



UNIVERSITAS AIRLANGGA
Excellence with Morality

DATA DAN PUSTAKA (SIP107)

Pengantar: Peranan dan Posisi Data dalam Sains dan Kehidupan
Sehari-hari



Kontrak Kuliah

Silakan mengunduh:

1. Rancangan Pembelajaran Semester (RPS)
2. Kontrak Perkuliahan, dan
3. Referensi untuk mata kuliah ini

..di ELearning HEBAT.



Cakupan Mata Kuliah

Di MK ini, mahasiswa akan belajar:

1. Prinsip-prinsip dasar menerjemahkan informasi deskriptif dalam data
2. Menyajikan (presentasi) data
3. Menggunakan data untuk memperkuat argumentasi

Yang *tidak* diajarkan dalam MK ini:

1. Menarik generalisasi atau mengambil keputusan berdasarkan data (analisis inferensial)
2. Manajemen data

Data yang dimaksud adalah **data kuantitatif** dan **data verba** (hasil wawancara) serta data/informasi yang terkandung dalam **referensi ilmiah**.



Data dan Kehidupan Manusia

- Seiring dengan meningkatnya interaksi manusia dengan teknologi digital, sebagian besar kehidupan manusia dimonitor dan direkam kemudian terkumpul menjadi *digital footprint*.
- Data tersebut terkumpul dalam volume yang besar, yang seringkali disebut dengan *big data*.
- Saat ini, teknologi komputasi juga memungkinkan untuk tidak hanya mengumpulkan data dalam jumlah besar, namun juga menganalisisnya dan membuat profil individual yang amat detail.
- *Big data* adalah sekumpulan set data (*datasets*) dengan volume data sangat besar (jutaan, bahkan milyaran) dalam bentuk digital dan diproses dengan teknologi komputasi.



Big Data

Dicirikan dengan 5 V

Volume

Variety

Velocity

Value

Veracity

Volume : Ukuran data berkapasitas besar

Variety : Jenis data sangat bervariasi, tidak hanya terdiri dari data terstruktur, namun juga data tidak terstruktur

Velocity : Kecepatan transfer data tinggi

Value : Aliran data membutuhkan resource yang besar

Veracity : kualitas dan keakuratan data



4 Tahapan Revolusi Industri

1784

1870

1969

2000

2019

Industri 1.0

Mesin uap, tenaga air, angin, dan, matahari



Industri 2.0

Energi listrik untuk produksi masal



Industri 3.0

Teknologi informasi dan elektronika yang diterapkan pada sistem otomatis produksi



Industri 4.0

Teknologi digital, teknologi wireless dan big data secara massif yang terintegrasi dengan kegiatan manufaktur



Society 5.0

Penerapan ilmu pengetahuan yang berbasis modern (AI, Robot, IoT)





Data dan Kehidupan Manusia

Contoh *big data*:

1. Sekuens genomik (lihat [GISAID](#))
2. Data rekam medik (lihat [Surgisphere](#), *caveat*: Surgisphere memiliki [banyak masalah dan skandal](#))
3. Data percakapan di Twitter dan Facebook
4. Data kependudukan yang bentuknya digital (e-KTP, data vaksinasi dan mobilitas di Peduli Lindungi)
5. Data transaksi dan data pribadi pengguna di *marketplace* digital



Aspek Legal atau Regulasi

- Undang-Undang No. 14 Tahun 2008 tentang Keterbukaan Informasi Publik, yang mewajibkan Badan Publik untuk membuka secara transparan data dan informasi yang mengandung kepentingan publik.
- Undang-Undang No. 11 Tahun 2008 (atau perubahannya UU No. 19 Tahun 2016) mengenai Informasi dan Transaksi Elektronik (ITE), yang mengatur lalu-lintas informasi digital, termasuk data dalam bentuk digital.
- Rancangan Undang-Undang mengenai Perlindungan Data Pribadi, yang melindungi data publik dari berbagai macam risiko pengungkapan.
 - Sayangnya, belum segera disahkan oleh DPR RI, meskipun kebutuhannya mendesak.



Literasi Data

- Data penting peranannya untuk membantu manusia untuk menerjemahkan pola, memahami realitas dan mengambil keputusan. Oleh karena itu, mahasiswa perlu **memiliki literasi data**.
- Seseorang yang memiliki **literasi yang baik** *bukan hanya berarti* terampil dalam membaca dan menulis saja, namun juga **menunjukkan kadar keingintahuan** dalam tingkatan tertentu, **memiliki pemahaman dasar** untuk **mencerap informasi, mengapresiasi**, bahkan mampu **menggunakan informasi** untuk kehidupannya sehari-hari.
- Memiliki **literasi data** artinya mahasiswa mampu **membaca** (menginterpretasi) **hasil analisis/penyajian data** sekaligus **menggunakannya untuk kepentingan sehari-hari** secara bertanggung jawab.



Mengapa Literasi Data Penting?

- Literasi data = Literasi Sains
- Sains modern adalah deskripsi dan prediksi yang sifatnya terpandu oleh data (*data-driven*), bukan berdasarkan intuisi, perasaan, atau *cocokologi*.
- Data membantu manusia menarik kesimpulan yang **paling sedikit** terkontaminasi bias.
- Literasi data juga membantu kehidupan kita sehari-hari. Contohnya:
 - Membantu orangtua untuk menilai perkembangan anak (membaca kurva perkembangan).
 - Memeriksa klaim hasil survei elektabilitas, *exit poll*, atau hitung cepat ketika musim pemilihan umum.
 - Memutuskan berkunjung/tidak berkunjung ke suatu daerah yang mengalami wabah.



Data dan Sains

- Dua kelompok peneliti menganalisis set data yang sama, namun kesimpulannya berbeda. Apakah mungkin?
- Sangat mungkin. Inkonsistensi penafsiran data adalah masalah yang sedang ingin diselesaikan oleh komunitas akademik.
- Problem ini disebut juga sebagai krisis reproducibility, yang disebabkan oleh berbagai masalah. Salah satunya adalah **gagalnya/keengganan peneliti mengungkapkan seluruh tahapan penelitian** yang telah dilakukan.
- Oleh karena itu, **transparansi dalam penggunaan data** adalah keniscayaan, agar temuan penelitian dapat dianggap kredibel/dapat dipercaya.
 - Meskipun secara normatif, transparansi data harus tetap mempertimbangkan aspek etis. Misalnya, peneliti hanya dapat berbagi set data yang anonim (*de-identified dataset*).



UNIVERSITAS AIRLANGGA
Excellence with Morality





Penalaran Statistik

Di MK ini, mahasiswa akan belajar mengenai Statistik Dasar, namun yang ditekankan adalah **penalaran statistik**, bukan sekadar kemampuan berhitung. Artinya, yang lebih penting dari berhitung adalah:

1. Memahami **logika dasar** informasi statistik dan **bagaimana menggunakannya**
2. Mampu **menarik benang merah** dan **mengkombinasikan** antara satu konsep dengan konsep yang lain (misalnya, tendensi sentral dan dispersi)
3. Menjelaskan **proses didapatnya suatu informasi statistik** sehingga mampu menerjemahkan hasil perhitungan statistik secara menyeluruh



Isu-Isu Etis

Prinsipnya, sebagai pemilik atau pengelola data, kita harus peka pada isu-isu mengenai:

1. Keamanan data

- a. Sudahkah data yang kita miliki/kelola disimpan dengan aman sehingga mencegah kemungkinan terjadinya pencurian atau penyalahgunaan data oleh pihak-pihak yang tidak bertanggung jawab?

2. Persetujuan (*consent*)

- a. Sebagai pengelola, pengumpul, atau pengguna data, sudahkah kita meminta persetujuan yang terpandu (*informed consent*) dari pemilik data untuk mengumpulkan atau menyimpan datanya?



Isu-Isu Etis

Prinsipnya, sebagai pemilik atau pengelola data, kita harus peka pada isu-isu mengenai:

3. Anonimitas dan Kerahasiaan

- a. Sebagai pengumpul, penyimpan dan pengguna data, sudahkah kita menjamin bahwa data yang dipercayakan pemiliknya kepada kita dapat dijamin kerahasiaan dan anonimitasnya? Pada saat seperti apa kita dibenarkan mengumpulkan data pribadi yang sifatnya sensitif (e.g. NIK, tanggal lahir, nama ibu kandung, riwayat kesehatan, dll)?

4. Risiko pengungkapan (*disclosure risk*)

- a. Sebagai pengumpul, penyimpan dan pengguna data, apakah pemilik data (yang mempercayakan datanya kepada kita) sudah diberi tahu mengenai risiko pengungkapan yang mungkin terjadi? Apa usaha yang seharusnya dilakukan untuk meminimalisasi risiko ini?



Sesi Diskusi

- Baru-baru ini, terungkap bahwa data perjalanan yang direkam oleh aplikasi Indonesia Health Alert Card (eHAC) dikabarkan bocor ke publik. Bagaimana pendapat anda?
- Studi mengenai efektivitas Ivermectin dalam meringankan gejala COVID-19 diambil dari set data rekam medik yang dikumpulkan oleh perusahaan *health data analytics* bernama Surgisphere, yang memiliki rekam jejak meragukan.
 - Set data yang sama digunakan untuk studi efektivitas *Hydroxychloroquine*, yang belakangan diketahui bahwa set data tersebut dipalsukan.
 - Bagaimana pendapat anda mengenai hal ini?
 - Bagaimana usaha yang mungkin dilakukan agar kasus ini tidak terjadi?



UNIVERSITAS AIRLANGGA
Excellence with Morality

Data Pribadi untuk Algoritma, Etiskah?





Kasus Cambridge Analytica

- Apakah ada yang familiar dengan kasus/skandal Cambridge Analytica (CA) di video *trailer* tersebut?
- Perusahaan ini terbelit masalah hukum di Eropa dan Amerika Serikat karena menyalahgunakan data pengguna Facebook untuk kepentingan kampanye politik.
 - CA adalah konsultan politik Donald Trump dan Partai Konservatif Britania Raya dalam kampanye Brexit.
- Bagaimana pendapat anda apabila pengelola situs media sosial (e.g. Twitter dan Facebook) menggunakan data pribadi (yang secara sukarela anda bagikan) untuk kepentingan pengayaan algoritma tanpa persetujuan anda?
 - Tujuannya untuk membangun profil personal sehingga anda secara otomatis mendapatkan iklan (atau informasi) yang sesuai dengan profil anda.



UNIVERSITAS AIRLANGGA
Excellence with Morality

**Terima Kasih dan
Selamat Menikmati!**