta mampu:

1. Memahami konsep dasar perencanaan manajemen data dengan benar;
2. Mengembangkan naskah dokumen perencanaan manajemen data benar.

Benjamin Franklin — *'If you fail to plan, you are planning to fail!'*

Dengan perencanaan yang tepat, maka aktivitas manajemen data akan menjadi lebih mudah dan terarah. Hal ini termasuk dengan meluangkan waktu untuk merencanakan bagaimana Anda beserta tim akan mengelola data semenjak awal kegiatan, sehingga Anda dan tim dapat mempertimbangkan pertanyaan penting yang akan muncul sebelum, selama, dan setelah proyek berjalan. Hal ini tentunya akan membantu dalam mematuhi (*compliance*) kebijakan apa pun yang berlaku dalam pengelolaan data yang dimiliki¹ serta menjamin efektifitas dan keberhasilan penyelenggaraan suatu kegiatan yang dilakukan (Corti, 2020).

Perencanaan Manajemen Data (*Data Management Planning*, selanjutnya disingkat DMP) sangat penting untuk memastikan kualitas penelitian dan mendorong kemajuan ilmiah, serta merupakan persyaratan khusus dari banyak penyandang dana maupun institusi. Analisis terhadap lembaga penyandang dana di Amerika Serikat mengungkapkan bahwa 17 dari 22 kebijakan memerlukan dokumen rencana manajemen data sejak tahap proposal². Namun, sayangnya banyak periset yang belum memahami bagaimana cara membuat dokumen tersebut. Dalam survei yang ditujukan pada periset di Eropa, hanya 25 persen yang sebelumnya pernah menulis DMP, dan 25 persen sisanya tidak tahu apa itu DMP³.

¹ <https://masterclasses.nature.com/managing-research-data/18320206>

² <https://doi.org/10.29173/istl1556>

³ <https://doi.org/10.2777/121253>

“A data management plan is a formal statement describing how you will manage and document data in a research project, and how you will ensure long-term preservation.”

– JISC UK⁴

Berdasarkan definisi dari JISC UK, DMP merupakan suatu pernyataan yang bersifat formal yang disampaikan dalam rangka menggambarkan bagaimana data akan dikelola dan di dokumentasi dalam suatu proyek riset, serta bagaimana data tersebut dapat terjamin untuk diakses jangka panjang sesuai prosedur preservasi data. Adapun keuntungan yang didapatkan apabila Anda dan tim menyiapkan dokumen DMP dengan baik, yakni diantaranya:

1. meningkatkan efisiensi dalam kegiatan penelitian/proyek yang dilakukan
2. memastikan data dan informasi yang dihasilkan terproteksi dengan baik
3. luaran atau hasil analisis dari kegiatan penelitian/proyek dapat dengan mudah diperiksa oleh tim
4. meningkatkan kesempatan untuk dapat berbagi data
5. memenuhi kebijakan penyandang dana

DMP adalah dokumen penting yang akan membantu Anda mengidentifikasi data dan membuat keputusan dalam mengelolanya. Hal ini termasuk manajemen data pada saat pengumpulan, pengelolaan, pemrosesan, penyimpanan, dan berbagi data selama dan setelah proyek berjalan. Struktur, panjang kata ataupun jumlah laman, dan format DMP akan sangat bervariasi menurut disiplin ilmu, penyandang dana, dan kebijakan terkait (Parham, 2016). Namun, sebagai pemahaman awal, DMP dapat disesuaikan pada lingkup berikut:

1. Jenis data apa yang akan dikumpulkan dan bagaimana data tersebut diolah
2. Penyimpanan dan keamanan data, serta
3. Rencana pelestarian dan akses data setelah proyek selesai.

Saat membuat dokumen DMP, penulisan tidak hanya terbatas pada mengadopsi *template* yang sudah diisi di kegiatan/proyek/penelitian lain. Hal ini dikarenakan

⁴ <https://www.jisc.ac.uk/guides/rdm-toolkit/data-management-planning>

setiap dokumen DMP bersifat unik, dan disesuaikan dengan kebutuhan masing-masing kegiatan. Dokumen DMP harus disampaikan secara jelas dan ringkas, proporsional, terfokus, dan terstruktur. Sebaiknya, dokumen DMP ini sudah mulai dibuat di awal kegiatan agar Anda dan tim dapat secara konsisten dalam memproduksi data dan mengelolanya, seperti semenjak 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.

Setiap struktur dan unsur DMP sangatlah bervariasi, tergantung pada disiplin ilmu, penyandang dana serta letak geografis dimana kegiatan/proyek dilakukan. Apabila Anda dan tim berada dibawah naungan suatu institusi, maka institusi tersebut diharapkan memiliki *template* DMP yang telah disesuaikan dengan tujuan pokok dan fungsi institusi tersebut, sehingga Anda dan tim akan lebih mudah mengimplementasikannya dalam kegiatan proyek. Anda dapat menemukan template DMP yang sudah dirangkum oleh beberapa situs halaman berikut:

1. DMP Online, Digital Curation Centre, UK
(https://dmponline.dcc.ac.uk/public_templates)
2. DMP Tool, University of California Curation Center
(https://dmptool.org/public_templates)
3. Situs masing-masing instansi/penyandang dana lainnya.

Sumber lainnya mengenai DMP dapat dieksplorasi pada poin berikut:

1. Checklist for a Data Management Plan⁵ (Digital Curation Centre, UK)
2. Data Management Plans⁶ (Digital Curation Centre, UK)

Pada umumnya, dokumen DMP memiliki 1-3 halaman dan terdiri dari beberapa poin berikut:

1. Deskripsi Data
2. Dokumentasi Data

⁵ <https://www.dcc.ac.uk/DMPs/checklist>

⁶ <https://www.dcc.ac.uk/dmps>

3. Penyimpanan Data
4. Preservasi dan Berbagi Data

A. Deskripsi Data

Dalam poin ini, Anda dan tim diminta untuk memberikan pernyataan berisi deskripsi dari 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 diam maupun video yang mungkin berupa format digital ataupun non-digital. Anda juga perlu menjelaskan perkiraan besaran data yang akan dihasilkan, baik dalam bentuk megabyte, gigabyte hingga terabyte atau lebih. Estimasi besaran data ini dapat membantu Anda dan tim dalam memprediksi berapa besar penyimpanan yang dibutuhkan nantinya. Dan tentunya, hal ini dapat saja berubah, karena disesuaikan dengan produksi data yang akan terus bertambah selama kegiatan proyek berlangsung.

B. Dokumentasi Data

Dalam poin ini, Anda dan tim diminta untuk menyampaikan bagaimana strategi dalam mengumpulkan dan mendokumentasikan data yang akan dan telah dihasilkan. Selama Anda dan tim memproduksi data, tentunya diperlukan metode pengumpulan dan pendokumentasian yang disepakati antar anggota tim, agar konsisten dapat ditemukan dan terorganisasi dengan baik. Hal ini termasuk bagaimana penamaan file diberlakukan, apakah setiap data tersebut memiliki *persistent identifier*, serta bagaimana manajemen akan perubahan data/ *versioning*. Untuk membantu tahapan dalam poin ini, Anda dan tim dapat menyiapkan *ReadMe file* yang umumnya berbentuk teks (.txt) dengan mencantumkan setiap detail dokumentasi yang dilakukan.

C. Penyimpanan Data

Dalam poin ini, Anda dan tim diminta untuk menyampaikan rencana penyimpanan data yang akan dan telah dihasilkan, dengan mempertimbangkan berbagai aspek keamanan, level akses serta sensitivitas data yang dimiliki. Beberapa bentuk pernyataan penyimpanan data yang

akan dimasukkan ke dalam repositori institusi yang terintegrasi atau memiliki interoperabilitas dapat menjadi pertimbangan utama karena bentuk penyimpanan ini menjadi hal yang lumrah dimanfaatkan dan dikelola secara formal dan berkelanjutan oleh institusi terkait.

D. Preservasi dan Berbagi Data

Dalam poin ini, Anda dan tim diminta untuk memberikan pernyataan terkait bagaimana data yang Anda dan tim kelola akan dapat diakses jangka panjang, terpelihara dengan baik, serta dapat dimanfaatkan kembali oleh pengguna lain sebagai dataset sekunder. Tahapan preservasi yang dimaksud dapat disampaikan dalam rentang waktu jangka pendek, maupun jangka panjang. Pada jangka waktu pendek, preservasi data dapat mencakup jaminan data tersimpan dengan aman dan dapat diakses dengan baik. Sedangkan pada jangka waktu panjang, data perlu dapat juga diakses walaupun kegiatan penelitian/proyek telah selesai. Hal ini dapat diakomodir apabila seluruh data tersimpan dalam repositori institusi karena keberlanjutan pemeliharaan yang dimiliki oleh repositori tersebut. Selain itu, pernyataan apakah data yang dihasilkan oleh Anda dan tim akan dibagikan atau tidak, menjadi pernyataan yang perlu disampaikan dalam dokumen DMP ini. Hal tersebut terkait dengan kebijakan institusi yang mengatur kebermanfaatan kegiatan penelitian/proyek yang didukung. Poin berbagi data ini dapat juga mencakup bagaimana lisensi terhadap data disematkan.

E. Rangkuman

Berdasarkan definisi dari JISC UK, Data Management Planning (DMP) merupakan suatu pernyataan yang bersifat formal yang disampaikan dalam rangka menggambarkan bagaimana data akan dikelola dan di dokumentasi dalam suatu proyek riset, serta bagaimana data tersebut dapat terjamin untuk diakses jangka panjang sesuai prosedur preservasi data. Komponen dokumen DMP umumnya terdiri dari penjelasan mengenai Deskripsi Data, Dokumentasi Data, Penyimpanan Data, Preservasi dan Berbagi Data.

F. Evaluasi

1. Apakah yang dimaksud dengan *data management planning*?
2. Silahkan pilih dan pelajari template dokumen DMP yang berasal dari situs halaman DMP Online atau DMP Tool:
 - a. Dokumen DMP mana yang Anda pilih? Berasal dari instansi mana?
 - b. Apa saja komponen yang dimiliki dokumen DMP tersebut?
 - c. Apa alasan Anda memilih dokumen DMP tersebut, apakah ada relasi dengan kegiatan yang Anda jalani saat ini?

MATERI POKOK 2:

SIKLUS DATA DALAM MANAJEMEN DATA

Indikator Hasil Belajar

Peserta mampu:

1. Memahami siklus manajemen data dengan benar
2. Memproses data dari pengumpulan data, integrasi, hingga diseminasi data dengan benar

Dalam suatu kegiatan, tentunya menjadi hal yang cukup lumrah bagi para manajer untuk melaksanakan aktivitas manajerial yang dimulai dari perencanaan hingga evaluasi ataupun pemantauan. Namun, karena saat ini membahas perihal data, maka siklus manajemen data akan disesuaikan dengan objek yang dikelola, baik itu data secara umum maupun data yang digunakan untuk keperluan riset dan pengambilan kebijakan. Adapun siklus data tersebut dapat disampaikan melalui Gambar 1 dibawah ini.



Gambar 1. Siklus data riset menurut JISC UK⁷

⁷ <https://www.jisc.ac.uk/guides/rdm-toolkit>

Dalam konteks umum, data dapat disesuaikan dengan tingkatan proses berikut:

1. Data mentah (*raw data*)
Merupakan data yang dihasilkan langsung dari kegiatan pengumpulan data tanpa ada perubahan berarti. Contoh dari data mentah ini diantaranya yakni hasil pengukuran atau hasil uji laboratorium harian. Data mentah tersebut umumnya dihasilkan pada tahap awal dari pembuatan data dari siklus hidup data (*data creation*).
2. 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.
3. Data hasil analisis (*analyzed data*)
Merupakan luaran data hasil analisis yang menunjukkan signifikansi atau temuan.
4. 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.

Untuk dapat memahami siklus manajemen data dan memproses data terkait penelitian/proyek yang dilakukan, maka Anda dan tim perlu melaksanakan beberapa tahapan siklus yang dicantumkan pada Gambar 1 berikut, yakni:

1. Perencanaan dan Desain Data
2. Pengumpulan dan Dokumentasi Data
3. Kolaborasi dan Analisis Data
4. Pengelolaan, Penyimpanan dan Preservasi Data
5. Berbagi dan Publikasi Data
6. Penemuan, Penggunaan Kembali dan Sitasi Data

A. Perencanaan dan Desain Data

Untuk dapat mengelola data, maka diperlukan tahapan perencanaan dan desain terhadap data tersebut. Hal ini dilakukan untuk menjamin data tersebut dikelola dengan baik, diakses oleh pihak yang memiliki otoritas dan

dapat diakses jangka panjang sesuai prinsip FAIR (*Findable, Accessible, Interoperable dan Reusable*)⁸. Prinsip ini menjadi dasar pelaksanaan pengorganisasian data, khususnya perihal data sensitif yang perlu mempertimbangkan etika dan regulasi terkait. Detail dari prinsip FAIR ini dapat dilihat sebagai berikut:

1. *Findable*: data yang dikelola adalah yang sudah memiliki metadata dan persistent identifier agar mudah ditemukan.
2. *Accessible*: data dan metadata yang dimaksud perlu untuk dapat ditemukan kembali dengan akses yang mudah, yang umumnya setidaknya metadata dapat ditemukan, walau akses pada data didalamnya tertutup.
3. *Interoperable*: metadata sebaiknya dilengkapi oleh subjek terkontrol, *machine-readable*, dan menggunakan standar metadata internasional.
4. *Reusable*: melalui standar metadata yang bersifat internasional, maka akan mudah untuk dapat digunakan kembali dimana data dan metadata tersebut telah disematkan lisensi sebagai ketentuan akses dan penggunaan kembalinya.

Anda dan tim diharapkan melihat kegiatan manajemen data ini seperti *project management*, dimana setiap tahapan perlu direncanakan dengan baik dan memiliki manajemen resiko tersendiri. Hal tersebut didasarkan pada kebijakan masing-masing institusi, apakah kategori data yang dikelola termasuk informasi publik ataupun non-publik. Regulasi terkait batasan informasi ini dapat dilihat pada Undang-Undang Nomor 14 Tahun 2018 Tentang Keterbukaan Informasi Publik⁹. Untuk perencanaan dan desain data ini, Anda dan tim dapat merujuk pada Materi Pokok 1 mengenai Perencanaan Manajemen Data.

⁸ <https://www.go-fair.org/fair-principles/>

⁹

https://jdih.kominfo.go.id/produk_hukum/view/id/172/t/undangundang+nomor+14+tahun++2008+tanggal+30+april+2008

B. Pengumpulan dan Dokumentasi Data

Setelah menyusun dokumen perencanaan dan desain data yang akan dikelola, Anda dan tim sudah dapat mengumpulkan dan mendokumentasikan data pada setiap aktivitas penelitian/proyek yang dilakukan. Proses ini bersifat layaknya pengarsipan terhadap data, dengan detail sebagai berikut:

1. Menggunakan sistem pendokumentasian yang mengakomodir pencadangan data;
2. Mendokumentasikan secara lengkap data sesuai isian field metadata yang dimiliki. Semakin akurat isian field metadata, maka akan semakin kaya informasi terkait data dan mudah untuk ditemukan kembali, digunakan kembali serta disitasi;
3. Menetapkan format file, penamaan file dan struktur folder yang konsisten.

C. Kolaborasi dan Analisis Data

Dalam tahapan ini, kegiatan mengelola data umumnya dilaksanakan secara berkelompok atau dalam organisasi tim. Oleh karena itu, tahapan terkait kolaborasi data mungkin dilakukan, dimana termasuk penentuan sumber data, penentuan peran dan tanggung-jawab dalam tim, serta penentuan perubahan/*versioning* data sebagai bahan analisis dalam tim. Hal ini menyangkut pengelolaan data aktif, yang mana merupakan data yang sudah bersih dan siap untuk dilakukan analisis. Aktivitas tersebut bisa termasuk bagaimana Anda dan tim menentukan anggota yang bertugas menjadi pengelola utama data (*data manager*), sehingga ia memiliki otoritas penuh untuk memastikan data yang dihasilkan dan digunakan merupakan data yang sesuai. Kegiatan analisis data pun kemudian dapat dilakukan, yang umumnya disampaikan melalui data terolah ataupun visualisasi. Sehingga, ragam jenis data yang dikelola dapat beragam, dari bentuk teks, audio, gambar, maupun bentuk lainnya.

D. Pengelolaan, Penyimpanan dan Preservasi Data

Dalam tahapan mengelola, menyimpan dan mempreservasi data, Anda dan tim diharapkan dapat melaksanakan kegiatan yang sesuai dengan dokumen

perencanaan manajemen data (DMP) yang telah diajukan. Hal ini termasuk bagaimana data secara teknis dikelola yang berdasarkan regulasi, mencakup Deskripsi Data, Dokumentasi Data, Penyimpanan Data, serta Preservasi dan Berbagi Data.

Terkait preservasi, banyak pemahaman yang perlu untuk diluruskan. Penyimpanan data dalam media portable, repositori institusi bahkan data yang dicadangkan pun bukanlah termasuk bagian dari preservasi data. Preservasi data digital merupakan serangkaian aktivitas yang dilakukan untuk menjamin keberlanjutan akses terhadap konten data digital, dimana didalamnya adalah memastikan keterbacaan file digital. Anda dapat melihat lebih detail aktivitas preservasi ini pada Materi Pokok 3 mengenai Reviu Konten Data, khususnya untuk preservasi.

E. Berbagi dan Publikasi Data

Tahapan ini mencakup bagaimana data siap untuk dipublikasi dan disebarkan kepada publik. Oleh karena data merupakan komponen yang dapat mengandung informasi sensitif didalamnya, maka pada tahapan ini, informasi sensitif tersebut sudah di filterisasi, sehingga siap dan layak untuk diakses sebagai informasi publik. Filterisasi data yang dimaksud disini adalah data tersebut sudah melalui tahap de-identifikasi, yakni tidak memiliki informasi personal setelah dilakukan anonimisasi, pseudonimisasi, maupun kodifikasi sesuai etika keilmuan terkait. Umumnya untuk mengakses data publik ini, dapat dilihat pada fasilitas penyimpanan yang tersedia, seperti repositori data ¹⁰, jurnal data ¹¹ maupun data pendukung publikasi (*supplementary materials*). Anda dan tim perlu menyematkan lisensi terhadap data yang dipublikasi dan siap dibagikan, untuk membantu pihak eksternal saat ingin menggunakan kembali data tersebut. Lisensi yang umum digunakan adalah Creative Commons ¹², sehingga Anda dan tim dapat diberikan atribusi atau rekognisi saat data digital tersebut digunakan oleh

¹⁰ <https://www.re3data.org/>

¹¹ <https://www.wiki.ed.ac.uk/display/dataset+peer+review>

¹² <https://creativecommons.org/choose/>

pihak eksternal.

F. Penemuan, Penggunaan Kembali dan Sitasi Data

Pada tahapan ini, Anda dan tim memiliki dua peranan yaitu sebagai pengguna data dan produsen data. Pengguna data yang dimaksud adalah saat Anda dan tim memanfaatkan dataset lain yang telah dipublikasi di repositori data dan menggunakannya sebagai bagian dari analisis pada kegiatan penelitian/proyek yang sedang dilakukan. Untuk memanfaatkannya, tentu Anda dan tim perlu melakukan pencarian dan penemuan kembali dataset yang diinginkan, yang mana akan mudah ditemukan pada Repositori Data yang telah terindeks di Re3Data¹⁰. Sedangkan, sebagai produsen data, Anda dan tim menghasilkan dataset yang nantinya akan disimpan di Repositori Data yang telah disepakati, sehingga pihak eksternal dapat menemukan dan menggunakannya kembali. Dalam prosesnya, terdapat aktivitas sitasi yang dilakukan, seperti halnya memberlakukan sitasi pada publikasi. Setiap data digital yang Anda dan tim produksi tentunya memiliki *digital object identifier* (DOI, atau *persistent identifier*) yang akan memudahkan pihak eksternal saat ingin merujuk dataset tersebut.

Diharapkan dengan semakin banyaknya implementasi siklus manajemen data yang baik, akan menghasilkan produksi dataset yang berkualitas dan memudahkan pemanfaatan kembali data berbasis sitasi, sehingga setiap kegiatan akan berbasiskan data yang FAIR: mudah ditemukan, dapat diakses, interoperabel (mudah dibagi pakaikan antar pengguna data) dan dapat digunakan kembali.

G. Rangkuman

Siklus manajemen data terdiri dari Perencanaan dan Desain Data, Pengumpulan dan Dokumentasi Data, Kolaborasi dan Analisis Data, Pengelolaan, Penyimpanan dan Preservasi Data, Berbagi dan Publikasi Data, serta Penemuan, Penggunaan Kembali dan Sitasi Data. Melalui implementasi siklus manajemen data, diharapkan dapat terciptanya prinsip FAIR data yang mendukung pemanfaatan kembali data di setiap kegiatan penelitian/proyek.

H. Evaluasi

1. Apa saja tahapan dari siklus manajemen data?
2. Bagaimana Anda menerapkan siklus manajemen data dalam program di institusi Anda?

MATERI POKOK 3:

REVIU KONTEN DATA

Indikator Hasil Belajar

Peserta mampu:

1. Membuat reviu konten data untuk preservasi
2. Membuat reviu konten data untuk analisis

A. Mereviu Konten Data untuk Preservasi

Preservasi data bertujuan untuk menjaga keamanan konten (data) dan konteks (metadata) untuk digunakan kembali di masa mendatang. Preservasi menjadi jauh lebih mudah jika dipertimbangkan dalam proses perencanaan manajemen data Anda. Anda dapat merujuk pada Materi Pokok 1 terkait Perencanaan Manajemen Data untuk lebih detail. Dalam melakukan reviu konten data dalam rangka preservasi, maka diperlukan informasi dari karakteristik format file untuk memastikan akses jangka panjang tersebut, yakni:

1. *Non-proprietary*

Format *non-proprietary* atau terbuka merupakan format yang dapat dibaca tidak hanya oleh peralatan dan/atau program yang membuatnya, sehingga dapat bebas untuk dikonversi. Umumnya bukan merupakan format yang dikenakan paten atau merek dagang.

2. Tidak terenkripsi / terkompresi

File yang tidak terenkripsi dan tidak terkompresi menawarkan prospek terbaik untuk akses jangka panjang. Jika file dienkripsi dan/atau dikompresi, metode enkripsi/kompresi yang digunakan harus dapat ditemukan dan digunakan untuk akses file di masa mendatang.

Berikut adalah beberapa contoh format pilihan untuk berbagai tipe data yang dikutip berdasarkan Data Archiving and Networked Service (DANS)¹³, yaitu:

¹³ <https://dans.knaw.nl/en/file->

Type	Preferred format(s)	Non-preferred format(s)
Text documents	<ul style="list-style-type: none"> ● PDF/A (.pdf) ● ODT (.odt) 	<ul style="list-style-type: none"> ● Microsoft Word (.doc) ● Office Open XML (.docx) ● Rich Text File (.rtf) ● PDF other than PDF/A (.pdf)
Plain text	<ul style="list-style-type: none"> ● Unicode text (.txt) 	<ul style="list-style-type: none"> ● Non-Unicode text (.txt)
Markup language	<ul style="list-style-type: none"> ● XML (.xml) ● HTML (.html) ● Related files: .css, .xslt, .js, .es 	<ul style="list-style-type: none"> ● SGML (.sgml) ● Markdown (.md)
Programming languages	<ul style="list-style-type: none"> ● MATLAB ● NetCDF ● Text-Fabric 	
Spreadsheets	<ul style="list-style-type: none"> ● ODS (.ods) ● CSV (.csv) 	<ul style="list-style-type: none"> ● Microsoft Excel (.xls) ● Office Open XML Workbook (.xlsx) ● PDF/A (.pdf)
Databases	<ul style="list-style-type: none"> ● SQL (.sql) ● SIARD (.siard) ● CSV (.csv) 	<ul style="list-style-type: none"> ● Microsoft Access (.mdb, .accdb) ● dBase (.dbf) ● HDF5 (.hdf5, .he5, .h5)
Statistical data	<ul style="list-style-type: none"> ● SPSS (.dat/.sps) ● STATA (.dat/.DO) ● R 	<ul style="list-style-type: none"> ● SPSS Portable (.por) ● SPSS (.sav) ● STATA (.dta) ● SAS (.7dat; .sd2; .tpt)
Raster images	<ul style="list-style-type: none"> ● JPEG (.jpg, .jpeg) ● TIFF (.tif, .tiff) ● PNG (.png) ● JPEG 2000 (.jp2) ● DICOM (.dcm) 	
Vector images	<ul style="list-style-type: none"> ● SVG (.svg) 	<ul style="list-style-type: none"> ● Adobe Illustrator (.ai) ● EPS (.eps) ● WMF/EMF (.wmf, .emf) ● CDR (.cdr)
Audio	<ul style="list-style-type: none"> ● BWF (.bwf) ● MXF (.mxf) ● Matroska (.mka) ● FLAC (.flac) ● OPUS 	<ul style="list-style-type: none"> ● WAVE (.wav) ● MP3 (.mp3) ● AAC (.aac, .m4a) ● AIFF (.aif, .aiff) ● OGG (.ogg)
Video	<ul style="list-style-type: none"> ● MXF (.mxf) ● Matroska (.mkv) 	<ul style="list-style-type: none"> ● MPEG-4 (.mp4, .m4a, .m4v, ...) ● MPEG-2 (.mpg, .mpeg, .m2v, .mpg2, ...) ● AVI ● QuickTime (.mov, .qt)
Computer Aided Design (CAD)	<ul style="list-style-type: none"> ● AutoCAD DXF version R12 (ASCII) (.dxf) ● SVG (.svg) 	<ul style="list-style-type: none"> ● Autocad DXF, version different than R12 (ASCII) (.dxf) ● DWG (.dwg) ● DGN (.dgn)

[formats/#:~:text=Preferred%20formats%20are%20file%20formats,always%20be%20accepted%20by%20DANS.](#)

Type	Preferred format(s)	Non-preferred format(s)
Geographical Information Systems (GIS)	<ul style="list-style-type: none"> ● GML (.gml) ● MIF/MID (.mif/.mid) ● GeoJSON (.json) 	<ul style="list-style-type: none"> ● Esri Shapefiles (.shp & related files) ● MapInfo (.tab & related files) ● KML (.kml, .kmz) ● Esri Geodatabase (.gdb) ● Project files / Workspaces (.mxd, .wor, .qgs)
Georeferenced images	<ul style="list-style-type: none"> ● GeoTIFF (.tif, .tiff) 	<ul style="list-style-type: none"> ● TIFF World File (.tfw & .tif, possibly with additional files) ● JPEG World File (.jgw & .jpg, possibly with additional files) ● ERDAS IMAGINE File Format (.img)
Raster GIS	<ul style="list-style-type: none"> ● ASCII GRID (.asc, .txt) 	<ul style="list-style-type: none"> ● Esri GRID (.grd & bijbehorende bestanden) ● Surfer Grid (.grd; .srf) ● ERDAS IMAGINE File Format (.img)
3D	<ul style="list-style-type: none"> ● WaveFront Object (.obj) ● Polygon file format (.ply) ● X3D (.x3d) ● COLLADA (.dae) 	<ul style="list-style-type: none"> ● Autodesk FBX (.fbx) ● Blender (.blend) ● 3D PDF (.pdf)
RDF	<ul style="list-style-type: none"> ● RDF/XML (.rdf) ● Trig (.trig) ● Turtle (.ttl) ● NTriples (.nt) ● JSON-LD 	
Computer Assisted Qualitative Data Analysis (CAQDAS)	<ul style="list-style-type: none"> ● REFI-QDA (Qualitative Data Analysis) 	<ul style="list-style-type: none"> ● ATLAS.TI copy bundle ● NVivo project file

Selain itu, untuk daftar format file umum, evaluasi kualitas format serta keberlanjutan jangka panjang, Anda dan tim dapat melihat Sustainability Digital Format¹⁴ yang disampaikan oleh Library of Congress. Beberapa metode pengarsipan yang dilakukan komunitas kearsipan juga menentukan format data optimal yang mereka gunakan untuk penyimpanan data jangka panjang. Salah satu contohnya adalah format file Arsip Data Inggris¹⁵. Daftar dan tabel tersebut berisi panduan tentang format file yang diterima oleh arsip data Inggris Raya untuk data yang disimpan guna kemudahan berbagi data, pemanfaatan kembali dan preservasi.

¹⁴ https://www.loc.gov/preservation/digital/formats/fdd/browse_list.shtml

¹⁵ <https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/recommended-formats/>

B. Mereviu Konten Data untuk Analisis

Selain digunakan untuk preservasi, data juga digunakan untuk analisis. Data yang digunakan sebagai analisis merupakan kumpulan atau gabungan data dari satu maupun beberapa sumber, yang disebut dengan dataset. Dalam *data mining*, kegiatan mereviu konten data untuk analisis disebut dengan *data preprocessing*, yang merupakan tahapan prapemrosesan data sebelum diolah menggunakan algoritma tertentu untuk di-analisis.

Kegiatan mereviu konten data untuk analisis dimaksudkan untuk mendapatkan data yang berkualitas agar menghasilkan analisis yang akurat. Data dengan kualitas rendah akan menghasilkan akurasi analisis yang rendah. Pun sebaliknya, data berkualitas tinggi, akan menghasilkan akurasi analisis yang tinggi. Kualitas data dipengaruhi oleh unsur-unsur: akurasi, kelengkapan, konsistensi, tepat waktu, dapat dipercaya, dan dapat diinterpretasikan (Han, J. et al. 2012).

Reviu konten data dilakukan menggunakan beberapa metode (Han, J. et al. 2012) yang terdiri dari:

a. Pembersihan data (*data cleaning/cleansing*)

Proses membersihkan data dilakukan dengan cara mengisi nilai yang kosong (*missing values*), menghaluskan data yang bermasalah atau tidak lengkap (*noisy data*), mengidentifikasi atau menghapus nilai pencilan (*outliers*), menyelesaikan inkonsistensi data serta mendeteksi dan menghapus data berulang (saat integrasi data).

b. Integrasi data (*data integration*)

Proses mendapatkan data dari berbagai sumber dengan mengintegrasikan beberapa basisdata (*databases*), *data cubes*, atau *files*.

c. Reduksi data (*data reduction*)

Proses ini menghasilkan pengurangan representasi dataset dalam ukuran yang lebih kecil, namun tetap menghasilkan hasil analisis yang sama (atau hampir sama). Strategi pengurangan data meliputi pengurangan dimensi dan pengurangan jumlah.

d. Transformasi data (*data transformation*)

Proses ini mengubah data dari satu bentuk ke bentuk lainnya melalui teknik representasi yang berbentuk normalisasi, diskretisasi, dan konsep generasi hirarki.

C. Rangkuman

Preservasi data bertujuan untuk menjaga keamanan konten (data) dan konteks (metadata) untuk digunakan kembali di masa mendatang.

Kegiatan mereviu konten data untuk analisis dimaksudkan untuk mendapatkan data yang berkualitas agar menghasilkan analisis yang akurat.

D. Evaluasi

1. Buatlah review konten untuk preservasi (data dapat menggunakan berbagai sumber)
2. Buatlah review konten data untuk analisis

MATERI POKOK 4:

PENGENALAN METADATA

Indikator Hasil Belajar:

Peserta mampu:

1. Memahami konsep metadata
2. Menentukan standar metadata yang akan digunakan

A. Deskripsi Metadata

Dalam konteks institusi pengelola data dan informasi seperti pusat data dan dokumentasi, perpustakaan, museum, galeri, dan unit kearsipan, keberadaan data dan informasi sebagai sumber daya harus dapat terdeskripsikan dengan baik dan tepat. Pendeskripsian sumber daya yang dimiliki harus dapat memenuhi kebutuhan informasi pengguna. Untuk tujuan tersebut, maka data dan informasi yang dimiliki memerlukan pengorganisasian mengenai informasi yang tersedia. Pada institusi-institusi tersebut, penggunaan metadata sebagai media dalam pendeskripsian sumber daya lazim digunakan untuk menghasilkan katalog. Institusi-institusi tersebut tentunya memiliki konsep dan implementasi yang berbeda, tergantung pada jenis sumber daya, pendekatan komunitas untuk mengatur informasi, dan kebutuhan penggunaannya (Joudrey et al., 2018).

Definisi metadata secara umum adalah data tentang data, dalam hal ini berarti adanya pendeskripsian suatu sumber daya informasi untuk menghasilkan data lainnya. NISO (2008) mendefinisikan metadata adalah informasi terstruktur yang terkait dengan objek tertentu dengan tujuan untuk penemuan, deskripsi, penggunaan, pengelolaan, dan pelestarian. Selain itu, Farnel & Shiri (2014) menyebutkan bahwa metadata adalah informasi terstruktur yang menyediakan konteks untuk semua jenis objek informasi, termasuk data penelitian. Konteks lain dari sumber daya digital adalah halaman web, file komputer, gambar, dokumen multimedia, dan sejenisnya

(Ramesh et al., 2015).

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

- *Metadata deskriptif*

Metadata deskriptif berisi identifikasi paling penting untuk tujuan penemuan, identifikasi, seleksi, dan akuisisi dari sumber informasi. Metadata deskriptif mencakup jenis informasi berdasarkan tujuan sebagai berikut:

- Mengidentifikasi data: judul, pernyataan tanggung jawab, tanggal publikasi atau distribusi, bahasa, format, jenis sumber daya, pengidentifikasi;
- Menghubungkan data: pencipta, kontributor, penerbit, narasumber, karya terkait, identifikasi hubungan lain di antara berbagai entitas; dan
- Mengisi data: subjek, abstrak, cakupan, daftar isi, notasi klasifikasi, dan kategori.

- *Metadata administratif*

Metadata administratif dibuat untuk tujuan pengelolaan, pengambilan keputusan, dan pencatatan. Metadata ini memberikan informasi tentang persyaratan teknis, pelestarian, dan penyimpanan sumber daya informasi, khususnya objek digital. Metadata administratif membantu memantau, mengakses, mereproduksi, mendigitalkan, dan back-up sumber daya digital. Metadata administratif dapat dibagi lagi menjadi beberapa sub-metadata, seperti metadata teknis, metadata preservasi, metadata kepemilikan dan akses, dan meta-metadata.

Metadata teknis: berfokus pada penggambaran fisik daripada karakteristik intelektual objek digital. Informasi teknis diperlukan untuk mengidentifikasi sumber daya, perangkat lunak dan perangkat keras yang digunakan untuk penciptaan sumber daya, dan apa saja yang diperlukan untuk mengakses sumber daya. Metadata teknis menjelaskan karakteristik, asal, dan siklus hidup dokumen digital dan menjadi acuan untuk pelestarian sumber daya agar dapat digunakan di waktu yang akan datang.

Metadata preservasi: berisikan informasi yang diperlukan untuk memastikan penyimpanan jangka panjang dan penggunaan sumber daya digital. Hal ini mencakup informasi tentang pemformatan ulang, migrasi, emulasi, konservasi, integritas file, dan sumber yang digunakan.

Metadata kepemilikan dan akses adalah informasi tentang siapa yang memiliki akses ke sumber daya informasi, siapa yang dapat menggunakannya, dan untuk tujuan apa. Metadata ini berurusan dengan masalah hak kekayaan intelektual pencipta dan perjanjian hukum yang memungkinkan pengguna untuk mengakses sumber daya.

Meta-metadata: Jika metadata adalah “data tentang data”, maka meta-metadata adalah “data tentang data tentang data”. Metadata tidak hanya dapat melacak data administratif tentang sumber daya, tetapi juga dapat melacak informasi administratif tentang metadatanya. Meta-metadata penting untuk memastikan keaslian metadata dan melacak prosesnya.

- Metadata struktural

Metadata struktural mengacu pada susunan atau pengaturan dari objek digital, dataset, atau sumber informasi lain yang sedang dideskripsikan. Ini adalah data yang diperlukan untuk memastikan bahwa sumber daya digital berfungsi dengan baik dan dapat digunakan serta dinavigasi oleh pengguna. Hal ini juga berarti bahwa metadata menjelaskan mengenai bagaimana file terkait dihubungkan agar berfungsi dengan baik, bagaimana file dapat ditampilkan di berbagai sistem, dan bagaimana file dapat disimpan dan disebarluaskan.

Disadur dari Research Data Netherlands (n.d.), berikut penjabaran jenis metadata secara ringkas:

Tabel 1. Jenis Metadata

Jenis Metadata	Tujuan	Contoh
Metadata deskriptif	Metadata minimum yang diperlukan untuk menemukan objek digital.	Penulis, judul, abstrak, tanggal, dsb.

Metadana struktural	Merekam hubungan antara objek digital individu/data yang membentuk satu unit.	<i>Link</i> ke objek digital terkait, (misalnya artikel yang ditulis berdasarkan data penelitian yang ditautkan).
Metadana teknis	Informasi tentang aspek teknis data.	Format data, hardware/software yang digunakan, kalibrasi, versi, otentikasi, enkripsi, standar metadana.
Metadana administratif	Berfokus pada hak penggunaan dan pengelolaan objek digital.	Lisensi, kemungkinan alasan embargo, pengabaian, log pencarian, pelacakan pengguna.

Pada dasarnya, penggunaan metadana sangat bergantung pada kebutuhan masing-masing komunitas dalam mengelola sumber daya. Namun yang perlu dipahami adalah bahwa metadana harus dapat mendukung seluruh aspek pengelolaan dari sumber daya digital. NISO (2008) memaparkan setidaknya metadana harus mencakup enam prinsip, yaitu:

1. Metadana yang baik sesuai dengan standar komunitas, yaitu sesuai dengan materi dalam koleksi, pengguna koleksi, dan penggunaan koleksi saat ini dan di masa mendatang.
2. Metadana yang baik mendukung interoperabilitas.
3. Metadana yang baik menggunakan kendali otoritas (authority control) dan standar konten untuk mendeskripsikan sumber daya dan menyusun sumber daya yang terkait.
4. Metadana yang baik mencakup pernyataan yang jelas tentang kondisi dan ketentuan penggunaan untuk sumber daya digital.
5. Metadana yang baik mendukung kurasi jangka panjang dan preservasi sumber daya.
6. Metadana yang baik adalah sumber daya itu sendiri. Oleh karena itu harus memiliki kualitas sumber daya yang baik, termasuk otoritas, keaslian, kemampuan pengarsipan, dan pengidentifikasi unik.

B. Ragam Metadata

a. Dublin Core

Adanya Dublin Core Metadata Element Set (DCMES) berawal dari pertemuan tahun 1995 di Dublin, Ohio, yang berfokus pada metadata untuk jaringan informasi elektronik. Dalam pertemuan ini lebih dari 50 orang yang terdiri dari penerbit, pakar bidang komputer, pustakawan, peneliti informasi, produsen perangkat lunak, dan pakar text-markup membahas bagaimana rangkaian semantik untuk sumber daya berbasis web / web-based resources akan sangat berguna untuk pencarian dan pengambilan yang lebih mudah. Pertemuan pertama ini menghasilkan 13 elemen inti yang didefinisikan, yang kemudian berkembang menjadi 15 elemen yang dikenal sebagai DCMES saat ini. Elemen-elemen tersebut adalah: contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title, type (The Dublin Core Metadata Initiative, n.d.).

Elemen-elemen inti dari DCMES dapat dijabarkan sebagai berikut:

Tabel 2. Elemen-elemen inti dari DCMES

Nama Istilah	Label	Deskripsi
contributor	Contributor	Entitas yang berkontribusi untuk membuat sumber daya
coverage	Coverage	Pernyataan mengenai tempat atau periode yang relevan dengan isi sumber daya
creator	Creator	Entitas utama yang membuat/menghasilkan sumber daya
date	Date	Periode waktu yang terkait dengan ketersediaan sumber daya.
description	Description	Deskripsi yang dapat mencakup namun tidak terbatas pada: abstrak, daftar isi, uraian
format	Format	Format file, media fisik, atau dimensi sumber daya
identifier	Identifier	Referensi yang unik dan tidak ambigu yang digunakan untuk mengidentifikasi ke sumber daya.
language	Language	Bahasa yang digunakan

publisher	Publisher	Pihak yang bertanggung jawab untuk menyediakan/mempublikasikan sumber daya
relation	Relation	Sumber daya lain yang berhubungan
rights	Rights	Informasi tentang hak cipta / hak milik yang dimiliki atas sumber daya.
source	Source	Informasi tentang asal dari sumber daya terkait
subject	Subject	Topik dari sumber daya
title	Title	Judul yang digunakan untuk sumber daya
type	Type	Jenis/bentuk dari sumber daya

Sumber: <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>

b. DataCite

Skema Metadata DataCite adalah daftar elemen metadata inti yang digunakan untuk mengidentifikasi sumber daya secara akurat dan konsisten untuk tujuan pengutipan dan temu kembali. Sumber daya yang diidentifikasi dapat berbagai jenis, tetapi biasanya berupa dataset penelitian. Berbeda dengan Dublin Core, skema metadata dari DataCite menyatakan tiga tingkatan dalam penggunaannya, yaitu a) elemen inti yang wajib disediakan, b) elemen yang dapat direkomendasikan dan sangat disarankan untuk interoperabilitas, dan c) opsional untuk disediakan untuk menambah nilai informasi dari dataset (Smaele et al., 2016).

- a) Elemen inti wajib: *Identifier, Creator, Title, Publisher, PublicationYear,* dan *ResourceType*
- b) Elemen yang direkomendasikan dan sangat disarankan: *Subject, Contributor, Date, RelatedIdentifier, Description,* dan *GeoLocation.*
- c) Elemen opsional: *Language, AlternateIdentifier, Size, Format, Version, Rights, FundingReference,* dan *RelatedItem.*

Banyak pengguna skema DataCite berlatarbelakang dari berbagai bidang ilmu. Selain itu, oleh karena DataCite banyak digunakan untuk dataset penelitian, skema ini mendukung adanya pengutipan terhadap data. DataCite menyediakan format pengutipan yang dapat dibaca manusia maupun mesin. Di bawah ini adalah contoh format pengutipan dari dataset

DataCite:

*Creator (PublicationYear): Title. Version. Publisher.
(resourceTypeGeneral). Identifier*

Irino, T; Tada, R (2009): Chemical and mineral compositions of sediments from ODP Site 127-797. V. 2.1. Geological Institute, University of Tokyo. (dataset).
<https://doi.org/10.1594/PANGAEA.726855>

C. Pemanfaatan Metadata untuk Proyek Penelitian

Saudara calon Analis Data Ilmiah, pada materi pokok 2 Anda telah mengetahui prinsip FAIR terhadap data, bahwa data harus lah dapat ditelusur, diakses, diintegrasikan dengan sistem lain, dan direproduksi kembali. Namun di samping itu, dalam suatu proyek penelitian, ada peran signifikan yang tidak kalah penting dari metadata. FAIR data kurang lebih sebagai pengejawantahan dari seberapa jauh metadata dapat berfungsi dengan baik dalam pengelolaan data.

Berdasarkan pemaparan dari GO FAIR, penjabaran lebih detail mengenai prinsip FAIR sebagai berikut (GO FAIR, n.d.):

Findable

Langkah pertama dalam menggunakan (kembali) data adalah dengan menemukannya. Metadata dan data harus mudah ditemukan oleh manusia dan komputer. Metadata yang dapat dibaca mesin sangat penting untuk penemuan terotomasi untuk dataset dan layanan. Prinsip *Findable* terdiri dari:

F1. (Meta)data diberi pengenal yang unik dan persisten secara global

F2. Data dijelaskan dengan metadata yang lengkap dan informatif

F3. Metadata secara jelas dan eksplisit menyertakan pengidentifikasi data yang digambarkan

F4. (Meta)data didaftarkan atau diindeks dalam sumber daya yang dapat dicari

Accessible

Setelah pengguna menemukan data yang diperlukan, mereka perlu mengetahui bagaimana data tersebut dapat diakses, termasuk juga autentikasi dan otorisasi. Prinsip *Accessible* terdiri dari:

A1. (Meta)data dapat diambil kembali oleh pengenalnya menggunakan protokol komunikasi terstandar

A1.1. Protokol bersifat terbuka, gratis, dan dapat diterapkan secara universal

A1.2. Jika diperlukan, protokol memungkinkan prosedur autentikasi dan otorisasi

A2. Metadata dapat diakses bahkan saat data tidak lagi tersedia

Interoperable

Data biasanya perlu diintegrasikan dengan data lain. Selain itu, data perlu beroperasi dengan aplikasi atau alur kerja untuk analisis, penyimpanan, dan pemrosesan. Prinsip *Interoperable* terdiri dari:

I1. (Meta)data menggunakan bahasa formal, dapat diakses, dibagikan, dan dapat diterapkan secara luas untuk merepresentasikan pengetahuan.

I2. (Meta)data menggunakan kosakata yang mengikuti prinsip FAIR

I3. (Meta)data menyertakan referensi yang memenuhi syarat ke (meta)data lainnya

Reusable

Tujuan akhir dari FAIR adalah untuk mengoptimalkan penggunaan kembali data. Untuk mencapai hal ini, metadata dan data harus dideskripsikan dengan baik sehingga dapat direplikasi dan/atau digabungkan dalam pengaturan yang berbeda.

R1. (Meta)data dideskripsikan secara lengkap dan informatif dengan sejumlah atribut yang akurat dan relevan

R1.1. (Meta)data dirilis dengan lisensi penggunaan data yang jelas dan dapat diakses

R1.2. (Meta)data dikaitkan dengan asal-usul (*provenance*) secara rinci


R1.3. (Meta)data memenuhi standar domain dari komunitas yang relevan


Prinsip-prinsip FAIR mengacu pada tiga jenis entitas yaitu data (atau objek digital apa pun), metadata (informasi tentang objek digital tersebut), dan infrastruktur. Misalnya, prinsip F4 mendefinisikan bahwa baik metadata maupun data didaftarkan atau diindeks dalam sumber daya yang dapat dicari (komponen infrastruktur).


Prinsip FAIR telah disepakati secara global dalam pengelolaan data. Engelhardt et al., (2022) menyebutkan ada banyak alasan di luar integritas penelitian dan reproduktifitas yang harus dipertimbangkan mengapa prinsip FAIR harus didorong dan dianut. Bagi peneliti, baik yang memiliki atau menghasilkan data maupun yang menggunakan kembali data yang diberikan oleh orang lain, kemampuan untuk menemukan, mengambil, dan menggunakan kembali data mempermudah pekerjaan dan akan meningkatkan visibilitas data sehingga nilainya meningkat. Selain itu, FAIR memungkinkan integrasi data yang lebih mudah di dalam dan lintas disiplin ilmu, mendukung upaya penelitian multidisiplin dan interdisiplin di seluruh dunia untuk mengatasi tantangan global seperti perubahan iklim, keadaan darurat kesehatan, atau realisasi tujuan pembangunan berkelanjutan.


Salah satu sistem pengelolaan data yang menerapkan prinsip FAIR baik terhadap data maupun metadata adalah Dataverse (<https://dataverse.org/>). Sistem ini merupakan platform penyedia pengelolaan data yang digunakan untuk deposit (penyimpanan), preservasi, penyebarluasan, berbagi, analisis, dan penggunaan kembali data. Sistem ini support terhadap beberapa standar metadata seperti Dublin Core, Data Documentation Initiative (DDI), DataCite, OpenAIRE, dsb. Berikut penerapan metadata dalam sistem Dataverse:


*Asterisks indicate required fields


Citation Metadata 


Title * 


Author * 


Name *  1) Family Name, Given Name or 2) Orga


Affiliation  Organization XYZ


Identifier Type  Select...


Identifier 

Point of Contact * 


Name  1) FamilyName, GivenName or 2) Organ


Affiliation  Organization XYZ


E-mail *  name@email.xyz


Description * 


This field supports only certain HTML tags.


Text * 


Date  YYYY-MM-DD


Subject * 


Keyword 


Term 


Controlled Vocabulary Name 


Controlled Vocabulary URL  https://


Related Publication 


Citation 


Identifier Type  Select...

Identifier 

URL  https://

Notes 

Depositor  1) FamilyName, GivenName or 2) Organization

Deposit Date  YYYY-MM-DD

D. Rangkuman

Metadata adalah data tentang data yang berupa informasi terstruktur yang terkait dengan objek tertentu dengan tujuan untuk penemuan, deskripsi, penggunaan, pengelolaan, dan pelestarian. Metadata lazim digunakan sebagai media dalam pendeskripsian sumber daya untuk menghasilkan katalog. metadata terdiri dari tiga jenis, yaitu metadata deskriptif, administratif, dan struktural. Standar metadata yang kerap digunakan dalam sistem pengelolaan data adalah DublinCore dan DataCite. Penerapan metadata pada pengelolaan data harus menerapkan prinsip FAIR.

E. Evaluasi

1. Jelaskan yang dimaksud dengan metadata
2. Sebutkan dan jelaskan prinsip-prinsip metadata
3. Jelaskan standar metadata yang biasa digunakan dalam pengelolaan data!

MATERI POKOK 5:

DOKUMENTASI DATA

Indikator Hasil Belajar:

Peserta mampu:

1. Memahami konsep dasar dokumentasi

A. Deskripsi dan atau informasi teknis dari suatu data

Dokumentasi data bertujuan untuk memperkenalkan data Anda, memberikan deskripsi yang detail tentang atribut utamanya, dan mengontekstualisasikannya agar orang lain memahami apa yang Anda lakukan sehingga mereka dapat menduplikasi/mereplikasi prosesnya. Dalam dokumentasi, Anda harus menyertakan informasi yang menjelaskan tentang apa yang dilakukan, bagaimana hal itu dilakukan, mengapa hal itu dilakukan, kapan dilakukan, di mana dilakukan, dan siapa yang melakukan pekerjaan tersebut. Beberapa contoh dokumentasi data adalah kamus data, file readme, dan metadata.

Pendokumentasian data dapat menggunakan berbagai format, meskipun semua format kurang lebih memiliki konten yang serupa. Semua bentuk dokumentasi harus menyertakan informasi dasar tentang data. Hal ini agar dapat diinterpretasikan dan digunakan kembali baik oleh Anda sendiri maupun untuk orang lain, baik di masa kini maupun di masa mendatang. Format dokumentasi data yang paling umum digunakan adalah File README. File README berisi informasi penting tentang file data, termasuk mengenai informasi kutipan, struktur organisasi file, definisi variabel, informasi metodologis, kode, informasi pengumpulan data, software yang digunakan dan versinya, informasi lisensi, dsb. File README seringkali dibuat dalam format file .txt.

Praktik terbaik dalam membuat file README adalah sebagai berikut. Namun tentunya Anda dapat menyesuaikan dengan kebutuhan masing-masing

sesuai dengan data yang Anda miliki.

Tabel 3. File README

Informasi Umum	<ul style="list-style-type: none">• 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	<ul style="list-style-type: none">• 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	<ul style="list-style-type: none">• Informasi Lisensi/Pembatasan• Publikasi/dataset terkait (URL)• Pengutipan yang direkomendasikan (misal APA, MLA, IEEE, dsb.)

Informasi Metodologi	<ul style="list-style-type: none">• 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• Prosedur jaminan kualitas
----------------------	--

Sumber: <https://libguides.nus.edu.sg/rdm>

B. Rangkuman

Dokumentasi bertujuan untuk memperkenalkan data, memberikan deskripsi yang detail tentang atribut utamanya, dan mengontekstualisasikannya agar orang lain dapat memahami sehingga mereka dapat menduplikasi prosesnya. Dalam dokumentasi perlu menyertakan informasi yang menjelaskan tentang apa yang dilakukan, bagaimana hal itu dilakukan, mengapa hal itu dilakukan, kapan dilakukan, di mana dilakukan, dan siapa yang melakukan pekerjaan tersebut. Format dokumentasi data yang paling umum digunakan adalah File README. File README berisi informasi penting tentang file data, termasuk mengenai informasi kutipan, struktur organisasi file, definisi variabel, informasi metodologis, kode, informasi pengumpulan data, software yang digunakan dan versinya, informasi lisensi, dsb.

C. Evaluasi

1. Jelaskan pengertian dokumentasi data!
2. Buatlah contoh sederhana untuk dokumentasi data!

MATERI POKOK 6:

ETIKA DAN INTEGRITAS DATA

Indikator Hasil Belajar:

Peserta mampu:

Menerapkan etika data: lisensi, hak cipta/copyright, kode etik pengambilan data, data security;

A. Definisi etika data

Dalam proses pengelolaan data sejak pengumpulan, penyimpanan, analisis, visualisasi, hingga penggunaan kembali tentunya ada beberapa aspek yang perlu diperhatikan dari sisi etika. Dari sisi pengumpulan data, apakah data yang dikumpulkan memang sudah sesuai prosedur dan standar etika yang ada? Dari sisi konten data, apakah data wajar untuk dapat dikelola, atau justru memiliki unsur-unsur yang sensitif jika digunakan, seperti adanya informasi pribadi dari seseorang atau informasi mengenai spesimen yang langka. Untuk itu maka diperlukan rambu-rambu atau norma agar data yang digunakan dan dikelola dapat dipertanggungjawabkan.

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. Etika data adalah segala hal tentang penggunaan data yang bertanggung jawab dan berkelanjutan. Etika data merujuk dan mematuhi prinsip dan nilai yang menjadi dasar hukum hak asasi manusia dan perlindungan data pribadi. Hal ini juga berkaitan dengan transparansi yang jujur dan bertanggungjawab dalam pengelolaan data (Tranberg, et al., 2018).

Lebih lanjut lagi, Shukla et al., (2022) membagi etika data menjadi tiga bagian, yaitu etika personal, etika umum, dan etika profesional. Etika

personal adalah tentang mengikuti norma-norma masyarakat yang baik, mengutamakan kepentingan orang lain, menerapkan kebajikan, dan melakukan pengorbanan yang dilakukan seseorang untuk perbaikan diri sendiri, keluarga, dan masyarakat pada umumnya dalam hal memperlakukan segala bentuk data. Sementara itu etika umum, atau biasa juga disebut sebagai etika sosial berawal dari etika personal lalu muncul ke permukaan masyarakat dikarenakan adanya pembatasan-pembatasan yang hadir di masyarakat. Etika ini biasanya dapat tertulis (mis. konstitusi negara) dan tidak tertulis (norma masyarakat). Sedangkan etika profesional adalah etika umum bagi individu yang terlibat dalam profesi tertentu. Hal ini berlaku untuk semua orang yang terlibat dalam profesi tertentu berdasarkan peran mereka yang terkait dengan organisasi atau profesi tertentu.

Dalam konteks ilmu data, data dianalisis secara ilmiah. Oleh karena itu, informasi penting tentang seseorang atau masyarakat berada di data saintis. Jika tidak dilindungi, data dapat disalahgunakan untuk kepentingan lain yang dapat merugikan individu lain.

B. Prinsip-prinsip etika data dalam pengelolaan dan pemanfaatan data

Tranberg, et al., (2018) memaparkan prinsip etika data yang terbagi menjadi lima, yaitu:

- 1) Mengutamakan manusia: Kepentingan manusia selalu diutamakan daripada kepentingan institusional dan komersial. Manusia bukanlah proses komputer atau perangkat lunak, tetapi unik dengan adanya empati, memiliki hak terhadap nasib sendiri, tidak dapat diprediksi, memiliki intuisi dan kreativitas sehingga memiliki status lebih tinggi daripada mesin.
- 2) Pengawasan data individu: Setiap individu memiliki kendali utama atas penggunaan data mereka, dalam konteks di mana datanya diproses dan bagaimana data itu diaktifkan.
- 3) Transparansi: Setiap aktivitas pemrosesan data dan keputusan harus dapat dimengerti oleh setiap individu. Aktivitas ini harus benar-benar transparan dan dapat dijelaskan. Tujuan dan kepentingan pemrosesan

data harus dipahami dengan jelas oleh individu dalam hal memahami risiko sosial, etika, dan kemasyarakatan.

- 4) Akuntabilitas: Akuntabilitas adalah bagian integral dari semua aspek pemrosesan data sebagai upaya untuk mengurangi risiko bagi individu.
- 5) Kesetaraan: Pemrosesan data dilakukan secara demokratis dengan berdasar pada pengawasan namun tanpa adanya diskriminasi atau stigmatisasi, misalnya karena kondisi ekonomi, sosial, budaya, atau kesehatan.

Sementara itu dalam konteks perlindungan data pribadi, Undang-undang Nomor 27 tahun 2022 dalam pasal 16 ayat 2 menyatakan pemrosesan data pribadi setidaknya mengikuti prinsip sebagai berikut:

- a. Data Pribadi dilakukan secara terbatas dan spesifik, sah secara hukum, dan transparan;
- b. pemrosesan Data Pribadi dilakukan sesuai dengan tujuannya;
- c. pemrosesan Data Pribadi dilakukan dengan menjamin hak Subjek Data Pribadi;
- d. pemrosesan Data Pribadi dilakukan secara akurat, lengkap, tidak diadaptasi, mutakhir, dan dapat dipertanggung jawabkan;
- e. pemrosesan Data Pribadi dilakukan dengan melindungi keamanan Data Pribadi dari pengaksesan yang tidak sah, pengungkapan yang tidak sah, pengubahan yang tidak sah, penyalahgunaan, perusakan, dan/atau penghilangan Data Pribadi;
- f. pemrosesan Data Pribadi dilakukan dengan memberitahukan tujuan dan aktivitas pemrosesan, serta kegagalan Pelindungan Data Pribadi;
- g. Data Pribadi dimusnahkan dan/atau dihapus setelah masa retensi berakhir atau berdasarkan permintaan Subjek Data Pribadi, kecuali ditentukan lain oleh peraturan perundang-undangan; dan
- h. pemrosesan Data Pribadi dilakukan secara bertanggung jelas.

C. Regulasi terkait etika data

Pada pembahasan sebelumnya dijelaskan bahwa permasalahan etika data

sangat berkaitan erat dengan data yang bersifat personal. Pihak pengguna data harus menerapkan prinsip-prinsip etika dalam menggunakan dan mengelola data yang dimiliki, terutama yang menyangkut konten pribadi. Permasalahan ini pula yang menjadikan pembahasan mengenai perlindungan terhadap data pribadi memerlukan regulasi yang jelas.

Adanya regulasi yang mengatur penggunaan data tidak hanya dalam konteks nasional, melainkan juga internasional. Di Indonesia, regulasi ini dituangkan dalam Undang-undang Nomor 27 tahun 2022 tentang Perlindungan Data Pribadi. Pada pembahasan sebelumnya telah dijabarkan mengenai pemrosesan data pribadi berdasarkan prinsip perlindungan data pribadi dalam pasal 16 ayat 2. Selain itu pasal 47 menambahkan bahwa pengendali data pribadi harus dapat mempertanggungjawabkan dalam setiap pemrosesan data dalam hal memenuhi kewajiban pelaksanaan prinsip perlindungan data pribadi.

Dalam konteks internasional, salah satu contoh regulasi yang mengatur etika data adalah General Data Protection Regulation (GDPR) yang dibuat oleh Uni Eropa. Regulasi ini menetapkan peraturan yang berkaitan dengan perlindungan orang perseorangan yang berhubungan dengan pemrosesan data pribadi dan peraturan yang berkaitan dengan gerakan bebas data pribadi. Regulasi ini melindungi hak dasar dan kebebasan individu dan khususnya hak mereka atas perlindungan data pribadi. Pergerakan bebas data pribadi di dalam Uni Eropa tidak akan dibatasi atau dilarang karena alasan yang terkait dengan perlindungan orang perseorangan sehubungan dengan pemrosesan data pribadi.

D. Ragam Lisensi Data

Pada era digitalisasi seperti saat ini tentunya kita dihadapi oleh semakin masifnya data yang beredar. Data yang telah beredar sangat memungkinkan bersifat terbuka. Namun pernahkah Anda berpikir ketika ingin menggunakan data tersebut, bagaimana kita dapat memperoleh izin untuk menggunakannya? Bagaimana syarat dan ketentuan dari pemilik data terkait akses dan penggunaan kembali? Mungkin jika ada informasi mengenai

pemilik data kita akan mudah dalam berkomunikasi, namun bagaimana jika informasi tersebut tidak ada? Salah satu cara yang dinilai efektif dan efisien adalah melalui pencantuman lisensi terhadap data. Lisensi data adalah pengaturan hukum 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).

Salah satu lisensi data adalah lisensi Open Data Commons dari Open Knowledge Foundation. Lisensi ini terdiri dari tiga jenis, yaitu:

1. Open Data Commons Open Database License (ODbL) — “Attribution Share-Alike for data/databases”

Penggunaan lisensi ini mengikuti format berikut:

This {DATA(BASE)-NAME} is made available under the Open Database License: <http://opendatacommons.org/licenses/odbl/1.0/>. Any rights in individual contents of the database are licensed under the Database Contents License: <http://opendatacommons.org/licenses/dbcl/1.0/>

- Anda bebas untuk berbagi (menyalin, mendistribusikan, dan menggunakan database).
- Anda bebas untuk membuat (menghasilkan karya dari database).
- Anda bebas untuk mengadaptasi (memodifikasi, mengubah, dan membangun database).

Selama Anda:

Atribusi: Anda harus memberi penjelasan setiap penggunaan database secara publik, atau karya yang dihasilkan dari database, dengan cara yang ditentukan dalam ODbL. Untuk setiap penggunaan atau pendistribusian ulang database, atau karya yang dihasilkan darinya, Anda harus menjelaskan kepada orang lain lisensi database tersebut dan menjaga keutuhan semua pemberitahuan pada database yang asli.

Berbagi-Serupa: Jika Anda secara publik menggunakan versi yang diadaptasi dari database ini, atau karya yang dihasilkan dari database yang diadaptasi, Anda juga harus menawarkan database yang diadaptasi tersebut di bawah ODbL.

Tetap terbuka: Jika Anda mendistribusikan ulang database, atau versi

yang diadaptasi darinya, maka Anda dapat menggunakan tindakan teknologi yang membatasi pekerjaan (seperti DRM) selama Anda juga mendistribusikan ulang versi tanpa tindakan tersebut.

2. Open Data Commons Attribution License — “Attribution for data/databases”

Penggunaan lisensi ini mengikuti format berikut:

This {DATA(BASE)-NAME} is made available under the Open Data Commons Attribution License: http://opendatacommons.org/licenses/by/1.0/

- Anda bebas untuk berbagi (menyalin, mendistribusikan, dan menggunakan database).
- Anda bebas untuk membuat (menghasilkan karya dari database).
- Anda bebas untuk mengadaptasi (memodifikasi, mengubah, dan membangun database).

Selama Anda memberi penjelasan setiap penggunaan database secara publik, atau karya yang dihasilkan dari database, dengan cara yang ditentukan dalam lisensi. Untuk setiap penggunaan atau pendistribusian ulang database, atau karya yang dihasilkan darinya, Anda harus menjelaskan kepada orang lain lisensi database tersebut dan menjaga keutuhan semua pemberitahuan pada database asal.

3. Open Data Commons Public Domain Dedication and License (PDDL) — “Public Domain for data/databases”

Penggunaan lisensi ini mengikuti format berikut:


This {DATA(BASE)} is made available under the Public Domain Dedication and License v1.0 whose full text can be found at: http://opendatacommons.org/licenses/pddl/1.0/
--



- Anda bebas untuk berbagi (menyalin, mendistribusikan, dan menggunakan database).
- Anda bebas untuk membuat (menghasilkan karya dari database).
- Anda bebas untuk mengadaptasi (memodifikasi, mengubah, dan membangun database).



Lisensi PDDL tidak memberlakukan batasan pada penggunaan database berlisensi PDDL.

Selain lisensi Open Data Commons, saat ini lisensi data yang paling lazim dan banyak digunakan adalah lisensi dari Creative Commons. Penggunaan lisensi ini bergantung pada kebutuhan dan keinginan dari masing-masing pemilik/pembuat data. Lisensi dan CC0 tidak dapat dicabut. Ini berarti sekali pemilik data menerapkan lisensi CC untuk sumber dayanya, siapa pun yang menerimanya dapat mengandalkan lisensi tersebut selama sumber daya tersebut dilindungi oleh hak cipta, bahkan jika pemilik kemudian berhenti mendistribusikannya. Creative Commons menyediakan setidaknya enam jenis lisensi dan satu domain publik.

Tabel 4. Jenis Lisensi

No.	Jenis Lisensi	Deskripsi
1	 <p>CC BY</p>	<p>Lisensi ini memungkinkan pengguna untuk mendistribusikan, mencampur ulang, mengadaptasi, dan membuat materi dalam media atau format apa pun, selama atribusi diberikan kepada pencipta. Lisensi ini memungkinkan untuk penggunaan komersial.</p> <p>CC BY mencakup elemen-elemen berikut:</p> <p>BY – Kredit harus diberikan kepada pencipta</p>

2	 <p>CC BY-SA</p>	<p>Lisensi ini memungkinkan pengguna untuk mendistribusikan, mencampur ulang, mengadaptasi, dan membuat materi dalam media atau format apa pun, selama atribusi diberikan kepada pencipta. Lisensi ini memungkinkan untuk penggunaan komersial. Jika pengguna mencampur ulang, mengadaptasi, atau membuat berdasarkan materi tersebut, pengguna harus melisensikan materi yang telah dimodifikasi tersebut dengan ketentuan yang sama.</p> <p>CC BY-SA mencakup elemen-elemen berikut:</p> <p>BY – Kredit harus diberikan kepada pencipta</p> <p>SA – Adaptasi harus dibagi dengan ketentuan yang sama</p>
3	 <p>CC BY-NC</p>	<p>Lisensi ini memungkinkan pengguna untuk mendistribusikan, mencampur ulang, mengadaptasi, dan membuat materi dalam media atau format apa pun hanya untuk tujuan nonkomersial dan hanya selama atribusi diberikan kepada pencipta.</p> <p>CC BY-NC mencakup elemen-elemen berikut:</p> <p>BY – Kredit harus diberikan kepada pencipta</p> <p>NC – Hanya penggunaan karya nonkomersial yang telah diizinkan</p>

4	 <p>CC BY-NC-SA</p>	<p>Lisensi ini memungkinkan pengguna untuk mendistribusikan, mencampur ulang, mengadaptasi, dan membangun materi dalam media atau format apa pun hanya untuk tujuan nonkomersial dan hanya selama atribusi diberikan kepada pencipta. Jika pengguna mencampur ulang, mengadaptasi, atau membuat berdasarkan materi tersebut, pengguna harus melisensikan materi yang dimodifikasi tersebut dengan ketentuan yang sama.</p> <p>CC BY-NC-SA mencakup elemen berikut:</p> <p>BY – Kredit harus diberikan kepada pencipta</p> <p>NC – Hanya penggunaan karya nonkomersial yang diizinkan</p> <p>SA – Adaptasi harus dibagi dengan persyaratan yang sama</p>
5	 <p>CC BY-ND</p>	<p>Lisensi ini memungkinkan pengguna untuk menyalin dan mendistribusikan materi dalam media atau format apa pun hanya dalam bentuk yang tidak diadaptasi dan hanya selama atribusi diberikan kepada pencipta. Lisensi ini memungkinkan untuk penggunaan komersial.</p> <p>CC BY-ND mencakup elemen-elemen berikut:</p> <p>BY – Kredit harus diberikan kepada pencipta</p> <p>ND – Tidak ada turunan atau adaptasi dari karya yang diizinkan</p>

6	 <p>CC BY-NC-ND</p>	<p>Lisensi ini memungkinkan pengguna untuk menyalin dan mendistribusikan materi dalam media atau format apa pun hanya dalam bentuk yang tidak diadaptasi, hanya untuk tujuan nonkomersial, dan hanya selama atribusi diberikan kepada pencipta.</p> <p>CC BY-NC-ND mencakup elemen-elemen berikut:</p> <p>BY – Kredit harus diberikan kepada pencipta</p> <p>NC – Hanya penggunaan karya nonkomersial yang diizinkan</p> <p>ND – Tidak ada turunan atau adaptasi dari karya yang diizinkan</p>
7	 <p>CC0 (CC Zero)</p>	<p>CC0 adalah alat dedikasi publik yang memungkinkan pencipta menyerahkan hak ciptanya dan memasukkan karya mereka ke dalam domain publik di seluruh dunia. CC0 memungkinkan pengguna ulang untuk mendistribusikan, mencampur ulang, mengadaptasi, dan membangun materi dalam media atau format apa pun tanpa syarat.</p>

Sumber: <https://creativecommons.org/about/cclicenses/>

E. Rangkuman

Etika data adalah segala hal tentang penggunaan data yang bertanggung jawab dan berkelanjutan. Etika data merujuk dan mematuhi prinsip dan nilai yang menjadi dasar hukum hak asasi manusia dan perlindungan data pribadi. Etika data menjadi tiga bagian, yaitu etika personal, etika umum, dan etika profesional. Prinsip etika data yang terbagi menjadi lima, yaitu Mengutamakan manusia, Pengawasan data individu, Transparansi, Akuntabilitas, dan Kesetaraan. Di Indonesia, regulasi ini dituangkan dalam Undang-undang Nomor 27 tahun 2022 tentang Perlindungan Data Pribadi. Sedangkan di lingkup internasional, regulasi yang mengatur etika data adalah General Data Protection Regulation (GDPR) yang dibuat oleh Uni Eropa. Sementara itu, lisensi data adalah pengaturan hukum antara pemilik/pembuat data dan pengguna atau tempat penyimpanan data untuk menentukan apa saja yang dapat dilakukan pengguna dengan data tersebut. Salah satu lisensi yang lazim digunakan adalah lisensi Creative Commons yang terdiri dari tujuh lisensi.

F. Evaluasi

1. Jelaskan definisi etika data!
2. Jelaskan prinsip dan regulasi terkait etika data!
3. Sebutkan dan jelaskan jenis lisensi data!

MATERI POKOK 7:

SENSITIVITAS DATA

Indikator Hasil Belajar:

Peserta mampu:

Memahami prinsip-prinsip perlindungan data sesuai aturan yang berlaku.

A. Pengklasifikasian data berdasarkan tingkat sensitivitasnya

Data sensitif adalah data yang berisi informasi yang harus dijaga keamanannya agar tidak terjadi kebocoran data yang berujung pada penyalahgunaan. Data sensitif dapat berupa data Kartu Tanda Penduduk (KTP), data perbankan (nomor rekening, 3 digit nomor cvv kartu kredit), data kesehatan (rekam medis, data biometrik), data keluarga (anak, suami, istri), agama, orientasi seksual dan data sensitif lainnya. Contoh tersebut merupakan data sensitif yang berkaitan dengan data pribadi.

Menurut Undang - Undang Republik Indonesia Nomor 27 tahun 2022 Tentang Perlindungan Data Pribadi, didefinisikan bahwa data pribadi adalah data tentang orang perseorangan yang teridentifikasi atau dapat diidentifikasi secara tersendiri atau dikombinasi dengan informasi lainnya baik secara langsung maupun tidak langsung melalui sistem elektronik atau nonelektronik. Data pribadi tersebut dibagi menjadi dua sifat yaitu spesifik dan umum. Data pribadi yang bersifat spesifik antara lain data dan informasi kesehatan, data biometrik, data genetika, catatan kejahatan, data anak, data keterangan pribadi, dan/ atau data lainnya sesuai dengan ketentuan peraturan. Sedangkan data pribadi yang bersifat umum meliputi nama lengkap, jenis kelamin, agama, status perkawinan, dan/ atau data pribadi yang dikombinasikan mengidentifikasi seseorang. Kedua jenis data tersebut perlu dilindungi mengingat data pribadi mengandung data-data sensitif yang rentan terhadap penyalahgunaan.

Data sensitif tidak hanya mengenai data pribadi. Beberapa data sama sekali tidak terkait ataupun merujuk ke individu manapun (Rumbold & Pierscioneck, 2018), namun data tersebut dapat digunakan untuk mengidentifikasi seseorang. *United Kingdom Anonymisation Network* (UKAN) mengklasifikasikan data menjadi empat jenis seperti terlihat pada tabel 5.

Tabel 5. Klasifikasi data menurut UKAN

<i>About people</i>	<i>Non-Identifiable data</i>	<i>Identifiable data</i>
<i>Yes</i>	<i>Anonymised Data</i>	<i>Primary Personal data</i>
<i>No</i>	<i>Apersonal Data</i>	<i>Secondary Personal Data</i>

Klasifikasi tersebut menekankan fakta bahwa data yang bukan tentang orang tetap bisa menjadi data pribadi. Contohnya yaitu data mengenai alamat, yang tidak berisi data mengenai individu, namun dapat membantu mengidentifikasi seseorang. Namun demikian, data mengenai alamat penyedia internet dinamis merupakan data yang mungkin dapat diidentifikasi tetapi mungkin juga tidak dapat diidentifikasi, tergantung pada data lain yang terkait dengannya.

Dari klasifikasi data pada tabel 5, data dikelompokkan ke dalam 6 (enam) kategori, yaitu data yang tidak terkait manusia/seseorang; interaksi manusia/mesin; demografi penduduk, perilaku, pemikiran/opini; kesehatan; serta karakteristik manusia yang mudah terlihat, baik yang dilindungi maupun tidak (Rumbold & Pierscioneck, 2018), seperti terlihat pada tabel 6.

Tabel 6 menggambarkan rincian potensi spektrum sensitivitas untuk subkategori data tertentu. Angka-angka dalam sel mengacu pada frekuensi relatif data yang akan masuk ke bagian spektrum untuk kategori data tersebut, contohnya pekerjaan jarang masuk ke dalam kategori data paling sensitif. Gambaran visual tabel 6 menjelaskan bahwa sensitivitas meningkat dari 0 hingga 10 dengan warna dari hijau ke merah. Frekuensi relatif dimana data akan muncul di bagian mana pun dari spektrum sensitivitas diberikan dengan angka 1 sampai dengan 4.

Tabel 6. Spektrum Sensitivitas Data

	10	9	8	7	6	5	4	3	2	1	0
A: Data not related to any human being											
Relating to objects											4
Anonymised data related to persons			1	1	1	1	2	3	3	4	3
B: Human/machine interaction											
Recordings of human/machine interaction						2	3	3	3	3	
Location data that act as proxy for human location			3	3	3	3	3	2			
C: Human demographics, behaviour, thoughts & opinions											
Purchasing habits			1	2	3	3	3	3	3	3	
Income			2	3	3	3	3	3			
Occupation	1	2	3	3	3	3	3	3			
Social class			3	3	3						
Address			2	3	3	3	3				
Location		1	2	3	3	3	3	3			
Opinions				3	3	3	3	3	3	3	
Religious or political beliefs			3	3	3	3	3				
Lifestyle or wellness data		3	3	3	3	3	3	3			
Sexual orientation			3	3	3	3	3				
Transgender status		3	3	3	3	2					
D: Readily apparent non-protected											
Facial images (non-processed)				3	3	3	3	3	2		
Body images (non-nude)			3	3	3	3	3	3	2		
Body images (nude)		3	3	3							
Any traits processed for biometrics		3	3	3							
E: Readily apparent protected											
Sex							3	3	3		
Age				3	3	3	3	3			
Pregnancy		3	3	3	3	3	3				
Race			3	3	3	3	3				
Ethnic group			3	3	3	3	3				
F: Medical or health data											
Wellness data					3	3	3	3			
Lifestyle data		3	3	3	3	3	3	3			
Genetic data	3	3	4	3	3						
Diagnoses		3	3	3	3						
Highly sensitive diagnoses	4	3									

Penilaian tersebut sangatlah kontekstual dan perlu memperhatikan hal berikut:

- risiko dengan data anonim yang berkaitan dengan risiko identifikasi ulang (ditambah sensitivitas data yang mendasarinya)
- sensitivitas data yang berkaitan dengan pendapat, keyakinan atau orientasi seksual akan berpotensi bergeser ke kiri dalam masyarakat otoriter atau teokratis

B. Klirens etik penelitian pada domain keilmuan¹⁶

Undang-Undang Nomor 11 Tahun 2019 tentang Sistem Nasional Ilmu Pengetahuan dan Teknologi Pasal 39 mengamanatkan bahwa semua kegiatan riset harus dilakukan sesuai dengan kode etik bidang ilmu. Untuk

¹⁶ <https://klirensetik.brin.go.id/prosedur>

menerapkan kode etik tersebut, dibentuklah Komisi Etik yang bertugas menelaah dan menetapkan kelayakan etik.

Klirens Etik Riset merupakan instrumen yang digunakan untuk mengukur keberterimaan suatu rangkaian proses riset secara etik. Persetujuan klirens etik riset dari Komisi Etik harus diperoleh sebelum riset dimulai. Klirens etik riset merupakan acuan bagi periset dalam menjunjung tinggi nilai integritas, kejujuran, dan keadilan dalam melakukan riset. Pemahaman atas klirens etik riset sangat diperlukan agar tidak menemui masalah dalam menjalankan riset dan mempublikasikan hasil risetnya.

Badan Riset dan Inovasi Nasional (BRIN) sebagai lembaga pemerintah yang berada di bawah dan bertanggung jawab kepada Presiden dalam menyelenggarakan penelitian, pengembangan, pengkajian dan penerapan, serta invensi dan inovasi, penyelenggaraan ketenaganukliran, dan penyelenggaraan keantariksaan yang terintegrasi, memberikan fasilitasi untuk menilai keberterimaan secara etik dari suatu proposal riset yang akan dilakukan oleh periset di dalam maupun di luar BRIN termasuk pihak asing yaitu kelembagaan ilmu pengetahuan dan teknologi asing dan/atau orang asing yang akan melakukan riset di Indonesia.

Klirens etik untuk penelitian yang difasilitasi oleh BRIN terdiri dari bidang ilmu sebagai berikut:

1. Kesehatan

Protokol penelitian kesehatan harus mencakup aspek perlindungan terhadap manusia sebagai subjek dan hewan yang terkait dengan penelitian kesehatan manusia. Hal ini meliputi risiko dan manfaat bagi subjek, prosedur dan metode pengambilan spesimen, percobaan eksperimental, dokumen dan proses *informed consent*/persetujuan setelah penjelasan (PSP), serta tambahan perlindungan apabila akan mengikutsertakan subjek rentan (*vulnerable*). Subjek rentan, sebagai contoh adalah lanjut usia, anak, terpidana, orang yang cacat kognitif, ibu hamil, janin, atau orang-orang yang tidak beruntung secara ekonomi atau pendidikan.

Komisi Etik Kesehatan Badan Riset dan Inovasi Nasional (KEK-BRIN)

merupakan Komisi Etik yang memiliki independensi dalam keputusan. KEK BRIN memiliki tugas pokok dan fungsi sebagai berikut:

- a. Melakukan kajian aspek etik dan ilmiah protokol penelitian kesehatan yang mengikutsertakan manusia sebagai subjek dan memanfaatkan hewan coba yang diajukan melalui BRIN;
- b. Memberikan klirens etik (*ethical clearance/ethical approval*);
- c. Melakukan monitoring dan evaluasi terhadap pelaksanaan penelitian yang telah memperoleh klirens etik;
- d. Melakukan sosialisasi pedoman Etik Penelitian Kesehatan baik di lingkungan BRIN maupun di institusi lain;
- e. Menyelenggarakan pelatihan Etik Penelitian Kesehatan baik di lingkungan BRIN maupun di institusi lain;
- f. Pelaksanaan butir-butir tersebut di atas mengacu kepada pedoman internasional dari *The Research Ethics Review Committee (ERC) World Health Organization (WHO)*, Deklarasi Helsinki, "*Standards and operational guidance for ethics review of health-related research with human participants (WHO, 2011)*", "*International Ethical Guidelines for Health Related Research Involving Human*" (CIOMS, 2016), ICH GCP Guideline E6, dan Pedoman Cara Uji Klinik yang Baik di Indonesia (CUKB, 2016) serta Komite Etik Penelitian dan Pengembangan Kesehatan Nasional, Kemkes (KEPPKN, 2021).

Komisi Etik Kesehatan akan mengelompokkan usulan riset yang masuk ke dalam 3 kategori berikut:

- a. **Hijau:** Tidak berisiko (tidak melibatkan subjek manusia atau memanfaatkan hewan coba; menggunakan informasi umum/public domain; menggunakan data sekunder/bahan biologi tersimpan yang tidak terkait dengan bioterrorism dan atau tidak linked dengan subjeknya);
- b. **Kuning:** Berisiko rendah (melibatkan subjek manusia dengan isu "tidak sensitif"; melibatkan subjek manusia atau memanfaatkan hewan coba dengan tindakan yang tidak melebihi risiko minimal);
- c. **Merah:** Berisiko tinggi (melibatkan subjek manusia dengan isu

"sensitif" atau melibatkan subjek rentan; melibatkan subjek manusia atau memanfaatkan hewan coba dengan tindakan invasif maupun non-invasif yang melebihi risiko minimal)

2. Nuklir

Memandang pentingnya pelaksanaan riset tenaga nuklir yang menerapkan aspek keselamatan, keamanan dan tujuan damai, serta dapat menghargai, melindungi dan berlaku adil terhadap pelaku riset/ subjek riset sesuai dengan norma-norma yang berlaku dalam masyarakat maka untuk memastikan hal tersebut diperlukan klirens etik riset sesuai dengan kelayakan etik riset.

Klirens Etik Riset bidang Tenaga Nuklir bertujuan untuk melindungi periset, masyarakat dan lingkungan dari bahaya radiasi dan konsekuensi hukum (dituntut) sebagai akibat melakukan kegiatan atau turut berpartisipasi dalam suatu riset tenaga nuklir.

Semua riset yang melibatkan tenaga nuklir harus memperhatikan filosofi dasar etika riset dan prinsip 3 (tiga) S (*Safety, Security, Safeguards*), yaitu:

- a. Keselamatan Nuklir, dalam pemanfaatan teknologi nuklir perlu dilakukan pencegahan bahaya nuklir terhadap pekerja, masyarakat, dan lingkungan dengan memenuhi persyaratan proteksi radiasi, yaitu: justifikasi, limitasi dan optimisasi. Dalam penerapannya institusi yang memanfaatkan teknologi nuklir antara lain harus memiliki: (a) Petugas keselamatan bersertifikasi; (b) Penilaian resiko, program proteksi dan keselamatan radiasi, serta prosedur yang sudah ditetapkan; (c) Operator alat/ fasilitas, supervisor dan petugas pendukung yang memenuhi kompetensi atau bersertifikasi; (d) Tersedianya APD dan peralatan keselamatan lainnya.
- b. Keamanan nuklir, dengan mencegah upaya pemilikan tidak sah, penguasaan, sabotase, penyalahgunaan, teror terhadap bahan nuklir dan zat radioaktif, termasuk instalasi nuklirnya, dengan

- memenuhi beberapa persyaratan sebagai berikut: (a) Mendapatkan izin otoritas akses terhadap bahan nuklir, zat radioaktif dan fasilitas nuklir; (b) Bersedia mematuhi prosedur kawasan nuklir yang berlaku.
- c. Safeguards, dengan memenuhi protokol safeguards dan tambahan pada perjanjian safeguards yang menjamin bahan nuklir digunakan untuk tujuan damai, dengan memenuhi beberapa persyaratan sebagai berikut: (a) Menginformasikan bahan nuklir, peralatan dual-used, dan prosedur yang digunakan; (b) Mematuhi ketentuan safeguards.
 - d. Menghormati individu (*respect for persons*), kemanfaatan (*beneficence*), dan berkeadilan (*distributive justice*), dengan memenuhi beberapa persyaratan sebagai berikut: (a) Menghormati otonomi; (b) Melindungi subjek riset; (c) Memberikan manfaat untuk masyarakat

Komisi Etik bidang Tenaga Nuklir akan mengelompokkan usulan riset yang masuk ke dalam 3 kategori berikut:

- a. Hijau : risiko rendah (potensi bahaya rendah)
- b. Kuning : risiko sedang (potensi bahaya sedang)
- c. Merah : risiko tinggi (potensi bahaya tinggi)

3. Kimia

Riset kimia telah membuat dampak positif yang signifikan pada dunia di sekitar kita. Namun, aspek dari penemuan bermanfaat tersebut juga dapat digunakan atau memiliki potensi untuk tujuan negatif. Dalam kondisi ini, periset atau ilmuwan kimia harus mampu menjadi pendukung dalam penerapan kimia yang aman dan terjamin untuk memberi manfaat bagi umat manusia dan melestarikan lingkungan untuk generasi mendatang, seperti tertuang dalam “*Global Chemists’ Code of Ethics*” yang disusun oleh ACS *International Activities* pada tahun 2016 [17]. Oleh

¹⁷ The Global Chemists' Code of Ethics, <https://www.acs.org/content/dam/acsorg/global/international/scifreedom/global-chemists-code-of-ethics-fi-2016.pdf>

karena itu, riset dalam kimia ini memperhatikan prinsip-prinsip dasar kode etik, yaitu [¹⁸, ¹⁹, ²⁰]:

- a. Kemanfaatan (*Beneficience*): kewajiban secara etik untuk memaksimalkan manfaat bagi umat manusia dan meminimalkan ancaman bahaya terhadap manusia dan lingkungan. Sehingga, rancangan riset harus jelas dan peneliti yang bertanggung jawab harus mempunyai kompetensi yang sesuai dan dapat melindungi pelaku dan subyek riset dari resiko yang tidak diinginkan.
- b. Keselamatan dan Keamanan (*Safety and Security*). Pelaksana riset harus mendorong aplikasi, penggunaan, dan pengembangan ilmu pengetahuan dan teknologi yang bermanfaat sambil mendorong dan memelihara budaya keselamatan, kesehatan, dan keamanan yang kuat.
- c. Keselamatan bagi manusia yang terlibat dalam riset (*Security of people involved in scientific research*). Dalam riset yang melibatkan manusia sebagai subyek risetnya, periset harus berusaha untuk meminimalkan kerusakan dan risiko serta memaksimalkan manfaat. Periset harus menghormati martabat manusia, privasi, dan otonomi, serta mampu melindungi individu/subyek riset yang memiliki keterbatasan atau kerentanan dari eksploitasi dan bahaya.
- d. Akuntabilitas (*Accountability*). Periset memiliki tanggung jawab untuk memastikan bahwa bahan kimia, peralatan, dan fasilitas dilindungi dari pencurian dan penyelewengan serta tidak digunakan untuk tujuan ilegal, berbahaya, atau merusak. Pelaksana riset harus mengetahui undang-undang dan peraturan yang berlaku yang mengatur pembuatan dan penggunaan bahan kimia, dan melakukan tindakan apabila ada penyalahgunaan bahan kimia, pengetahuan ilmiah, peralatan, dan fasilitas.

¹⁸ CIOMS, International Ethical Guidelines for Biomedical Research Involving Human Subjects

¹⁹ The Hague Ethical Guidelines, <https://www.opcw.org/hague-ethical-guidelines>

²⁰ Pranas Žukauskas, Jolita Vveinhardt, and Regina Andriukaitienė, Research Ethics, DOI: 10.5772/intechopen.70629

Tujuan utama mendapatkan persetujuan klirens etik pada bidang kimia adalah untuk melindungi pelaku riset, subyek riset, serta lingkungan dari bahaya yang dapat timbul sebagai akibat dari riset yang dilakukan.

Periset dapat mengajukan klirens etik riset bidang kimia apabila risetnya termasuk dalam ruang lingkup sebagai berikut:

- a. Riset yang terkait dengan psikotropika.
- b. Riset yang menggunakan senyawa yang termasuk dalam Bahan Kimia Daftar 1, 2, dan 3 Konvensi Senjata Kimia (Schedule 1, 2, dan 3 Konvensi Senjata Kimia).
- c. Riset yang terkait dengan Bahan Kimia Berbahaya (bahan alam atau sintesis).
- d. Riset yang menggunakan bahan dengan kandungan senyawa non-halal dalam hal riset farmasi, pangan, nutrasetikal, dan kosmetik.
- e. Riset yang menggunakan senyawa aditif berbahaya untuk farmasi, pangan, nutrasetikal, dan kosmetik.
- f. Riset yang menggunakan bahan kimia pembentuk kemasan yang berbahaya untuk objek uji.
- g. Riset yang melakukan uji organoleptik.
- h. Riset yang menggunakan bahan bioproses (melibatkan mikroorganisme) yang membahayakan manusia dan/atau lingkungan.
- i. Riset yang menghasilkan hal baru baik berupa senyawa, sintesis, metode atau prosesnya.

Komisi Etik Riset Bidang Kimia akan mengelompokkan usulan riset yang masuk ke dalam 3 kategori berikut:

- a. Hijau: Tidak berisiko (tidak ada keterlibatan manusia/menggunakan data sekunder atau subjek dan/atau isu riset tidak "sensitif")
- b. Kuning: Berisiko (subjek dan/atau isu riset "sensitif")

- c. Merah: Berisiko tinggi (subjek dan/atau isu riset sangat “sensitif”).

Riset yang termasuk dalam klasifikasi Merah adalah riset dalam lingkup:

- Uji organoleptik yang menggunakan bahan aditif yang dilarang dan menggunakan bahan aditif yang diperbolehkan namun melebihi ambang batas yang diperbolehkan.
- Melibatkan psikotropika Golongan 1 berdasarkan Peraturan Menteri Kesehatan No 2 Tahun 2021.
- Bahan senjata kimia Bahan Kimia Daftar 1 Undang-Undang Republik Indonesia No. 9 Tahun 2008.
- Mikroba cemaran yang melebihi ambang batas yang diperbolehkan berdasarkan Peraturan Badan Pengawas Obat dan Makanan No. 13 Tahun 2019.
- Bahan yang dilarang untuk digunakan sebagai bahan baku dalam obat tradisional, sesuai dengan daftar yang dikeluarkan oleh Direktorat Registrasi Obat Tradisional, Suplemen Kesehatan, dan Kosmetik - Badan Pengawas Obat dan Makanan.
- Bahan kemasan pangan yang mengandung timbal, kromium, nikel, cadmium, dan merkuri berdasarkan Peraturan Badan Pengawas Obat dan Makanan No. 20 Tahun 2019.

4. Hewan coba

Semua riset yang menggunakan hewan harus melalui proses Klirens Etik Riset. Klirens Etik Riset ini bertujuan untuk menjamin implementasi konsep kesejahteraan hewan khususnya pada ranah riset terutama riset menggunakan hewan.

Penerapan kesejahteraan hewan pada pemanfaatan hewan baik untuk kepentingan riset, pengujian dan pendidikan selain akan memberikan hasil yang lebih baik karena rendahnya tingkat intervensi stres juga komunitas ilmiah dapat menegaskan kesadaran moral akan perikemanusiaan terhadap hewan serta tetap dapat menjunjung kewajibannya untuk umat manusia demi kemajuan ilmu pengetahuan dan teknologi. Adapun

penerapan kesejahteraan hewan mengikuti prinsip 3R, yaitu :

- a. *Replacement* atau penggantian mengacu kepada metode mensubstitusi hewan dengan program komputer, kultur sel atau Hewan Coba dengan tingkatan sensitifitas (*sentient*) lebih rendah;
- b. *Reduction* atau pengurangan melibatkan strategi menggunakan jumlah hewan minimal tanpa mengurangi validitas data atau berupa pengurangan perlakuan yang menimbulkan sakit dan stres; dan
- c. *Refinement* atau perbaikan berkenaan dengan modifikasi sistem pemeliharaan atau prosedur percobaan untuk meningkatkan kesejahteraan hewan atau meminimalisasi sakit dan stres.

Kesejahteraan hewan menurut Asosiasi Dokter Hewan Amerika (*American Veterinary Medical Association – AVMA*) merupakan kehidupan hewan yang sejahtera diindikasikan dengan parameter ilmiah sesuai konsep lima kebebasan kesejahteraan hewan (*Five Freedom Animal Welfare*). Kelima konsep tersebut antara lain:

- a. bebas dari rasa haus, lapar dan malnutrisi;
- b. bebas dari ketidaknyamanan;
- c. bebas dari rasa sakit, luka dan penyakit;
- d. bebas untuk mengekspresikan perilaku normal dengan menyediakan tempat yang memadai, aktivitas yang sesuai dan koloni; dan
- e. bebas dari rasa takut dan stress dengan menjamin kondisi yang menghindarkan dari penderitaan mental.

Konsep kesejahteraan hewan ini harus diterapkan pada hewan terutama hewan yang telah hidup dengan campur tangan manusia seperti hewan peliharaan, hewan ternak, satwa liar dan hewan laboratorium.

Komisi Etik Riset Hewan Coba BRIN mengelompokkan usulan yang masuk ke dalam tiga kelompok, yaitu:

- a. Hijau: tidak ada resiko (tidak ada intervensi)
- b. Kuning: minimal atau resiko rendah (ada intervensi namun bersifat *non-invasive*)

- c. Merah: resiko tinggi (ada intervensi yang bersifat invasif dan menggunakan metode *euthanasia*)

5. Sosial humaniora

Dalam buku "*International Ethical Guidelines for Biomedical Research Involving Human Subjects*" (CIOMS, 2002) dikatakan bahwa semua riset yang melibatkan manusia tidak boleh melanggar standar etik yang berlaku universal, tetapi juga harus memperhatikan berbagai aspek sosial budaya masyarakat yang diteliti. Oleh karena itu, semua riset yang melibatkan manusia harus melalui proses klirens etik riset.

Tujuan utama mendapatkan persetujuan klirens etik pada bidang ilmu pengetahuan sosial dan kemanusiaan adalah untuk melindungi subjek riset/responden dari bahaya secara fisik (ancaman), psikis (tertekan, penyesalan), sosial (stigma, diasingkan dari masyarakat) dan konsekuensi hukum (dituntut) sebagai akibat turut berpartisipasi dalam suatu riset. Oleh karena itu, semua riset yang melibatkan manusia harus memperhatikan tiga prinsip dasar kode etik (CIOMS, 2002) yaitu:

- a. Menghormati individu (*Respect for persons*), setidaknya ada 2 (dua) etika dasar yang perlu diperhatikan, yaitu (a) Menghormati otonomi (*Respect for autonomy*): menghargai kebebasan seseorang terhadap pilihan sendiri; dan (b) Melindungi subyek riset (*Protection of persons*): melindungi individu/subyek riset yang memiliki keterbatasan atau kerentanan dari eksploitasi dan bahaya
- b. Kemanfaatan (*Beneficence*): kewajiban secara etik untuk memaksimalkan manfaat dan meminimalkan bahaya. Prinsip ini menekankan bahwa semua riset harus bermanfaat bagi masyarakat. Oleh karena itu rancangan riset harus jelas dan peneliti yang bertanggung jawab harus mempunyai kompetensi yang sesuai dan dapat melindungi subyek riset dari resiko yang tidak diinginkan.
- c. Berkeadilan (*Distributive justice*): Keseimbangan antara beban dan manfaat ketika berpartisipasi dalam riset. Prinsip ini menekankan bahwa setiap individu yang berpartisipasi dalam riset harus

diperlakukan sesuai dengan latar belakang dan kondisi masing-masing. Perbedaan perlakuan antara satu individu/kelompok dengan lain dapat dibenarkan bila dapat dipertanggung jawabkan secara moral dan dapat diterima oleh masyarakat.

Periset dapat mengajukan klirens etik riset sosial humaniora apabila risetnya termasuk dalam kriteria berikut:

- a. Semua riset yang melibatkan manusia harus melalui proses klirens etik. Riset yang dikecualikan dari proses klirens etik adalah riset yang menggunakan data sekunder, review literatur atau data yang sudah dipublikasi seperti: (a) Surat kabar, *website*, majalah, laporan publik, pernyataan publik, film, program televisi, pertunjukan di depan publik, pameran di publik, pidato publik; (b) Karya yang telah dipublikasi, sistematik review, review literatur, dan lain-lain; (c) Materi-materi lama (seperti manuskrip, arsip) yang disimpan dan boleh digunakan untuk umum.
- b. Riset yang menggunakan review dari materi-materi yang bersifat konfidensial (catatan kesehatan rumah sakit/klinik kesehatan) harus melalui proses klirens etik (meskipun masih memerlukan ijin dari institusi yang mengeluarkan). Demikian juga riset yang menggunakan informasi dari media tertutup/tidak bersifat umum, seperti review statistik dari suatu lembaga (karyawan, klien, pasien, pengguna, penyedia jasa, catatan pelayanan, catatan keuangan, catatan perusahaan) harus melalui klirens etik.
- c. Studi yang menggunakan metode tambahan, yang berhubungan langsung dengan manusia seperti wawancara, *Focus Group Discussion* (FGD) dan lain-lain, tetap memerlukan proses klirens etik, meskipun metode utamanya menggunakan review materi yang ada di publik.

Untuk riset yang melibatkan manusia, status riset yang memerlukan atau tidak memerlukan proses klirens etik diputuskan oleh Komisi Etik Riset Sosial Humaniora, bukan oleh peneliti atau lembaga lainnya. Komisi Etik

akan mengelompokkan usulan riset yang masuk ke dalam 3 kategori berikut:

- a. Hijau: Tidak berisiko (tidak ada keterlibatan manusia/menggunakan data sekunder)
- b. Kuning: Berisiko rendah (subjek dan isu riset tidak “sensitif”)
- c. Merah: Berisiko tinggi (subjek dan/atau isu riset sangat “sensitif”).

Riset yang termasuk dalam klasifikasi Merah adalah riset yang melibatkan kelompok rentan, yaitu anak-anak, lansia, wanita hamil, orang lemah mental dan intelektual, disabilitas, kelompok LGBT, orang yang pernah mengalami trauma, pengidap HIV, orang dengan ketergantungan obat, orang yang mengalami kesulitan berkomunikasi, narapidana, residivis, dan teroris.

6. Bidang lainnya.

Untuk riset yang tidak termasuk dalam kategori riset sosial humaniora, hewan coba, kesehatan, kimia, dan nuklir namun dalam tim risetnya terdapat Warga Negara Asing (WNA) atau bekerjasama dengan Lembaga Riset Asing, maka wajib mengajukan Izin Riset melalui permohonan Klirens Etik Riset. Permohonan pada kategori ini akan dikenai biaya sesuai dengan tarif PNBPN yang berlaku sebagai berikut:

Tabel 7. Biaya tarif PNPB Klirens Etik untuk mengajukan izin riset

KATEGORI (Category)	Izin Baru (<1 bulan) (New Permit <1 month)	Izin Baru (<6 bulan) (New Permit <6 month)	Izin Baru (6-12 bulan) (New Permit 6-12 month)	Izin Perpanjangan (<1 bulan) (Renewal Permit <1 month)	Izin Perpanjangan (<6 bulan) (Renewal Permit <6 month)	Izin Perpanjangan (6-12 bulan) (Renewal Permit 6-12 month)	Dokumen Perjalanan untuk Spouse & Dependent (Travel Document for Spouse & Dependent)	Perpanjangan Dokumen Perjalanan untuk Spouse & Dependent ((Travel Document for Spouse & Dependent))
Biaya Pendaftaran (Registration Fee)	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
Perguruan Tinggi Asing (foreign College)	1.750.000	2.500.000	5.000.000	750.000	1.250.000	2.500.000	1.000.000	500.000
Lembaga Litbang Asing (foreign R&D Institute)	1.750.000	2.250.000	5.000.000	750.000	1.250.000	2.500.000	1.000.000	500.000
Badan Usaha Asing (Foreign Corporation)	3.000.000	5.000.000	10.000.000	1.750.000	2.500.000	5.000.000	1.000.000	500.000
Orang Asing (Foreigners)	750.000	1.300.000	1.500.000	500.000	650.000	750.000	1.000.000	500.000

Bebas Tarif dapat diberikan oleh Badan Riset dan Inovasi Nasional (BRIN) apabila periset asing bekerjasama dengan periset BRIN. Untuk mengajukan bebas tarif, periset harus menyerahkan kelengkapan dokumen sebagai berikut:

- rencana kegiatan penelitian dan pengembangan;
- surat keterangan rekomendasi atau persetujuan dari lembaga penjamin;
- dokumen kerjasama dengan mitra kerja dari BRIN; dan
- surat permohonan kepada Deputi Bidang Fasilitasi Riset dan Inovasi BRIN paling lambat 7 (tujuh) hari kerja sebelum pelaksanaan.

Semua dokumen diunggah pada website klirensetik.brin.go.id.

C. Anonimisasi atau pseudonimisasi

Data pribadi merupakan data sensitif yang rawan penyalahgunaan. Oleh karena itu, perlu adanya perlindungan terhadap data pribadi secara hukum. Negara Uni Eropa memiliki *General Data Protection Regulation* (GDPR) untuk perlindungan data pribadi warganya yang berada di dalam maupun di luar Uni Eropa, serta yang dikelola pihak mana pun di dalam maupun di luar

wilayah Uni Eropa. Di Indonesia, perlindungan terhadap data pribadi diatur dalam Undang-Undang Nomor 27 Tahun 2022 Tentang Perlindungan Data Pribadi. Menurut undang-undang tersebut, Pelindungan Data Pribadi adalah keseluruhan upaya untuk melindungi Data Pribadi dalam rangkaian pemrosesan Data Pribadi guna menjamin hak konstitusional subjek Data Pribadi.

Salah satu teknik perlindungan data, yaitu anonimisasi atau pseudonimisasi. Anonimisasi merupakan teknik perlindungan data dengan mengubah data pengenal pribadi sehingga tidak dapat diidentifikasi baik secara langsung maupun tidak langsung. Sedangkan pseudonimisasi merupakan metode perlindungan data dengan menggunakan nama samaran atau identifikasi buatan. Walaupun terdapat perbedaan mekanisme perlindungan data, namun kedua cara tersebut sama-sama dapat menyembunyikan maupun menyamarkan data asli sehingga data pribadi tidak terpapar ke pihak yang tidak berhak.

D. Rangkuman

Data sensitif adalah data yang berisi informasi rahasia yang harus dijaga keamanannya agar tidak terjadi kebocoran data yang berujung pada penyalahgunaan. Data sensitif dapat berupa data Kartu Tanda Penduduk (KTP), data perbankan (nomor rekening, 3 digit nomor cvv kartu kredit), data kesehatan (rekam medis, data biometrik), data keluarga (anak, suami, istri), agama, orientasi seksual dan data rahasia lainnya. Contoh tersebut merupakan data sensitif yang berkaitan dengan data pribadi.

Undang-Undang Nomor 11 Tahun 2019 tentang Sistem Nasional Ilmu Pengetahuan dan Teknologi Pasal 39 mengamanatkan bahwa semua kegiatan riset harus dilakukan sesuai dengan kode etik bidang ilmu.

Klirens Etik Riset merupakan instrumen yang digunakan untuk mengukur keberterimaan suatu rangkaian proses riset secara etik. Persetujuan klirens etik riset dari Komisi Etik harus diperoleh sebelum riset dimulai.

E. Evaluasi

1. Sebutkan dan jelaskan Klirens etik untuk penelitian yang difasilitasi oleh BRIN.

DAFTAR PUSTAKA

- Corti L. Eynden V. van den Bishop L. Woollard M. Haaker M. & Summers S. (2019). *Managing and sharing research data : a guide to good practice* (2nd ed.). SAGE Publications.
- Creative Commons. (n.d.). About CC Licenses. Creative Commons. <https://creativecommons.org/about/cclicenses/>
- Deutz, Daniella Bayle, Buss, Mareike C. H., Hansen, Jitka Stilund, Hansen, Karsten Kryger, Kjellmann, Kristian G., Larsen, Asger Væring, Vlachos, Evgenios, & Holmstrand, Katrine Flindt. (2020, June 30). How to FAIR: a website to guide researchers on making research data more FAIR. Zenodo. <https://doi.org/10.5281/zenodo.3712065>
- Engelhardt, C., Biernacka, K., Coffey, A., Cornet, R., Danciu, A., Demchenko, Y., Downes, S., Erdmann, C., Garbuglia, F., Germer, K., Helbig, K., Hellström, M., Hettne, K., Hibbert, D., Jetten, M., Karimova, Y., Kryger Hansen, K., Kuusniemi, M. E., Letizia, V., ... Zhou, B. (2022). *How to be FAIR with your data: A teaching and training handbook for higher education institutions*. <https://doi.org/10.5281/zenodo.6674301>
- Farnel, S., & Shiri, A. (2014). Metadata for Research Data: Current Practices and Trends. *Proceedings of the International Conference on Dublin Core and Metadata Applications*, 74–82.
- GO FAIR. (n.d.). *FAIR Principles*. Retrieved December 12, 2022, from <https://www.go-fair.org/fair-principles/>
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining: concepts and techniques* (Third Ed.). Burlington: Elsevier.
- Hasselbalch, G. (2021). *Data Ethics of Power*. Edward Elgar Publishing.
- Joudrey, D. N., Taylor, A. G., & Wisser, K. M. (2018). *The Organization of Information* (Fourth Edi). Libraries Unlimited.
- <https://klirensetik.brin.go.id/prosedur>. Diakses pada tanggal 12 Desember 2022, 22.45 WIB.
- Ramesh, P., Vivekavardhan, J., & Bharathi, K. (2015). Metadata Diversity, Interoperability and Resource Discovery Issues and Challenges. *DESIDOC Journal of Library and Information Technology*, 35(3), 193–199. <https://doi.org/10.14429/djlit.35.3.8074>
- National Information Standards Organization. (2008). *Metadata Framework*. niso.org; National Information Standards Organization. <https://framework.niso.org/24.html>

- Parham, S. W., Carlson, J., Hswe, P., Westra, B., Whitmire, A. (2016). Using data management plans to explore variability in research data management practices across domains. *International Journal of Digital Curation*, 11(1), 53–67. <https://doi.org/10.2218/ijdc.v11i1.423>
- Research Data Netherlands. (n.d.). Standardised metadata. RDNL - Essentials 4 Data Support. Retrieved December 6, 2022, from <https://datasupport.researchdata.nl/en/start-the-course/iv-harvest-phase/data-archives/standardised-metadata>
- Rumbold, J. M. M., & Pierscionek, B. K. (2018). What are data? A categorization of the data sensitivity spectrum. In Elsevier *In Big Data Research*. Vol. 12., July 2018, 45-49. <https://doi.org/10.1016/j.bdr.2017.11.001>
- Shukla, S., George, J. P., Tiwari, K., & Kureethara, J. V. (2022). Data ethics and challenges. In *SpringerBriefs in Applied Sciences and Technology*. Springer Singapore.
- Smaele, M. de, Dasler, R., Ashton, J., Martínez, I. B., Burger, M., Fenner, M., Habermann, T., Ilik, V., Jacobson, M., Raugh, A., Roi, A. la, Roy, S., Yahia, M., & Zolly, L. (2016). *DataCite Metadata Schema for the Publication and Citation of Research Data (Version 4.3)* (4.3). <https://doi.org/https://doi.org/10.14454/7xq3-zf69>
- The Dublin Core Metadata Initiative. (n.d.). *DCMI Metadata Terms*. Retrieved January 4, 2021, from <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>
- Tranberg, P., Hasselbalch, G., Olsen, B. K., & Byrne, C. S. (2018). *Dataethics: Principles and Guidelines for Companies, Authorities & Organisations* (1st ed.). Spintype.com.