

Data Infrastructure & Data Governance untuk Republik Indonesia

*Fondasi Negara Cerdas 2045: Data, AI, Keamanan, dan
Efisiensi Anggaran*

Onno W. Purbo
onno@indo.net.id
onno@itts.ac.id
twitter onnowpurbo

**Institut Teknologi Tangerang Selatan (ITTS)
OnnoCenter
2025**

Hak Cipta

Karya ini dilisensikan di bawah lisensi **Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)**. Artinya, **siapapun bebas untuk menggunakan, menyalin, membagikan, dan mengadaptasi materi ini, dengan syarat:**

- **Memberikan atribusi yang sesuai** (menyebut sumber asli),
- **Menyertakan lisensi yang sama** jika dimodifikasi atau dikembangkan lebih lanjut.

Lihat detail lisensi di sini <https://creativecommons.org/licenses/by-nd/4.0/legalcode.id>

Disclaimer

Naskah ini disusun sebagai panduan teknis untuk membantu pemerintah Indonesia membangun fondasi **data infrastructure** dan **data governance** yang kuat, efisien, dan berdaulat, dengan mengutamakan open source. Isi dokumen ini merupakan rekomendasi profesional penulis, **bukan kebijakan resmi pemerintah**, sehingga setiap implementasi tetap memerlukan kajian keamanan, regulasi, dan penyesuaian oleh instansi terkait. Seluruh contoh arsitektur, prosedur, dan tools disampaikan untuk tujuan edukasi, dan penulis tidak bertanggung jawab atas risiko operasional atau hukum yang timbul dari penggunaan langsung tanpa adaptasi. Naskah ini diniatkan sebagai kontribusi terbuka agar Indonesia semakin mandiri, aman, dan siap menghadapi transformasi digital nasional.

Daftar Isi

Hak Cipta.....	2
Disclaimer.....	2
Daftar Isi.....	3
Kata Pengantar.....	15
Executive Summary.....	16
Langkah Strategis 1 Tahun Pertama.....	17
Ringkasan 1 Tahun (Tabel Singkat).....	17
1. Tetapkan Standar Nasional Tunggal (3 bulan pertama).....	17
2. Bentuk Kelembagaan Inti Data Nasional (bulan 1–6).....	18
3. Bangun GovCloud Dasar & Node Data Daerah (bulan 1–12).....	18
4. Etalase Dashboard Nasional (bulan 4–12).....	18
5. Aman, Enkripsi, dan Monitoring Nasional (bulan 1–12).....	19
6. Percepatan SDM dan Pelatihan Nasional (bulan 3–12).....	19
7. Integrasi Data Prioritas (bulan 4–12).....	19
8. Rancangan Perpres + Perda Data Governance (bulan 1–12).....	20
Pesan Inti.....	20
BAGIAN I — LANDASAN & URGENSI NASIONAL.....	21
Bab 1 — Mengapa Data Governance Penting untuk Republik Indonesia.....	21
Data Buruk Menimbulkan Biaya Ekonomi yang Besar.....	21
Keputusan Pemerintah Hanya Sebaik Kualitas Datanya.....	22
Contoh Nyata Kegagalan Program Akibat Data Tidak Akurat.....	22
Single Source of Truth (SSOT) adalah Pondasi Negara Digital.....	23
Tabel Pentingnya Data Governance.....	24
Pesan Inti.....	24
Bab 2 — Kerangka Regulasi & Visi Nasional 2045.....	25
Perpres Satu Data Indonesia (SDI).....	25
SPBE — Kerangka Digital Pemerintahan Modern.....	25
UU PDP — Pengamanan Data Pribadi Masyarakat.....	25
GovCloud — Infrastruktur Negara Digital 2045.....	26
Model Kematangan Data Governance (Maturity Model).....	26
Penyelarasan dengan RPJMN & Visi Indonesia Emas 2045.....	27
Tabel Kerangka Kebijakan.....	28
Pesan Inti.....	28
BAGIAN II — ORGANISASI & TATA KELOLA DATA.....	29
Bab 3 — Arsitektur Kelembagaan Data Nasional.....	29
Dewan Data Nasional, CDO, DPO, CIO — “Komando Utama Data Negara”.....	29
Dewan Data Nasional.....	29
CDO (Chief Data Officer).....	29
DPO (Data Protection Officer).....	29
CIO (Chief Information Officer).....	29
Data Owner, Data Steward, Wali Data — “Pelaksana Harian Data Pemerintah”.....	30

Data Owner (Pemilik Data).....	30
Data Steward.....	30
Wali Data.....	30
SOP Pengambilan Keputusan Data-Driven (Decision Framework).....	31
Langkah-langkah Decision Framework yang sederhana dan kritis:.....	31
Tabel Arsitektur Kelembagaan Data.....	32
Pesan Inti.....	32
Bab 4 — Standar Data Nasional & Interoperabilitas.....	33
Metadata Nasional & Kamus Data.....	33
Sistem Data Catalog & Data Lineage.....	34
Data Catalog menyediakan:.....	34
Data Lineage menjawab pertanyaan kritis:.....	34
Format Standar Pertukaran Data Nasional (JSON Schema).....	34
Policy PKS/SOP Antar-Lembaga.....	35
Tabel Standard Data Nasional.....	35
Pesan Inti.....	36
BAGIAN III — INFRASTRUKTUR DATA & TEKNOLOGI PEMERINTAH.....	37
Bab 5 — Arsitektur Infrastruktur Data Indonesia.....	37
Data Warehouse, Data Lake, Data Mesh, Lakehouse.....	37
Data Warehouse — untuk laporan resmi pemerintah.....	37
Data Lake — untuk data besar (big data) dan data mentah.....	37
Data Mesh — pembagian per-domain di dalam K/L besar.....	38
Lakehouse — gabungan terbaik Warehouse + Lake.....	38
Pusat Data Nasional (PDN) & Pusat Data Daerah (PDD).....	38
Pusat Data Nasional (PDN).....	38
Pusat Data Daerah (PDD).....	38
Government Cloud (Hybrid / Multi-cloud).....	39
Hybrid Cloud.....	39
Multi-cloud.....	39
Jaringan, Server, Komputasi CPU/GPU/NPU — Mesin AI Nasional.....	39
Jaringan.....	39
Server (Compute & Storage).....	40
CPU (General Processing).....	40
GPU (Artificial Intelligence Computing).....	40
NPU (Neural Processing Unit, Edge AI).....	40
Tabel Arsitektur Infrastruktur Data.....	41
Pesan Inti.....	41
Bab 6 — Sistem Pertukaran Data Nasional (National Data Exchange).....	42
API Gateway Nasional — Jalan Tol Data Indonesia.....	42
Standar JSON Nasional & Schema Versioning.....	43
Message Bus, ESB, Batch vs Real-Time.....	43
Message Bus — jalur cepat data real-time.....	43
ESB (Enterprise Service Bus) — orkestrator integrasi antar sistem.....	43
Batch vs Real-Time.....	44

Arsitektur Microservices untuk Pemerintah.....	44
Open-source pendukung microservices:.....	45
Tabel Sistem Pertukaran Data Nasional.....	45
Pesan Inti.....	45
Bab 7 — Aplikasi Open Source untuk Infrastruktur Pemerintah.....	46
DATABASE (PostgreSQL, MariaDB, ClickHouse).....	46
PostgreSQL — standar emas untuk data pemerintah.....	46
MariaDB — alternatif cepat dan ringan.....	46
ClickHouse — database analitik super cepat.....	46
ETL/ELT (n8n, Apache NiFi, Airflow).....	47
n8n — sangat mudah untuk pemda.....	47
Apache NiFi — kuat untuk data besar dan real-time.....	47
Apache Airflow — untuk jadwal batch (harian/mingguan).....	47
Dashboard (Superset, Metabase).....	48
Apache Superset — visualisasi tingkat nasional.....	48
Metabase — untuk pemda dan unit kecil.....	48
Big Data (Hadoop, Spark, Kafka).....	48
Hadoop — penyimpanan skala besar.....	48
Apache Spark — analitik cepat.....	48
Apache Kafka — “jalan tol” data real-time.....	48
AI Platform (Kubeflow, HuggingFace, Ollama).....	49
Kubeflow — pipeline AI pemerintah.....	49
HuggingFace — pustaka model AI terlengkap.....	49
Ollama — LLM lokal di server pemerintah.....	49
Cybersecurity (Wazuh, Suricata, pfSense, WireGuard).....	49
Wazuh — deteksi ancaman.....	49
Suricata — IDS/IPS open source.....	49
pfSense — firewall enterprise gratis.....	49
WireGuard — VPN aman antar-instansi.....	49
Tabel Open Source untuk Pemerintah.....	50
Pesan Inti.....	50
Bab 8 — Contoh Implementasi Nyata Open-Source untuk Pemerintah.....	51
ETL Bansos Daerah → Pusat Menggunakan n8n.....	51
Dashboard Kesehatan & Stunting Berbasis Superset.....	52
Data Warehouse Pendidikan Menggunakan PostgreSQL.....	53
AI Lokal dengan Ollama di Server Pemda.....	54
Monitoring Keamanan Wazuh untuk Kabupaten/Kota.....	55
Tabel Contoh Implementasi Open-Source.....	56
Pesan Inti.....	56
BAGIAN IV — PENGUMPULAN DATA NASIONAL & INTEROPERABILITAS.....	57
Bab 9 — Strategi Pengumpulan Data dari Daerah → Provinsi → Nasional.....	57
Model Sentralisasi vs Desentralisasi.....	57
A. Model Sentralisasi.....	57
B. Model Desentralisasi.....	58

Model terbaik untuk Indonesia: Hibrida (Hybrid).....	58
Validasi, Verifikasi, dan Deduplication (Kab/Kota → Provinsi → Pusat).....	58
Validasi (cek format).....	58
Verifikasi (cek kebenaran data).....	59
Deduplication (hapus data ganda).....	59
Hasil akhirnya:.....	59
Arsitektur Pelaporan Real-Time.....	60
Komponen Real-Time Nasional:.....	60
Contoh Real-Time:.....	60
Integrasi Sistem Daerah Warisan (Legacy).....	61
Strategi integrasi yang paling realistis:.....	61
Open-source pendukung integrasi Legacy → Modern:.....	61
Tabel Strategi Pengumpulan Data.....	62
Pesan Inti.....	62
Bab 10 — Mekanisme Data Sharing antar Kementerian & Lembaga.....	63
Insentif Berbagi Data — Cara Paling Efektif Menghilangkan Ego-Sektoral.....	63
Insentif efektif untuk K/L/D:.....	63
Insentif berbasis teknologi (open source):.....	64
Trust Framework Nasional — Membangun Kepercayaan Antar Instansi.....	64
Elemen penting Trust Framework Nasional:.....	64
Open source pendukung Trust Framework:.....	64
Federated Data Governance — Cara Praktis Menyatukan Data Tanpa Menyatukan Sistem.....	65
Konsep “Data Exchange Day” — Forum Operasional untuk Koordinasi Data Nasional	66
Tabel Mekanisme Data Sharing.....	67
Pesan Inti.....	67
BAGIAN V — DATA ANALYTICS, AI, & FORECASTING KEBIJAKAN.....	68
Bab 11 — Dasar-Dasar Forecasting untuk Pemerintah.....	68
Forecast Kesejahteraan Rakyat, Ekonomi, dan Pendidikan.....	69
A. Kesejahteraan Rakyat.....	69
B. Ekonomi Daerah.....	69
C. Pendidikan.....	69
Early Warning System Berbasis Data.....	70
Cara kerja EWS berbasis open-source:.....	70
Contoh “alarm data”:.....	70
Value for Money (VfM) Berbasis Data.....	70
A. Economy.....	71
B. Efficiency.....	71
C. Effectiveness.....	71
Open-source pendukung VfM:.....	71
Tabel Dasar Forecasting & EWS.....	72
Pesan Inti.....	72
Bab 12 — Pemanfaatan AI untuk Klasifikasi & Prediksi.....	73

Klasifikasi Kelompok Rentan & Wilayah Prioritas.....	73
Prediksi Anggaran, Dropout, Stunting, UMKM Workload.....	74
A. Prediksi Anggaran.....	74
B. Prediksi Dropout Sekolah.....	75
C. Prediksi Stunting.....	75
D. Prediksi Beban UMKM / Workload.....	75
Studi Kasus GovAI Global — Apa yang Bisa Kita Tiru?.....	76
🇸🇬 Singapura – Deteksi Dini Kebutuhan Sosial.....	76
🇬🇧 Inggris – Kesehatan Masyarakat (NHS).....	76
🇨🇳 Tiongkok – Perencanaan Transportasi dan Kota.....	76
🇨🇱 Chile – Pencegahan Putus Sekolah.....	76
Kesimpulan dan Pelajaran Praktis untuk Indonesia.....	76
Tabel GovAI.....	77
Pesan Inti.....	77
Bab 13 — Arsitektur AI Berjenjang Republik Indonesia.....	78
AI Level Desa → Kabupaten → Provinsi → Pusat.....	78
A. AI Level Desa/Kelurahan (AI Mikro – Mendekat ke Warga).....	78
B. AI Level Kabupaten/Kota (AI Operasional – Urusan Pelayanan Publik).....	79
C. AI Level Provinsi (AI Taktis – Koordinasi Lintas Daerah).....	79
D. AI Level Pusat (AI Strategis – Untuk Kebijakan Nasional).....	80
Tools nasional (open-source):.....	80
AI Operasional, AI Taktis, AI Strategis.....	80
A. AI Operasional (Kabupaten/Kota).....	80
B. AI Taktis (Provinsi).....	81
AI Strategis (Nasional).....	81
Tabel Arsitektur AI Berjenjang.....	82
Pesan Inti.....	82
Bab 14 — Penyiapan Data Training AI Nasional.....	83
Membangun Ground Truth Nasional.....	83
Labeling oleh ASN & Mahasiswa (Crowdsourcing Nasional).....	84
Synthetic Data, Data Augmentation, dan Federated Learning.....	85
A. Synthetic Data.....	85
B. Data Augmentation.....	85
C. Federated Learning.....	85
Audit Bias & Fairness.....	86
Tabel Penyiapan Data Training AI Nasional.....	87
Pesan Inti.....	87
Bab 15 — Analitik Berjenjang untuk Pengambilan Keputusan.....	88
Dashboard Operasional (Kabupaten/Kota).....	88
Jenis analitik yang dibutuhkan kabupaten/kota:.....	88
Karakter dashboard operasional:.....	88
Dashboard Taktis (Provinsi).....	89
Jenis analitik provinsi:.....	89
Karakter dashboard taktis:.....	89

Dashboard Strategis (Pusat/Nasional).....	90
Jenis analitik di pusat:.....	90
Karakter dashboard strategis:.....	90
Mengapa Open Source Sangat Penting?.....	91
Tabel Analitik Berjenjang Indonesia.....	91
Pesan Inti.....	91
BAGIAN VI — KEAMANAN, PRIVASI, & PERLINDUNGAN DATA PRIBADI.....	92
Bab 16 — Keamanan Infrastruktur Data Pemerintah.....	92
Mengapa Keamanan Data Pemerintah Sangat Kritis?.....	92
Zero Trust Architecture (ZTA).....	92
Enkripsi — Fondasi Mutlak Keamanan Nasional.....	93
IAM (Identity Access Management) — Kendalikan Siapa Boleh Akses Apa.....	94
RBAC (Role-Based Access Control).....	94
Logging — Semua Aktivitas harus Terekam.....	95
Open-source logging:.....	95
SIEM — Deteksi Serangan secara Real-Time.....	95
Open-source SIEM:.....	95
Tabel Keamanan Infrastruktur Data Pemerintah.....	96
Pesan Inti.....	96
Bab 17 — Cyber Security yang Efektif dan Efisien untuk Pemerintah.....	97
Realitas Hari Ini: Pemerintah Diserang Setiap Hari, tetapi Banyak Pemda Tidak Siap.	97
Hardening Sistem — Langkah Murah yang Paling Mengurangi Risiko.....	97
Security-in-a-Box untuk Pemda Kecil.....	98
GovCSIRT Nasional — Pusat Respon Insiden untuk Seluruh Pemerintah.....	99
Red Team – Blue Team & Bug Bounty untuk Pemerintah.....	99
Red Team.....	99
Blue Team.....	100
Bug Bounty untuk Pemerintah.....	100
Tabel Cyber Security Efektif & Efisien.....	101
Pesan Inti.....	101
Bab 18 — Perlindungan Data Pribadi sesuai UU PDP.....	102
UU PDP: Bukan Dokumen Hukum, Tetapi “Perisai Negara Digital”.....	102
Prinsip Utama UU PDP yang Wajib Dipahami ASN.....	102
A. Lawfulness & Purpose Limitation — Data digunakan untuk tujuan jelas.....	102
B. Data Minimization — Hanya ambil yang diperlukan.....	102
C. Accuracy — Data harus akurat dan terbaru.....	102
D. Storage Limitation — Tidak boleh menyimpan data selamanya.....	102
E. Integrity & Confidentiality — Data harus dilindungi.....	102
F. Accountability — Pemerintah bertanggung jawab atas pelanggaran.....	102
Masking, Hashing, Tokenisasi — Teknik Perlindungan Data yang Wajib Dipakai....	103
A. Masking — Menyembunyikan sebagian data.....	103
B. Hashing — Mengubah data menjadi kode permanen.....	103
C. Tokenisasi — Mengganti data asli dengan token sementara.....	103

Incident Response & Breach Handling — Harus Ada SOP Nasional.....	104
Langkah incident response untuk pemerintah:.....	104
Klasifikasi Data Sensitif & Rahasia Negara.....	105
A. Data Pribadi Umum.....	105
B. Data Pribadi Sensitif (kategori UU PDP).....	105
C. Data Rahasia Negara.....	106
Tabel Perlindungan Data Pribadi.....	106
Pesan Inti.....	107
BAGIAN VII — SDM, KAPASITAS, & KURIKULUM NASIONAL.....	108
Bab 19 — Kompetensi SDM Pemerintah di Era Data.....	108
SDM adalah Masalah Utama. Teknologi Bukan Penghambat.....	108
Data Literacy ASN — Pondasi Pemerintah Modern.....	108
Kompetensi dasar Data Literacy ASN:.....	108
Tools open-source untuk pembelajaran ASN:.....	109
Membaca Dashboard, Insight, dan Rekomendasi Kebijakan.....	109
Kompetensi yang diperlukan:.....	109
Tools open-source pendukung:.....	109
Pemahaman API & Keamanan Data — Wajib untuk Semua Wali Data.....	110
Pemahaman API yang wajib untuk pejabat data:.....	110
Apa risiko jika tidak paham API?.....	110
Pemahaman keamanan data yang harus dimiliki:.....	110
Tools open-source untuk memahami API dan keamanan:.....	110
Tabel Kompetensi SDM Pemerintah.....	111
Pesan Inti.....	111
Bab 20 — Kurikulum Kampus untuk Mendukung Data Governance RI.....	112
Indonesia Butuh SDM yang Bisa Bangun Sistem Data, Bukan Hanya Pengguna Aplikasi.....	112
Program Data Engineering & AI.....	112
Kompetensi inti Data Engineering untuk Indonesia:.....	112
Open source yang wajib ada dalam kurikulum:.....	112
Keamanan Siber & Privasi (UU PDP).....	113
Kompetensi keamanan yang dibutuhkan pemerintah:.....	113
Open-source yang wajib masuk kurikulum:.....	113
Open Source for Government.....	114
Kompetensi open-source untuk pemerintah:.....	114
Open-source wajib dalam kurikulum pemerintah:.....	114
Micro-credential untuk ASN.....	114
A. Data Literacy (Wajib).....	114
B. Dasar ETL.....	115
C. Dasar API & Pertukaran Data.....	115
D. Keamanan Data Dasar.....	115
Open-source yang cocok untuk micro-credential:.....	115
Tabel Kurikulum untuk Data Governance RI.....	115
Pesan Inti.....	116

Bab 21 — Model Pelatihan Nasional Bertingkat.....	117
Kenapa Indonesia Butuh Pelatihan Bertingkat?.....	117
Pelatihan Level Kabupaten, Provinsi, Kementerian.....	117
A. Pelatihan Level Kabupaten/Kota (Level Operasional).....	117
B. Pelatihan Level Provinsi (Level Taktis).....	118
C. Pelatihan Level Kementerian (Level Strategis).....	118
Train-the-Trainer Nasional (TtT).....	119
Prinsip Train-the-Trainer:.....	119
Kompetensi yang diberikan kepada Trainer Nasional:.....	119
Open-source untuk TtT:.....	119
Government Data Community — Komunitas Praktisi Pemerintah.....	119
Tabel Model Pelatihan Nasional Bertingkat.....	120
Pesan Inti.....	121
BAGIAN VIII — IMPLEMENTASI, ROADMAP, & STUDI KASUS.....	122
Bab 22 — Roadmap 10 Tahun Data Governance Nasional.....	122
Roadmap 10 Tahun: Perlu agar Indonesia Tidak Berjalan Tanpa Arah.....	122
TAHUN 1–2 — Standardisasi & Interoperability (Fondasi).....	122
TAHUN 3–4 — Konsolidasi Provinsi & GovCloud.....	123
TAHUN 5–7 — AI Nasional (AI untuk Semua Sektor).....	124
TAHUN 8–10 — Predictive Government (Pemerintah Prediktif).....	124
TABEL Roadmap 10 Tahun Data Governance Nasional.....	125
Pesan Inti.....	125
Bab 23 — Indikator Keberhasilan & Audit Sistem Data Nasional.....	126
Indeks Maturitas Data Daerah.....	126
Indeks Ketahanan Siber Pemerintah.....	127
Penghematan Anggaran dari Data Governance.....	128
Tabel Indikator & Audit Data Nasional.....	129
Pesan Inti.....	129
Bab 24 — Studi Kasus Pemerintah Republik Indonesia.....	130
Bansos — Dari Salah Sasaran ke Presisi Nasional.....	130
Pendidikan — Deteksi Dini Dropout & Kualitas Sekolah.....	131
Kesehatan — Stunting, Puskesmas, dan Early Warning System.....	131
UMKM — Prediksi Risiko dan Dukungan Tepat Sasaran.....	132
Fiskal Daerah — Perencanaan Anggaran Berbasis Data.....	132
Tabel Studi Kasus Pemerintah RI.....	133
Pesan Inti.....	133
Penutup & Rekomendasi Nasional.....	134
Kesimpulan Nasional: Data = Infrastruktur Negara.....	134
Empat Prinsip Utama yang Harus Dipertahankan Indonesia.....	134
1. Standarisasi nasional adalah wajib, bukan opsional.....	134
2. Infrastruktur data harus open-source dan bisa dikelola daerah.....	134
3. SDM pemerintah adalah fondasi yang harus dibangun terus-menerus.....	135
4. Audit berkala adalah kunci keberlanjutan.....	135
Rekomendasi Nasional.....	135

A. Untuk Presiden & Kementerian Koordinator.....	135
B. Untuk Kementerian Teknis (Kemenkes, Kemendikbud, Kemensos, UMKM, dll.) 135	
C. Untuk Provinsi.....	135
D. Untuk Kabupaten/Kota.....	135
E. Untuk Perguruan Tinggi.....	135
F. Untuk BSSN & GovCSIRT.....	135
Peran Open-Source dalam Kedaulatan Digital Indonesia.....	136
Pesan Akhir: Data Governance sebagai Fondasi Indonesia Emas 2045.....	136
Tabel Rekomendasi Nasional.....	137
Pesan Singkat.....	137
LAMPIRAN — JSON Schema Nasional.....	139
Apa Itu JSON Schema Nasional.....	139
Mengapa JSON Schema Nasional Sangat Penting?.....	140
Struktur Umum JSON Schema Nasional.....	140
Contoh JSON Schema Nasional (Kategori Besar).....	140
Manfaat Praktis untuk Pemerintah.....	142
Contoh Implementasi Open Source.....	142
Tabel JSON Schema Nasional.....	143
Pesan Inti.....	143
LAMPIRAN — Template JSON Schema.....	144
0. Blok Standar (bisa di-copy ke semua schema).....	144
A. BANSOS.....	145
1) bansos_rumah_tangga.json.....	145
2) bansos_individu.json.....	146
3) bansos_penyaluran.json.....	147
4) bansos_verifikasi.json.....	148
B. PENDIDIKAN.....	149
5) edu_siswa.json.....	149
6) edu_sekolah.json.....	149
7) edu_guru.json.....	150
8) edu_dropout_risk.json.....	150
9) edu_kehadiran.json.....	151
C. KESEHATAN.....	152
10) health_balita.json.....	152
11) health_status_gizi.json.....	152
12) health_faskes.json.....	153
13) health_vaksinasi.json.....	154
14) health_ews_case.json.....	154
D. UMKM.....	156
15) umkm_usaha.json.....	156
16) umkm_transaksi.json.....	157
17) umkm_risiko.json.....	158
18) umkm_bantuan.json.....	159

19) umkm_legalitas.json.....	160
E. FISKAL DAERAH.....	161
20) fiskal_belanja_daerah.json.....	161
21) fiskal_pad.json.....	162
22) fiskal_pembiayaan.json.....	163
23) fiskal_proyeksi.json.....	164
24) fiskal_realisasi_bulanan.json.....	165
F. METADATA & PELAPORAN.....	166
25) metadata_pelaporan.json.....	166
Ringkas: Cara Pakai di Open Source.....	167
LAMPIRAN: Template Peraturan Daerah Data Governance.....	168
Mengapa Perda Data Governance Dibutuhkan?.....	168
Isi Utama Template Perda Data Governance (Disederhanakan).....	169
Ketentuan Umum.....	169
Tujuan & Ruang Lingkup.....	169
Kelembagaan Data Daerah.....	169
Standar Data & Interoperabilitas.....	169
Infrastruktur & Keamanan.....	169
Pendanaan & SDM.....	169
Ketentuan Pengawasan & Sanksi.....	169
Penekanan Penggunaan Open Source di Dalam Perda.....	170
Bagaimana Pemda Menggunakan Template Ini?.....	171
Poin Kritis yang Harus Dijaga dalam Perda.....	171
Pesan Inti untuk Kepala Daerah & DPRD.....	172
LAMPIRAN — Tool Open Source + Panduan Deploy.....	173
Filosofi Tools Open Source.....	173
Kategori Tools.....	174
A. Database dan Data Warehouse.....	174
B. ETL / Integrasi Data.....	174
C. Dashboard & BI untuk Pimpinan.....	175
D. Big Data & Streaming.....	175
E. AI & Machine Learning.....	175
F. Data Lake & Storage.....	175
G. API Gateway & Interoperabilitas.....	175
H. Cybersecurity.....	176
I. DevOps & Automation.....	176
J. Kolaborasi & Arsip Dokumen.....	176
10 Tools Paling Penting Bagi Kabupaten/Kota.....	177
Top 10 tools:.....	177
Panduan Deploy Singkat (Untuk Daerah).....	178
1) Install Docker.....	178
2) Deploy PostgreSQL.....	178
3) Deploy n8n.....	178
4) Deploy Superset.....	178

5) Deploy Wazuh (keamanan).....	178
Tabel Open Source Tools & Fungsinya.....	179
Pesan Inti.....	179
LAMPIRAN — Draft Perpres Data Governance & GovCloud.....	180
Tujuan Utama Draft Perpres (Disederhanakan).....	180
Struktur Draft Perpres (Disederhanakan).....	181
A. Ketentuan Umum.....	181
B. Kebijakan Data Nasional.....	181
C. Arsitektur Data Nasional.....	181
D. GovCloud Nasional.....	181
E. Keamanan Siber.....	181
F. Standar Data dan JSON Nasional.....	181
G. Pemanfaatan Data & AI Nasional.....	181
H. Pengawasan & Audit.....	182
Prinsip “Open Source First” dalam Draft Perpres.....	182
Infrastruktur GovCloud.....	182
Data Governance.....	182
Keamanan Nasional.....	182
Dampak Praktis Perpres Ini bagi Pemerintah.....	183
Untuk Presiden dan Menteri.....	183
Untuk Provinsi.....	183
Untuk Kabupaten/Kota.....	183
Untuk BSSN.....	183
Untuk K/L sektor.....	183
Contoh Poin-Poin Kritis yang Perpres Wajib Tegaskan.....	183
Tabel Ringkas — Isi Draft Perpres.....	184
Pesan Inti.....	184
LAMPIRAN — Contoh Dashboard Nasional.....	185
Prinsip Dashboard Nasional yang Baik.....	185
Contoh Kategori Dashboard Nasional.....	185
A. Dashboard Bansos Nasional.....	185
B. Dashboard Pendidikan Nasional.....	186
C. Dashboard Kesehatan Nasional.....	186
D. Dashboard UMKM Nasional.....	187
E. Dashboard Fiskal Nasional.....	187
Mengapa Dashboard Nasional Penting untuk Birokrat?.....	187
Karena birokrat perlu tool yang:.....	187
Cara Cepat Membangun Dashboard Nasional (Pola Open Source).....	188
Langkah 1 — Data Masuk.....	188
Langkah 2 — Penyimpanan.....	188
Langkah 3 — Query Analitik.....	188
Langkah 4 — Visualisasi.....	188
Langkah 5 — Monitoring.....	188
Contoh Hasil Dashboard (disederhanakan).....	189

Contoh “Dashboard Nasional – Satu Layar”:	189
Tabel Ringkas Dashboard Nasional :	190
Pesan Inti :	190
Glossary :	191
Referensi :	200
Tentang Penulis :	202

Kata Pengantar

Alhamdulillah, dengan izin Allah SWT, buku ini dapat terselesaikan sebagai ikhtiar kecil untuk membantu pemerintah Indonesia membangun **fondasi data nasional yang kuat, rapi, aman, dan terstandar**. Kami mengucapkan **terima kasih** yang sebesar-besarnya kepada para birokrat, wali data, analis, akademisi, dan para pegiat teknologi yang terus berjuang memperbaiki tata kelola data di seluruh tingkatan pemerintahan. Tujuan utama penyusunan buku ini sangat sederhana namun krusial: *membangun dasar yang kokoh agar Indonesia dapat memiliki **data infrastructure** dan **data governance** yang modern, efisien, dan mampu mendukung keputusan berbasis bukti*. Buku ini juga bertujuan menumbuhkan **kemandirian dan kedaulatan digital**, dengan mendorong penggunaan **open source** yang lebih murah, lebih aman, bebas vendor lock-in, dan dapat dijalankan oleh daerah maupun pusat.

Harapan dan doa kami, semoga buku ini menjadi panduan praktis yang bisa langsung digunakan kementerian, provinsi, kabupaten, dan desa dalam upaya memperbaiki kualitas data, memperkuat interoperabilitas, meningkatkan keamanan, dan membangun sistem yang benar-benar bermanfaat bagi rakyat. Semoga Allah SWT memudahkan setiap langkah, memberikan keberkahan pada setiap kebijakan yang diambil, dan menjadikan upaya ini bagian kecil dari ikhtiar besar menuju Indonesia yang lebih siap, lebih mandiri, dan lebih berdaulat secara digital. Aamiin.

Jakarta, November 2025
Onno W. Purbo

Executive Summary

Indonesia perlu segera membangun fondasi kuat untuk Data Infrastructure dan Data Governance nasional melalui tiga langkah strategis: **menyatukan seluruh standar data pemerintah**—mulai dari JSON Schema, metadata, hingga API wajib—agar kabupaten, provinsi, dan kementerian dapat bertukar data secara konsisten; **mendirikan GovCloud nasional beserta node pusat data daerah** berbasis open-source untuk menghilangkan silo, menekan biaya, meningkatkan keamanan, dan memastikan data siap digunakan oleh AI; serta **menguatkan kelembagaan, keamanan, dan SDM** melalui Dewan Data Nasional, Wali Data di setiap OPD, baseline keamanan BSSN, dan program pelatihan besar-besaran bagi ASN dan kampus. Dengan menjalankan langkah-langkah ini secara terstruktur, pemerintah dapat meletakkan landasan menuju tata kelola data yang efisien, aman, interoperable, dan mampu mewujudkan pemerintahan prediktif berbasis bukti.

Langkah Strategis 1 Tahun Pertama

Ringkasan 1 Tahun (Tabel Singkat)

Fokus Strategis	Output Tahun Pertama	Dampak
Standardisasi	JSON Schema Nasional 1.0 & API wajib	Data seragam seluruh Indonesia
Kelembagaan	CDO/Wali Data terbentuk	Koordinasi berjalan
GovCloud	Cloud nasional + 10 provinsi pilot	Data tidak tercecer
Dashboard	Dashboard Nasional 1 layar	Keputusan lebih cepat
Keamanan	Wazuh SIEM + VPN nasional	Data lebih aman
SDM	Pelatihan ASN & kampus	Kapasitas meningkat
Integrasi Data	5 Data Prioritas tersatukan	Efisiensi anggaran & presisi program
Regulasi	Draft Perpres & template Perda	Kepatuhan meningkat

1. Tetapkan Standar Nasional Tunggal (3 bulan pertama)

Fondasi data negara tidak bisa dibangun tanpa standardisasi.

Aksi konkret:

- Sahkan **Standar Nasional JSON Schema** untuk 20–30 data prioritas (bansos, pendidikan, kesehatan, UMKM, fiskal).
- Terbitkan **Kamus Data Nasional** versi 1.0.
- Wajibkan setiap sistem pemerintah baru untuk:
 - menyediakan **API**,
 - menggunakan **open standard**,
 - mengikuti struktur data nasional yang sama.

Mengapa penting?

Sebelum data disatukan, formatnya harus seragam terlebih dahulu.

2. Bentuk Kelembagaan Inti Data Nasional (bulan 1–6)

Organisasi lebih menentukan daripada teknologi.

Aksi konkret:

- Aktifkan **Dewan Data Nasional** sebagai pengarah strategi.
- Tunjuk **Wali Data** di setiap kementerian/provinsi/kabupaten.
- Tunjuk **CDO (Chief Data Officer)** di setiap K/L strategis.
- Buat **SOP pengiriman data daerah** → provinsi → pusat.

Mengapa penting?

Tanpa struktur organisasi, data tidak akan pernah bergerak.

3. Bangun GovCloud Dasar & Node Data Daerah (bulan 1–12)

Tidak harus besar; yang penting *ada dan berjalan*.

Aksi konkret:

- Luncurkan **GovCloud Nasional** berbasis open source (OpenStack/Harvester).
- Bangun **Pusat Data Daerah (PDD)** minimal 10 provinsi pilot.
- Terapkan MinIO sebagai **storage nasional murah & aman**.
- Wajibkan backup harian ke GovCloud.

Mengapa penting?

Langkah ini menghentikan server liar, data tercecer, dan risiko kebocoran.

4. Etalase Dashboard Nasional (bulan 4–12)

Mulai dari indikator sempit, tapi konsisten.

Aksi konkret:

- Buat **Dashboard Nasional 1-Layar** untuk 5 sektor prioritas:
 1. Bansos
 2. Kesehatan
 3. Pendidikan
 4. UMKM
 5. Fiskal
- Gunakan tools open source: **Superset, Metabase, Grafana**.
- Pastikan data dari daerah mengalir harian.

Mengapa penting?

Ini memberi sinyal kuat ke seluruh daerah bahwa data dipakai untuk keputusan.

5. Aman, Enkripsi, dan Monitoring Nasional (bulan 1–12)

Aksi konkret:

- Terapkan **Zero Trust dasar** untuk seluruh server pemerintah.
- Implementasi **Wazuh SIEM Nasional** (open source).
- Terapkan **WireGuard VPN** untuk pertukaran data antar instansi.
- Audit keamanan bulanan untuk 100 instansi (batch awal).

Mengapa penting?

Data governance tidak bisa berdiri kalau keamanan bocor.

6. Percepatan SDM dan Pelatihan Nasional (bulan 3–12)

Tanpa SDM, semua strategi gagal.

Aksi konkret:

- Buat **bootcamp ASN**: API, dashboard, ETL, keamanan.
- Wajibkan Wali Data ikut pelatihan 40 jam.
- Kerja sama kampus untuk **micro-credential data engineering**.
- Pelatihan desa: input data, validasi, verifikasi.

Mengapa penting?

Standar bisa dibuat, tapi manusia yang menjalankan.

7. Integrasi Data Prioritas (bulan 4–12)

Tidak semua data harus langsung disatukan. Mulai dari “5 Data Emas”.

Aksi konkret:

Integrasikan data:

1. **Kependudukan**
2. **Bansos**
3. **Kesehatan balita/stunting**
4. **Pendidikan (siswa/sekolah)**
5. **UMKM**

Gunakan tools open-source:

- n8n / NiFi (ETL),
- PostgreSQL / ClickHouse (DW),
- MinIO (data lake).

Mengapa penting?

Data ini paling berdampak langsung pada kesejahteraan rakyat.

8. Rancangan Perpres + Perda Data Governance (bulan 1–12)

Aksi konkret:

- Finalisasi **Perpres Data Governance & GovCloud** sebagai payung hukum nasional.
- Siapkan template **Perda Satu Data Daerah** → provinsi/kabupaten tinggal mengadopsi.
- Buat PKS/SOP baku untuk data sharing antar instansi.

Mengapa penting?

Tanpa dasar hukum, instansi enggan berbagi data.

Pesan Inti

- Tahun pertama bukan tentang membangun sistem besar.
- Tahun pertama adalah tentang **menyatukan standar, menyiapkan SDM, mengamankan data, dan menjalankan integrasi kecil namun signifikan.**
- Jika pondasi ini berhasil, 3–5 tahun berikutnya Indonesia dapat memiliki **GovCloud kuat, dashboard nasional real-time, dan AI pemerintah yang benar-benar presisi.**

BAGIAN I — LANDASAN & URGENSI NASIONAL

Bab 1 — Mengapa Data Governance Penting untuk Republik Indonesia

Data bukan sekadar tabel. Data adalah **dasar keputusan** yang menentukan hidup orang banyak. Tanpa data yang rapi dan akurat, pemerintah bekerja dalam keadaan setengah buta.

Contoh yang terjadi sehari-hari:

- Data kemiskinan tidak sinkron → bantuan tidak tepat sasaran.
- Data pendidikan tidak lengkap → sekolah tidak mendapat dukungan sesuai kebutuhan.
- Data kesehatan lambat → stunting terlambat ditangani.

Kesimpulan sederhana:

Semakin baik kualitas data, semakin cepat dan tepat program pemerintah meningkatkan kesejahteraan rakyat.

Data Buruk Menimbulkan Biaya Ekonomi yang Besar

Data yang salah bukan hanya mengacaukan laporan. Data yang salah membuat negara **kehilangan uang**.

Kerugian nyata:

- Anggaran bansos salah sasaran, tidak tepat ke penerima.
- Pemerintah membeli obat, tool, bantuan, tapi tidak berdasarkan kebutuhan sebenarnya.
- Proyek pendidikan tidak efektif karena perencanaan salah.

Dampaknya:

- Pemborosan anggaran.
- Program tidak efektif.
- Kepercayaan publik menurun.

Biaya terbesar negara bukan eksekusi program. Biaya terbesar adalah akibat data yang buruk.

Keputusan Pemerintah Hanya Sebaik Kualitas Datanya

Setiap keputusan pejabat — dari camat sampai menteri — **bergantung penuh** pada data.

Jika data salah:

- kebijakan salah
- anggaran salah
- target salah
- hasil tidak tercapai

Jika data benar:

- keputusan cepat
- program tepat sasaran
- pengawasan mudah
- kinerja meningkat

Ini prinsip sederhana:

Garbage In = Garbage Out.

Data buruk → kebijakan buruk.

Data baik → kebijakan baik.

Contoh Nyata Kegagalan Program Akibat Data Tidak Akurat

Empat bidang ini paling sering terdampak:

Bansos

- Nama ganda
- NIK tidak valid
- Alamat tidak jelas
 - bantuan tidak sampai ke warga yang benar.

Stunting

- Data berat badan/tinggi anak tidak terupdate
- Pencatatan manual
 - intervensi terlambat.

PIP (Program Indonesia Pintar)

- Data siswa tidak lengkap
- NISN tidak valid
 - bantuan pendidikan salah sasaran.

UMKM

- Data omzet dan kategori usaha tidak akurat
→ kebijakan permodalan dan subsidi tidak tepat.

Titik masalahnya: semua berawal dari data yang tidak standar, tidak terverifikasi, dan tidak terhubung antar lembaga.

Single Source of Truth (SSOT) adalah Pondasi Negara Digital

SSOT berarti: **semua instansi melihat data yang sama, versi yang sama, dan standar yang sama.**

Mengapa penting?

- Menghilangkan duplikasi data.
- Menghilangkan perbedaan angka antar kementerian.
- Menghilangkan saling menyalahkan.
- Mempercepat keputusan.

Contoh:

Jika Kemenkeu, Kemendagri, Kemenkes, dan Bappenas memakai data kemiskinan yang berbeda → mustahil membuat kebijakan yang konsisten.

SSOT membuat:

- Perencanaan nasional lebih kuat.
- Anggaran lebih efisien.
- Program lebih tepat sasaran.

Kesimpulan:

Tanpa SSOT, negara digital hanyalah slogan. Dengan SSOT, Indonesia bisa bergerak cepat menuju pemerintahan modern yang otomatis dan berbasis bukti.

Tabel Pentingnya Data Governance

Isu Utama	Dampak Lapangan	Risiko untuk Negara	Solusi Data Governance
Data tidak akurat	Bansos salah sasaran, laporan tidak cocok	Pemborosan anggaran besar	Validasi, verifikasi, deduplication
Data tersebar di banyak instansi	Angka K/L berbeda-beda	Kebijakan tidak sinkron	Single Source of Truth
Metode pencatatan manual	Stunting terlambat diketahui	Intervensi tidak tepat waktu	Digitalisasi pencatatan
Tidak ada standar	Format data kacau, sulit integrasi	Tumpang tindih sistem	Standar JSON Nasional
Data tidak update	PIP salah target, UMKM salah klasifikasi	Program tidak efektif	Integrasi real-time daerah → pusat

Pesan Inti

- Data adalah fondasi kesejahteraan rakyat.
- Data buruk = anggaran bocor = program gagal.
- Keputusan pemerintah hanya sebaik kualitas datanya.
- Empat sektor paling terdampak: bansos, stunting, PIP, UMKM.
- Single Source of Truth wajib untuk Indonesia menjadi negara digital.

Bab 2 — Kerangka Regulasi & Visi Nasional 2045

Regulasi adalah “rel kereta” bagi Data Governance Indonesia. Tanpa rel, setiap instansi berjalan sendiri, membuat sistem sendiri, dan menciptakan data sendiri. Hasilnya adalah **kekacauan data nasional**, tumpang tindih anggaran, dan keputusan yang tidak konsisten.

Bagian ini menekankan bahwa Indonesia sudah memiliki dasar hukum kuat. Tantangannya bukan membuat aturan baru, tetapi **memastikan aturan yang sudah ada bisa benar-benar berjalan di lapangan**.

Perpres Satu Data Indonesia (SDI)

Perpres SDI adalah **pondasi utama** integrasi data pemerintah. Poin penting bagi birokrat:

- SDI mewajibkan **standar data** dan **interoperabilitas**.
- Ada peran jelas: Pembina Data, Wali Data, dan Produsen Data.
- Setiap data harus memiliki **metadata**, format baku, dan referensi yang jelas.
- SDI mengharuskan koordinasi pusat–daerah, bukan jalan sendiri-sendiri.

Intinya: SDI adalah aturan yang memaksa seluruh instansi berhenti membuat data versi masing-masing.

SPBE — Kerangka Digital Pemerintahan Modern

SPBE adalah kerangka untuk memastikan layanan pemerintah **terhubung, efisien, dan digital by default**. Hal yang paling relevan untuk Data Governance:

- SPBE mendorong integrasi layanan lintas K/L.
- SPBE menuntut konsolidasi aplikasi agar tidak ada 10 aplikasi untuk fungsi yang sama.
- SPBE menempatkan data sebagai “aset strategis”, bukan sekadar dokumen.

SPBE + SDI bekerja seperti dua sisi mata uang:

SDI mengatur data → SPBE mengatur layanan digital yang memakai data.

UU PDP — Pengamanan Data Pribadi Masyarakat

UU PDP adalah perlindungan dasar agar data warga tidak disalahgunakan. Yang harus dipahami pejabat:

- Data pribadi harus diproses dengan **legal basis** yang jelas.
- Ada kewajiban **keamanan, minimasi data, dan transparansi**.
- Ada kewajiban **respons terhadap insiden kebocoran data**.
- Ada ancaman **sanksi administratif dan pidana** bagi pelanggaran.

Artinya: Data Governance tidak boleh hanya bicara integrasi, tetapi juga privasi dan keamanan.

GovCloud — Infrastruktur Negara Digital 2045

Tidak mungkin Data Governance berjalan tanpa infrastruktur yang kokoh. GovCloud memberi:

- Pusat Data Nasional (PDN) dan Cloud pemerintah.
- Integrasi pusat–daerah dengan standar keamanan tinggi.
- Lokasi penyimpanan data yang lebih aman dan terkendali.
- Infrastruktur untuk AI pemerintahan dan layanan digital masa depan.

GovCloud = fondasi fisik negara digital.

Tanpanya, integrasi dan keamanan hanya menjadi konsep di atas kertas.

Model Kematangan Data Governance (Maturity Model)

Maturity Model membantu pemerintah menilai seberapa siap mereka mengelola data. Ini bukan teori — ini tool praktis. Tingkatannya biasanya seperti ini:

1. **Level 1 – Ad hoc**
 - Data tersebar
 - Format berbeda
 - Tidak ada standar
2. **Level 2 – Terstandar**
 - Ada format
 - Ada metadata
 - Belum terintegrasi
3. **Level 3 – Terintegrasi**
 - Data antar instansi mulai sinkron
 - Ada API exchange
4. **Level 4 – Terkelola & Konsisten**
 - Master Data Nasional
 - Pengawasan kualitas data
5. **Level 5 – Data Driven Government**
 - AI digunakan pada setiap perencanaan
 - Dashboards lintas K/L real-time
 - Single Source of Truth nasional

Tujuan Indonesia 2045:

- Minimal seluruh kementerian & provinsi mencapai **Level 4**
- Pemerintah pusat mencapai **Level 5**

Penyelarasan dengan RPJMN & Visi Indonesia Emas 2045

Data Governance bukan proyek IT. Data Governance adalah **fondasi pembangunan nasional**. Kaitannya dengan RPJMN & Indonesia Emas 2045 sangat jelas:

A. RPJMN butuh data yang akurat untuk:

- Menentukan target pengentasan kemiskinan
- Menyusun kebijakan pendidikan
- Menentukan alokasi fiskal daerah
- Memonitor kesehatan dan stunting
- Menilai efektivitas program nasional

Jika data salah, seluruh RPJMN meleset.

B. Indonesia Emas 2045 membutuhkan:

- Negara digital yang efisien
- Pelayanan publik otomatis
- Penggunaan AI secara nasional
- Anggaran tepat sasaran
- Pelayanan berbasis data real-time

Semua hanya mungkin bila:

- Data terintegrasi
- Regulasi berjalan
- GovCloud beroperasi
- Privasi dan keamanan dijaga

Tanpa Data Governance, Indonesia Emas 2045 hanya menjadi slogan.

Tabel Kerangka Kebijakan

Komponen	Fungsi Inti	Masalah yang Dipecahkan	Dampak ke Pemerintah
Perpres Satu Data Indonesia	Standarisasi & interoperabilitas	Data beda-beda antar instansi	Satu format, satu bahasa data nasional
SPBE	Integrasi layanan digital	Aplikasi tumpang tindih	Layanan lebih efisien & terhubung
UU PDP	Perlindungan data pribadi	Kebocoran & penyalahgunaan data	Keamanan & kepercayaan publik meningkat
GovCloud	Infrastruktur tunggal pemerintah	Server tercerai-berai	Pusat Data Nasional yang aman & modern
Maturity Model	Mengukur kesiapan data	Tidak tahu posisi saat ini	Roadmap peningkatan data yang jelas
RPJMN & 2045	Arah pembangunan jangka panjang	Perencanaan tidak berbasis data	Kebijakan nasional lebih akurat & efektif

Pesan Inti

- Indonesia sudah punya aturan kuat: SDI, SPBE, UU PDP, GovCloud.
- Tantangan kita adalah **implementasi**, bukan membuat aturan baru.
- Maturity Model membantu mengukur kemajuan.
- Semua harus selaras dengan RPJMN dan visi Indonesia Emas 2045.
- Tanpa Data Governance, pembangunan nasional tidak pernah tepat sasaran.

BAGIAN II — ORGANISASI & TATA KELOLA DATA

Bab 3 — Arsitektur Kelembagaan Data Nasional

Arsitektur kelembagaan adalah “siapa mengerjakan apa”. Tanpa struktur yang jelas, Data Governance tidak akan berjalan meskipun infrastrukturnya sudah bagus.

Masalah yang sering terjadi:

- Banyak instansi merasa “bukan tugas saya”.
- Tidak ada penanggung jawab yang tegas.
- Tidak ada mekanisme koreksi ketika data salah.
- Keputusan dibuat berdasarkan kebiasaan, bukan data.

Dewan Data Nasional, CDO, DPO, CIO — “Komando Utama Data Negara”

Struktur teratas ini adalah pengarah kebijakan dan pengambil keputusan strategis.

Dewan Data Nasional

- Memastikan seluruh kementerian/lembaga memakai standar data yang sama.
- Menetapkan kebijakan integrasi dan interoperabilitas.
- Menghilangkan ego-sektoral dalam pengelolaan data.

Analogi sederhana: Dewan Data Nasional adalah “dirigen”. Tanpa dirigen, semua instansi memainkan musik sendiri-sendiri.

CDO (Chief Data Officer)

- Memimpin strategi data di tiap kementerian/lembaga.
- Mengawasi kualitas data, integrasi, dan keamanan.
- Bertanggung jawab memastikan kebijakan didukung data yang valid.

DPO (Data Protection Officer)

- Mengawasi kepatuhan terhadap UU PDP.
- Memastikan privasi warga terjaga.
- Menangani insiden kebocoran data.

CIO (Chief Information Officer)

- Mengatur teknologi, aplikasi, server, cloud.
- Menjamin infrastruktur mendukung Data Governance.

Intinya:

- **CDO** fokus pada data.
 - **CIO** fokus pada teknologi.
 - **DPO** fokus pada privasi dan keamanan.
- Ketiganya harus bekerja **sebagai satu paket**.

Data Owner, Data Steward, Wali Data — “Pelaksana Harian Data Pemerintah”

Kelompok ini bekerja di lapangan. Merekalah yang membuat integrasi data benar-benar berjalan.

Data Owner (Pemilik Data)

- Biasanya Eselon I atau pimpinan unit.
- Penanggung jawab utama kualitas dan akurasi data.
- Memutuskan siapa yang boleh mengakses data.

Data Steward

- Petugas teknis yang menjaga kualitas data setiap hari.
- Memastikan data bersih, lengkap, dan mengikuti standar SDI.
- Mengelola metadata dan kamus data.

Wali Data

- Peran yang ditetapkan dalam Perpres SDI.
- Bertugas mengkoordinasikan pengumpulan data dari unit-unit lain.
- Penghubung antara Produsen Data dan Dewan Data Nasional.

Kesalahan terbesar pemerintah selama ini:

Tidak ada yang bertanggung jawab penuh pada data.
Dengan struktur ini, tidak ada lagi data “tidak bertuan”.

SOP Pengambilan Keputusan Data-Driven (Decision Framework)

Pemerintah tidak boleh lagi mengambil keputusan berdasarkan “feeling”, “kebiasaan”, atau “kira-kira”. Keputusan harus melewati **alur tetap**.

Langkah-langkah Decision Framework yang sederhana dan kritis:

1. **Identifikasi masalah**
 - Apa isu kebijakan?
 - Dampaknya ke masyarakat?
2. **Ambil data dari SSOT (Single Source of Truth)**
 - Tidak boleh pakai data versi masing-masing instansi.
 - Harus dari sumber terverifikasi.
3. **Validasi & Cek kualitas data**
 - Apakah datanya lengkap?
 - Apakah formatnya sesuai standar SDI?
4. **Analisis & skenario**
 - Apa yang terjadi jika A?
 - Apa yang terjadi jika B?
5. **Rekomendasi berbasis bukti**
 - Keputusan harus didukung angka, bukan asumsi.
6. **Dokumentasi keputusan**
 - Catat sumber data, tanggal, dan pihak yang menyetujui.
7. **Monitoring & evaluasi**
 - Apakah keputusan tadi memberi dampak?
 - Perlu revisi atau tidak?

Prinsip emas:

Tidak ada data → tidak ada keputusan.

Tabel Arsitektur Kelembagaan Data

Komponen	Peran Utama	Risiko Jika Tidak Ada	Dampak Jika Berjalan Baik
Dewan Data Nasional	Koordinasi data seluruh K/L	Ego sektoral, data terpecah	Data nasional konsisten
CDO	Mengatur strategi data	Kualitas data buruk	Kebijakan berbasis bukti
DPO	Menjaga privasi & UU PDP	Kebocoran data	Kepercayaan publik naik
CIO	Infrastruktur & teknologi	Sistem tidak terhubung	Integrasi nasional lancar
Data Owner	Pemilik tanggung jawab data	Tidak ada accountability	Data rapi & dapat diaudit
Data Steward	Menjaga kualitas data	Data kotor, tidak standar	Data siap dipakai AI & analitik
Wali Data	Penghubung pusat–daerah	Data tidak terkumpul	Pelaporan nasional akurat
Decision Framework	Keputusan berbasis data	Kebijakan salah	Anggaran tepat sasaran

Pesan Inti

- Struktur kelembagaan data menentukan keberhasilan Data Governance.
- Peran harus jelas: siapa mengelola data, siapa menjaga standar, siapa membuat keputusan.
- Tanpa CDO–DPO–CIO dan Wali Data, integrasi nasional tidak mungkin terjadi.
- Decision Framework memastikan semua keputusan pemerintah berbasis bukti, bukan perasaan.

Bab 4 — Standar Data Nasional & Interoperabilitas

Standar data adalah “bahasa bersama” antar instansi. Tanpa standar, data tidak bisa digabung, tidak bisa dibandingkan, dan tidak bisa dipakai untuk kebijakan nasional.

Masalah paling sering terjadi:

- Setiap instansi memakai format sendiri.
- Nama variabel berbeda-beda untuk hal yang sama.
- Data ganda tidak bisa dideteksi.
- Integrasi pusat–daerah selalu tersendat.

Karena itu, bagian ini menegaskan bahwa **standar data bukan pilihan — tetapi keharusan nasional**.

Metadata Nasional & Kamus Data

Metadata adalah “penjelasan tentang data”.

Kamus Data adalah “arti dari setiap kolom”.

Tanpa keduanya, data seperti dokumen tanpa judul: ada isinya, tapi tidak jelas maknanya.

Metadata Nasional memastikan:

- Setiap dataset memiliki definisi jelas.
- Sumber data tercatat dengan rapi.
- Metode pengumpulan data transparan.
- Tanggal update, versi, dan kualitas dapat dilacak.

Kamus Data memastikan:

- Semua instansi memakai nama kolom yang sama.
- Tidak ada lagi 10 variasi nama untuk “NIK”, “nama_kepala_keluarga”, atau “pendapatan”.
- Data dapat dikombinasikan antar instansi tanpa konflik.

Inti penting bagi birokrat:

Metadata & Kamus Data adalah tool untuk memastikan *konsistensi nasional*.

Tanpa ini, Data Governance tidak akan bergerak.

Sistem Data Catalog & Data Lineage

Data Catalog adalah “daftar semua data pemerintah”.

Data Lineage adalah “jejak perjalanan data”.

Keduanya sangat penting untuk mengatasi masalah klasik:

- Data tersebar di mana-mana.
- Tidak tahu versi data yang paling baru.
- Tidak tahu siapa pemilik data.
- Tidak tahu asal-usul data ketika muncul anomali.

Data Catalog menyediakan:

- Satu portal pusat yang mencatat semua dataset K/L/D.
- Informasi kualitas data.
- Metadata lengkap.
- Pemilik data (Data Owner) dan penjaga data (Data Steward).

Data Lineage menjawab pertanyaan kritis:

- Dari mana data ini berasal?
- Siapa yang mengubah?
- Kapan diperbarui?
- Apakah pernah terjadi manipulasi atau kesalahan input?

Manfaat langsung:

Ketika ada masalah data, birokrat tidak perlu menebak-nebak.

Data Lineage langsung menunjukkan sumber masalah.

Format Standar Pertukaran Data Nasional (JSON Schema)

JSON Schema adalah format baku untuk bertukar data antar instansi. Masalah besar selama ini:

- Setiap instansi memakai struktur berbeda.
- Kabupaten kirim data dalam Excel, provinsi pakai CSV, pusat pakai database.
- Integrasi memakan waktu karena harus membersihkan format setiap kali.

Dengan JSON Schema Nasional:

- Semua data memakai struktur yang sama.
- Validasi bisa otomatis.
- Kesalahan input bisa dicegah.
- Integrasi API jadi mudah.
- Pelaporan pusat–daerah menjadi cepat dan rapi.

JSON Schema = fondasi teknis interoperabilitas.

Policy PKS/SOP Antar-Lembaga

Standar teknis saja tidak cukup. Perlu aturan tertulis agar instansi mau berbagi data dengan aman.

PKS (Perjanjian Kerja Sama) dan SOP antar-instansi memastikan:

- Siapa boleh mengakses data.
- Data apa yang boleh dibagi.
- Frekuensi pembaruan data.
- Kewajiban menjaga kerahasiaan.
- Penanganan insiden data.
- Sanksi jika terjadi pelanggaran.

Intinya:

Interoperabilitas bukan hanya teknologi — tetapi juga komitmen organisasi.
Tanpa PKS/SOP, integrasi data hanya menjadi wacana.

Tabel Standard Data Nasional

Komponen	Fungsi Inti	Risiko Jika Tidak Ada	Dampak Jika Berjalan Baik
Metadata Nasional	Menjelaskan dataset	Data tidak bisa dipahami	Data mudah dianalisis
Kamus Data	Menyatukan definisi kolom	Data antar instansi tidak cocok	Integrasi nasional lancar
Data Catalog	Menemukan dan mengelola data	Data tersebar & tidak diketahui	Transparansi dan efisiensi data
Data Lineage	Melacak asal-usul data	Sulit audit & koreksi	Data terpercaya & mudah diaudit
JSON Schema	Format standar bertukar data	Format kacau, integrasi mahal	Validasi otomatis & data bersih
PKS/SOP antar-Lembaga	Aturan berbagi data	Ego-sektoral & kebuntuan birokrasi	Akses data cepat, aman, dan legal

Pesan Inti

- Standar data adalah syarat utama interoperabilitas nasional.
- Metadata & Kamus Data membuat data bisa dipahami dan dipakai.
- Data Catalog & Lineage memastikan data rapi, dapat dilacak, dan tidak “gelap”.
- JSON Schema menyelesaikan problem format yang selama ini menghambat integrasi.
- PKS/SOP membuat kerja sama data antar-instansi legal, jelas, dan aman.

BAGIAN III — INFRASTRUKTUR DATA & TEKNOLOGI PEMERINTAH

Bab 5 — Arsitektur Infrastruktur Data Indonesia

Infrastruktur data adalah “mesin” di balik seluruh Data Governance. Kalau mesinnya lemah, seluruh kebijakan berbasis data menjadi lambat, salah, atau gagal.

Masalah nyata di lapangan:

- Data terpecah di banyak server.
- Tidak ada standar penyimpanan.
- AI tidak bisa dijalankan karena server tidak memadai.
- Daerah memakai aplikasi berbeda-beda.

Karena itu, arsitektur data harus **seragam**, **terukur**, dan **menggunakan teknologi yang kuat namun efisien**, termasuk **open-source**.

Data Warehouse, Data Lake, Data Mesh, Lakehouse

Empat komponen ini adalah “tempat tinggal data”. Bahasannya sering teknis, tapi kita sederhanakan agar mudah dipahami birokrat.

Data Warehouse — untuk laporan resmi pemerintah

- Dipakai untuk angka final seperti kemiskinan, pendidikan, kesehatan.
- Data sudah rapi, bersih, terstruktur.
- Tingkat keakuratannya harus tinggi.

Open-source contoh:

- **PostgreSQL** sebagai data warehouse pemerintah.
- **Metabase / Apache Superset** untuk dashboard menteri/gubernur.

Data Lake — untuk data besar (big data) dan data mentah

- Menampung CCTV, IoT, file besar, data belum dibersihkan.
- Digunakan untuk analitik mendalam dan pelatihan AI.

Open-source contoh:

- **Hadoop** untuk penyimpanan skala besar.
- **Apache Spark** untuk analitik cepat.
- **MinIO** untuk penyimpanan objek (S3-compatible).

Data Mesh — pembagian per-domain di dalam K/L besar

- Setiap direktorat memiliki domain data, tapi tetap mengikuti standar nasional.
- Mengurangi beban pusat karena domain mengelola datanya sendiri.

Open-source contoh:

- **GitLab CE** sebagai repositori domain data.
- **OpenMetadata** sebagai metadata domain.

Lakehouse — gabungan terbaik Warehouse + Lake

- Platform tunggal untuk data mentah dan data terstruktur.
- Memudahkan integrasi AI, laporan, dan analitik.

Open-source contoh:

- **Databricks open-source alternatives: Apache Iceberg / Delta Lake / Apache Hudi.**

Kesimpulan:

Warehouse = laporan

Lake = data mentah

Mesh = organisasi

Lakehouse = integrasi nasional jangka panjang

Pusat Data Nasional (PDN) & Pusat Data Daerah (PDD)

PDN dan PDD adalah “rumah besar data negara”.

Pusat Data Nasional (PDN)

- Semua data strategis negara disimpan di sini.
- Keamanan tinggi, standar nasional, terpusat.
- Menghapus kebutuhan setiap K/L membuat server sendiri.

Pusat Data Daerah (PDD)

- Menangani data kab/kota dan provinsi.
- Mencegah overload ke PDN.
- Memberi ruang bagi daerah untuk analitik mandiri.

Manfaat besar:

- Menghilangkan server “liar” di daerah.
- Mengurangi biaya server masing-masing instansi.
- Data provinsi dan kabupaten tetap bisa berjalan saat jaringan pusat terganggu.

Open-source contoh implementasi PDN/PDD:

- **Proxmox / OpenStack** sebagai platform virtualisasi.
- **Ceph** untuk penyimpanan terdistribusi.
- **Wazuh** untuk monitoring keamanan PDN/PDD.

Government Cloud (Hybrid / Multi-cloud)

Cloud adalah cara modern menjalankan layanan pemerintah. Bukan lagi membeli server, tetapi menggunakan infrastruktur elastis.

Hybrid Cloud

- Kombinasi PDN + cloud nasional + cloud swasta.
- Cocok untuk data yang butuh fleksibilitas tetapi tetap aman.

Multi-cloud

- Menggunakan lebih dari satu cloud.
- Menghindari ketergantungan satu vendor (vendor lock-in).
- Meningkatkan ketahanan layanan (high availability).

Keuntungan utama:

- Skalabilitas instan.
- Biaya jauh lebih efisien.
- SIAP menghadapi lonjakan data seperti Pemilu, PPDB, dan Bansos.

Open-source contoh:

- **OpenStack** sebagai fondasi private Government Cloud.
- **Kubernetes** sebagai orkestrator aplikasi.
- **Harbor** untuk container registry lokal.

Jaringan, Server, Komputasi CPU/GPU/NPU — Mesin AI Nasional

Jika Data Governance adalah kerangka kerja, maka komputasi adalah “tenaganya”.

Jaringan

- Minimal 1–10 Gbps antar PDN–PDD.
- Harus terenkripsi end-to-end.
- Wajib ada monitoring dan logging.

Open-source contoh:

- **pfSense** atau **VyOS** untuk router/firewall pemerintah.
- **WireGuard** untuk VPN aman antar instansi.

Server (Compute & Storage)

- Tidak boleh lagi pakai server kecil yang tidak tahan beban.
- Harus mendukung skalabilitas dan redundansi.

Open-source contoh:

- **ZFS / Ceph** untuk penyimpanan pemerintah.
- **Prometheus + Grafana** untuk monitoring.

CPU (General Processing)

- Cocok untuk database, layanan web, dan aplikasi umum.
- Murah, hemat energi, stabil.

GPU (Artificial Intelligence Computing)

- Wajib untuk AI, deep learning, computer vision.
- Tanpa GPU, semua proyek AI pemerintah akan sangat lambat.

Open-source contoh:

- **Ollama** untuk menjalankan LLM lokal di GPU.
- **Kubeflow** untuk pipeline AI di cloud/hybrid.

NPU (Neural Processing Unit, Edge AI)

- Untuk inference AI cepat di perangkat kecil.
- Cocok untuk daerah yang koneksi internetnya lemah.
- Dipakai untuk CCTV cerdas, sensor pertanian, deteksi kesehatan.

Open-source contoh:

- **OpenVINO** untuk optimasi NPU dan Edge AI.

Tabel Arsitektur Infrastruktur Data

Komponen	Fungsi	Risiko jika tidak ada	Dampak jika berjalan	Contoh Open-source
Data Warehouse	Laporan resmi	Data tidak sinkron	Laporan rapi	PostgreSQL, Metabase
Data Lake	Data besar mentah	AI tidak bisa jalan	Analitik kuat	Hadoop, Spark, MinIO
Data Mesh	Domain data per unit	Beban pusat overload	Koordinasi efisien	GitLab CE, OpenMetadata
Lakehouse	Integrasi Lake+Warehouse	Data terpecah	Platform terpadu	Apache Iceberg / Delta
PDN/PDD	Pusat Data Negara	Server liar & bocor	Integrasi nasional	OpenStack, Ceph
Hybrid/Multi-cloud	Fleksibilitas & kecepatan	Gangguan layanan	Hemat biaya & tahan beban	Kubernetes, Harbor
CPU/GPU/NPU	Komputasi AI & data	Analitik lambat	AI nasional berjalan	Ollama, Kubeflow, OpenVINO
Jaringan	Konektivitas aman	Data bocor	Alur data stabil	pfSense, WireGuard

Pesan Inti

- Infrastruktur data adalah pondasi negara digital.
- Data Warehouse, Lake, Mesh, Lakehouse harus dipakai sesuai fungsinya.
- PDN–PDD menyatukan data pusat dan daerah.
- Government Cloud membuat layanan pemerintah lebih cepat, aman, dan efisien.
- CPU–GPU–NPU adalah mesin utama untuk AI nasional.
- Open-source membuat semua ini **lebih murah, lebih aman, dan lebih mandiri**.

Bab 6 — Sistem Pertukaran Data Nasional (National Data Exchange)

Indonesia tidak kekurangan data. Yang kurang adalah **alur data yang rapi, cepat, dan standar** antar instansi.

Masalah nyata hari ini:

- Tiap instansi punya API sendiri, format sendiri, standar sendiri.
- Data kabupaten → provinsi → pusat sering tidak sinkron.
- Pertukaran data lambat karena masih manual (Excel, WA, email).
- Duplikasi data terjadi di banyak tempat.
- Pusat kesulitan mengambil keputusan karena datanya tidak konsisten.

Bagian ini adalah jawaban: **kita harus punya Sistem Pertukaran Data Nasional.**

API Gateway Nasional — Jalan Tol Data Indonesia

API Gateway adalah “gerbang resmi” untuk mengakses data pemerintah. Tanpa gateway, akses data seperti jalan kampung: kecil, lambat, rawan macet.

Peran API Gateway Nasional:

- Menyatukan seluruh akses data antar K/L/D.
- Menentukan siapa yang boleh mengakses apa.
- Melakukan *rate limiting*, audit, logging, dan keamanan.
- Menghilangkan API liar yang dibuat tanpa standar.

Keuntungan untuk pemerintah:

- Data tidak perlu dikirim berkali-kali.
- Instansi tidak perlu saling meminta file manual.
- Data bisa diambil kapan saja dan selalu terkini.

Contoh Open Source yang relevan:

- **Kong Gateway** (populer, stabil, mendukung skala nasional).
- **WSO2 API Manager** (banyak dipakai sektor publik).
- **KrakenD** (cukup ringan untuk PDN/PDD).

Standar JSON Nasional & Schema Versioning

Pertukaran data mustahil berjalan tanpa format baku. Selama kabupaten, provinsi, dan pusat memakai format berbeda, integrasi tidak akan pernah selesai.

Standar JSON Nasional memastikan:

- Semua data memakai format yang sama.
- Nama kolom konsisten di seluruh Indonesia.
- Data mudah divalidasi secara otomatis.
- Bug dan error bisa dicegah sebelum data dikirim.

Schema Versioning menghindari kekacauan versi:

- Versi 1 → untuk tahun ini.
- Versi 2 → untuk skema yang diperbarui.
- Sistem tetap kompatibel ke belakang (*backward-compatible*).

Contoh Open Source untuk validasi JSON:

- **JSON Schema Validator** (node/python/go)
- **OpenAPI/Swagger** untuk dokumentasi skema
- **Ajv** untuk validasi cepat di gateway

Manfaat langsung:

Tidak ada lagi “file ini tidak bisa dibaca” atau “formatnya beda” pada saat integrasi.

Message Bus, ESB, Batch vs Real-Time

Pertukaran data tidak cukup dengan API saja. Indonesia butuh mekanisme *push*, bukan hanya *pull*.

Message Bus — jalur cepat data real-time

- Mengalirkan data seperti aliran listrik.
- Cocok untuk data transaksi, kesehatan, bencana, bansos.
- Mendukung jutaan pesan per detik.

Contoh open source:

- **Apache Kafka** (standar emas pertukaran data besar).
- **RabbitMQ** (stabil untuk beban menengah).

ESB (Enterprise Service Bus) — orkestrator integrasi antar sistem

- Menghubungkan berbagai aplikasi lama (legacy) dengan sistem baru.
- Menerjemahkan format data yang berbeda.
- Sangat cocok untuk kementerian dengan banyak aplikasi warisan.

Contoh open source:

- **WSO2 ESB**
- **Talend Open Studio**

Batch vs Real-Time

- **Batch:** baik untuk laporan harian/mingguan. Murah, sederhana.
- **Real-time:** penting untuk respon cepat (stunting, bencana, kesehatan, bansos).

Kapan dipakai real-time?

- Data gizi anak
- Data rumah sakit
- Data harga pangan
- Data CCTV/IoT
- Data bantuan sosial

Kapan batch cukup?

- Laporan sekolah
- Laporan fiskal
- Rekap kependudukan

Arsitektur Microservices untuk Pemerintah

Aplikasi pemerintah yang besar dan monolitik sangat sulit dipelihara. Microservices membuat sistem lebih fleksibel.

Microservices berarti:

- Setiap layanan kecil dan berdiri sendiri.
- Jika satu layanan error, layanan lain tetap berjalan.
- Lebih cepat dikembangkan dan diperbaiki.
- Bisa diperbarui tanpa shutdown seluruh sistem.

Contoh Microservices yang relevan untuk pemerintah:

- Layanan validasi NIK
- Layanan cek status bansos
- Layanan data sekolah
- Layanan data kesehatan
- Layanan dashboard provinsi

Open-source pendukung microservices:

- **Kubernetes** untuk orkestrasi.
- **Istio / Linkerd** untuk service mesh.
- **Docker** untuk kontainer aplikasi.

Intinya:

Microservices membuat layanan digital pemerintah **lebih cepat, lebih tahan gangguan, dan mudah diperluas**.

Tabel Sistem Pertukaran Data Nasional

Komponen	Fungsi	Risiko jika tidak ada	Dampak jika berjalan	Contoh Open Source
API Gateway Nasional	Gerbang data bersama	API liar, tidak aman	Akses data cepat & aman	Kong, WSO2, KrakenD
Standar JSON & Versioning	Format baku nasional	Data tidak bisa digabung	Integrasi otomatis	JSON Schema, OpenAPI
Message Bus	Data real-time	Data telat tiba	Early warning berjalan	Kafka, RabbitMQ
ESB	Integrasi legacy	Sistem lama tidak terhubung	Interoperabilitas penuh	WSO2 ESB, Talend
Batch vs Real-Time	Mekanisme pertukaran	Data terlambat	Kebijakan responsif	Airflow (batch), Kafka (RT)
Microservices	Arsitektur modern	Sistem besar mudah jatuh	Layanan stabil & scalable	Kubernetes, Istio

Pesan Inti

- National Data Exchange adalah “urat nadi” Data Governance.
- API Gateway Nasional menghapus API liar dan mempercepat akses data.
- Standar JSON + Versioning adalah fondasi interoperabilitas.
- Message Bus membuat data nasional mengalir real-time.
- Microservices membuat aplikasi pemerintah lebih cepat, stabil, dan mudah diperluas.
- Open-source adalah cara paling efisien, aman, dan mandiri untuk membangun semuanya.

Bab 7 — Aplikasi Open Source untuk Infrastruktur Pemerintah

Open source bukan sekadar “gratis”. Open source memberikan tiga keuntungan besar yang sangat penting bagi negara:

- **Mandiri:** tidak bergantung vendor luar negeri.
- **Aman:** kode terbuka, mudah diaudit.
- **Hemat anggaran:** tidak perlu bayar lisensi tahunan.
- **Scalable:** bisa dipakai dari kabupaten kecil sampai kementerian besar.

Dalam konteks Data Governance, open source adalah **jalan tercepat** untuk membangun sistem nasional yang kuat, murah, dan modern.

DATABASE (PostgreSQL, MariaDB, ClickHouse)

Database adalah pusat penyimpanan data pemerintah. Pemerintah tidak boleh bergantung pada teknologi tertutup untuk menyimpan data strategis negara.

PostgreSQL — standar emas untuk data pemerintah

- Stabil, aman, matang, dipakai di seluruh dunia.
- Cocok untuk data kependudukan, bansos, pendidikan, kesehatan.
- Mendukung GIS → ideal untuk analisis wilayah.

Contoh penggunaan:

- Database kemiskinan nasional.
- Data sekolah dan NISN.
- Data stunting berbasis desa.

MariaDB — alternatif cepat dan ringan

- Cocok untuk aplikasi dengan trafik tinggi.
- Bisa dipakai oleh kabupaten/kota dengan server kecil.

Contoh:

- Aplikasi pelayanan publik (perizinan, adminduk).
- Sistem informasi rumah sakit.

ClickHouse — database analitik super cepat

- Dirancang untuk query besar dan laporan real-time.
- Cocok untuk dashboard nasional.

Contoh:

- Laporan harga pangan harian.

- Pengawasan fiskal daerah.

ETL/ELT (n8n, Apache NiFi, Airflow)

ETL/ELT adalah tool untuk **mengambil, membersihkan, mengubah, dan mengirim data** dari daerah → provinsi → pusat.

n8n — sangat mudah untuk pemda

- Antarmuka drag-and-drop.
- Sangat cocok untuk kabupaten/kota dengan SDM terbatas.
- Bisa otomatis ambil data dari Excel/CSV.

Contoh:

- Otomatis kirim data stunting harian ke provinsi.
- Integrasi data UMKM dari berbagai perangkat.

Apache NiFi — kuat untuk data besar dan real-time

- Cocok untuk kementerian atau PDN.
- Bisa menangani aliran data puluhan ribu per detik.

Contoh:

- Aliran data kesehatan real-time.
- Pengiriman data bencana ke pusat.

Apache Airflow — untuk jadwal batch (harian/mingguan)

- Ideal untuk rekap data, laporan rutin, atau data berat.

Contoh:

- Rekap kemiskinan nasional mingguan.
- Laporan pendidikan bulanan.

Dashboard (Superset, Metabase)

Dashboard adalah cara paling cepat bagi pimpinan untuk melihat kondisi lapangan.

Apache Superset — visualisasi tingkat nasional

- Cocok untuk kementerian, provinsi, dan PDN.
- Mendukung user role, audit, dan multi-database.

Contoh:

- Dashboard stunting nasional.
- Dashboard bansos per kabupaten.

Metabase — untuk pemda dan unit kecil

- Sangat mudah dipakai ASN tanpa kemampuan teknis.
- Cocok untuk kabupaten/kota.

Contoh:

- Dashboard pelayanan publik.
- Monitoring sekolah dan Puskesmas.

Big Data (Hadoop, Spark, Kafka)

Digunakan untuk data ukuran besar, cepat, kompleks, dan real-time.

Hadoop — penyimpanan skala besar

- Cocok untuk PDN dan kementerian besar.

Apache Spark — analitik cepat

- Untuk prediksi, data mining, dan AI skala besar.

Apache Kafka — “jalan tol” data real-time

- Cocok untuk early-warning system.

Contoh kombinasi Big Data:

- Data CCTV, IoT, sensor pertanian.
- Monitoring harga pangan setiap menit.
- Deteksi dini penyakit menular.

AI Platform (Kubeflow, HuggingFace, Ollama)

AI membutuhkan infrastruktur khusus. Tanpa platform ini, AI tidak akan jalan di pemerintah.

Kubeflow — pipeline AI pemerintah

- Cocok untuk PDN, PDD, dan kementerian.
- Mengatur pelatihan model, deployment, dan evaluasi.

HuggingFace — pustaka model AI terlengkap

- Ribuan model open source.
- Cocok untuk NLP, klasifikasi dokumen, chatbot layanan publik.

Ollama — LLM lokal di server pemerintah

- Bisa menjalankan model AI tanpa internet.
- Sangat cocok untuk privasi dan keamanan data.

Contoh penggunaan AI open-source:

- Analisis surat masuk pemerintah.
- Chatbot pelayanan desa/kecamatan.
- Prediksi kemiskinan atau stunting.

Cybersecurity (Wazuh, Suricata, pfSense, WireGuard)

Data Governance tidak ada artinya tanpa keamanan siber.

Wazuh — deteksi ancaman

- Pantau server PDN, PDD, dan daerah.

Suricata — IDS/IPS open source

- Menghentikan serangan jaringan.

pfSense — firewall enterprise gratis

- Cocok untuk kementerian dan pemda.

WireGuard — VPN aman antar-instansi

- Kecepatan tinggi, enkripsi modern.

Contoh:

- Semua akses API antar K/L melalui VPN terenkripsi.
- Monitoring PDN/PDD 24 jam.

Tabel Open Source untuk Pemerintah

Kategori	Fungsi	Masalah yang Diselesaikan	Contoh Open Source	Cocok Untuk
Database	Penyimpanan data	Data terpisah, lambat	PostgreSQL, MariaDB, ClickHouse	Pusat & daerah
ETL/ELT	Alur data	Data manual, lambat	n8n, NiFi, Airflow	Kab/Kota → Pusat
Dashboard	Visualisasi	Laporan sulit dibaca	Superset, Metabase	Pimpinan
Big Data	Data besar & real-time	API overload	Hadoop, Spark, Kafka	PDN / Kementerian
AI Platform	Prediksi & otomasi	Analisis lambat	Kubeflow, HuggingFace, Ollama	Nasional
Cyber Security	Keamanan data	Kebocoran & serangan	Wazuh, Suricata, pfSense, WireGuard	Semua level

Pesan Inti

- Open source adalah fondasi infrastruktur data pemerintah yang kuat, murah, dan mandiri.
- PostgreSQL, Superset, Kafka, Kubeflow, Ollama, dan Wazuh sudah cukup untuk membangun **Data Governance Nasional**.
- Seluruh aplikasi ini bisa berjalan dari **kabupaten kecil** sampai **Pusat Data Nasional** tanpa biaya lisensi mahal.
- Inilah teknologi yang memungkinkan Indonesia membangun **negara digital yang berdaulat**.

Bab 8 — Contoh Implementasi Nyata Open-Source untuk Pemerintah

Bab 8 adalah bukti bahwa teknologi open-source **bukan teori**, tetapi **sudah sangat siap** dipakai oleh pemerintah. Semua contoh ini bisa diterapkan di:

- Kabupaten/kota
- Provinsi
- Kementerian
- PDN / Pusat Data Daerah
- Balai dan UPTD

Biayanya rendah, SDM tidak perlu ahli, dan implementasinya cepat.

ETL Bansos Daerah → Pusat Menggunakan n8n

Masalah klasik bansos:

- Data dari kelurahan/kecamatan tidak rapi.
- Pengiriman manual (Excel/WA/email).
- Pusat menerima data terlambat.
- Banyak data ganda atau tidak valid.

Solusi: n8n sebagai tool ETL otomatis.

Cara kerja sederhana n8n untuk bansos:

- n8n mengambil data dari file Excel/CSV atau API daerah.
- Membersihkan data (hapus duplikasi, cek NIK, cek alamat).
- Mencocokkan dengan DTKS dan database pusat.
- Mengirim data otomatis ke server provinsi atau pusat.

Manfaat:

- Tidak ada lagi input manual.
- Validasi otomatis.
- Data bansos update harian (bukan bulanan).
- Mengurangi salah sasaran.

Open-source:

n8n (gratis, mudah, bisa dipasang di server kecil pemda)

Dashboard Kesehatan & Stunting Berbasis Superset

Superset adalah tool dashboard open-source yang sangat cocok untuk laporan kesehatan.

Masalah di daerah:

- Data stunting tersebar di Puskesmas.
- Laporan harus rekap manual.
- Angka sering berbeda antara kabupaten dan provinsi.
- Pimpinan tidak punya dashboard real-time.

Solusi: Superset untuk visualisasi kesehatan.

Contoh dashboard yang bisa dibuat:

- Grafik jumlah balita stunting per desa.
- Tren penimbangan bulanan.
- Peta wilayah rawan stunting.
- Monitoring Puskesmas yang belum lapor.

Manfaat:

- Bupati/Walikota bisa melihat kondisi harian.
- Provinsi bisa memantau kabupaten secara real-time.
- Data kesehatan cepat dianalisis untuk intervensi.

Open-source:

Apache Superset (dashboard nasional dan daerah)

Data Warehouse Pendidikan Menggunakan PostgreSQL

Pendidikan punya data sangat besar:

siswa, guru, rombel, nilai, sekolah, pemetaan wilayah, absensi, bantuan PIP, dll.

Masalah yang sering muncul:

- Data berserakan.
- Banyak versi.
- Sering terjadi perbedaan angka antara kabupaten dan provinsi.

Solusi: Data warehouse pendidikan berbasis PostgreSQL.

Cara kerja:

- Semua data pendidikan dimasukkan ke PostgreSQL.
- Dibangun tabel indikator seperti:
 - jumlah siswa
 - jumlah guru
 - angka putus sekolah
 - rombel aktif
- Data bersih dan konsisten untuk dashboard provinsi/pusat.

Manfaat:

- Laporan PIP, BOS, dan PPDB jauh lebih cepat.
- Data dapat digunakan untuk prediksi (AI).
- Provinsi bisa memantau kualitas sekolah secara objektif.

Open-source:

PostgreSQL (stabil, aman, bebas lisensi)

Tools tambahan: pgAdmin, PostGIS

AI Lokal dengan Ollama di Server Pemda

AI tidak harus memakai cloud luar negeri. Pemerintah bisa menjalankan AI secara **lokal**, tanpa data keluar. Ollama adalah platform open-source untuk menjalankan LLM lokal.

Apa yang bisa AI lokal lakukan di pemda?

- Membaca dan merangkum surat masuk.
- Membuat draft jawaban otomatis.
- Melakukan klasifikasi aduan masyarakat.
- Mengambil insight dari dokumen PDF.
- Membuat rekap laporan kecamatan.

Mengapa Ollama penting?

- **Data tidak keluar dari pemda** → aman.
- Bisa berjalan di server biasa dengan GPU standar.
- Sangat cepat dan mudah dipasang.

Contoh aplikasi AI pemda:

- Chatbot layanan kelurahan/kecamatan.
- AI untuk membaca laporan Puskesmas.
- AI untuk analisis pengaduan via WhatsApp/Email.

Open-source:

Ollama (LLM lokal)

Model open-source seperti Llama3, Phi3, Mistral

Monitoring Keamanan Wazuh untuk Kabupaten/Kota

Banyak pemda tidak punya tool monitoring keamanan. Padahal ancaman siber meningkat.

Solusi: Wazuh sebagai pusat deteksi ancaman.

Apa yang dilakukan Wazuh?

- Mendeteksi serangan ke server pemda.
- Memberi peringatan jika ada aktivitas mencurigakan.
- Memantau log dari aplikasi dan API.
- Memberi rekomendasi keamanan.

Contoh implementasi:

- Monitoring server Disdukcapil.
- Monitoring API bansos.
- Monitoring aplikasi perizinan daerah.
- Monitoring PDN/PDD node daerah.

Manfaat:

- Pemda tidak buta terhadap serangan siber.
- Insiden dapat ditangani lebih cepat.
- Keamanan data masyarakat meningkat drastis.

Open-source:

Wazuh (SIEM & XDR open-source)

Suricata (IDS/IPS tambahan)

Tabel Contoh Implementasi Open-Source

Implementasi	Masalah yang Dipecahkan	Dampak Cepat	Contoh Open-Source
ETL Bansos dengan n8n	Data ganda, kirim manual	Data bersih & otomatis	n8n
Dashboard Stunting Superset	Laporan lambat	Monitoring real-time	Superset
Data Warehouse Pendidikan	Angka tidak konsisten	Laporan cepat & rapi	PostgreSQL
AI Lokal Pemda	Surat & dokumen menumpuk	Analisis cepat, otomatis	Ollama, LLM OS
Monitoring Keamanan Wazuh	Serangan tidak terdeteksi	Sistem pemda lebih aman	Wazuh, Suricata

Pesan Inti

- Seluruh contoh di bagian ini **sudah bisa dipakai hari ini** tanpa biaya lisensi.
- Open-source memungkinkan pemda/kementerian membangun sistem cepat, murah, dan aman.
- n8n, Superset, PostgreSQL, Ollama, dan Wazuh adalah **starter pack nasional** untuk Data Governance.
- Inilah bukti bahwa teknologi open-source dapat mempercepat transformasi digital Indonesia **tanpa menunggu anggaran besar**.

BAGIAN IV — PENGUMPULAN DATA NASIONAL & INTEROPERABILITAS

Bab 9 — Strategi Pengumpulan Data dari Daerah → Provinsi → Nasional

Data nasional tidak akan pernah akurat jika data daerah kacau. Masalah terbesar pengumpulan data hari ini bukan teknologi, tetapi **fragmentasi**:

- Kabupaten/kota menggunakan aplikasi berbeda.
- Provinsi punya format lain.
- Pusat menerima data “berbeda versi” dari tiap daerah.
- Pengiriman data masih banyak memakai Excel, WA, dan email.

Akibatnya:

- Laporan lambat.
- Angka pusat dan daerah tidak cocok.
- Kebijakan nasional salah sasaran.

Bagian ini menjelaskan bagaimana membangun **alur data yang rapi, standar, dan otomatis** dari daerah → provinsi → pusat.

Model Sentralisasi vs Desentralisasi

Tidak ada model tunggal yang cocok untuk semua. Pemerintah perlu menggabungkan dua pendekatan ini secara seimbang.

A. Model Sentralisasi

Semua data dikumpulkan di server pusat.

Kelebihan:

- Standar sangat kuat.
- Pusat mudah mengendalikan kualitas data.
- Lebih cocok untuk program nasional (bansos, kesehatan, pendidikan).

Risiko:

- Beban jaringan besar.
- Jika pusat bermasalah, seluruh daerah terdampak.
- Inovasi daerah lebih sulit.

B. Model Desentralisasi

Daerah mengelola sistem masing-masing, kemudian mengirim data ke provinsi/pusat.

Kelebihan:

- Lebih fleksibel.
- Daerah bisa berinovasi sesuai konteks lokal.
- Beban pusat lebih ringan.

Risiko:

- Format data bisa berantakan.
- Validasi sulit jika tidak ada standar JSON nasional.
- Perbedaan sistem antar daerah menghambat sinkronisasi.

Model terbaik untuk Indonesia: Hibrida (Hybrid)

- **Desentralisasi pada input data** (kecamatan/kabupaten tetap bebas menggunakan aplikasi).
- **Sentralisasi pada standarisasi, validasi, dan penyimpanan nasional.**

Teknologi Open Source Pendukung:

- **OpenAPI + JSON Schema** → standarisasi format.
- **Kafka** → pengiriman data real-time pusat–daerah.
- **n8n / NiFi** → ETL otomatis daerah ke provinsi.

Validasi, Verifikasi, dan Deduplication (Kab/Kota → Provinsi → Pusat)

Tiga langkah ini adalah “filter nasional” supaya data tidak kotor.

Validasi (cek format)

Daerah mengirim data → sistem otomatis memeriksa apakah:

- Format sesuai JSON Nasional?
- Semua kolom wajib terisi?
- Tipe data benar?

tool Open-Source:

JSON Schema Validator, Ajv, OpenAPI Validator

Verifikasi (cek kebenaran data)

Provinsi memeriksa apakah:

- NIK valid?
- Lokasi benar?
- Puskesmas/sekolah terdaftar?

Contoh: Cross-check data stunting dengan master data provinsi.

tool Pendukung:

PostgreSQL + stored procedures

Apache Superset untuk memantau anomali

Deduplication (hapus data ganda)

Masalah terbesar:

- Balita muncul dua kali.
- Peserta bansos terdaftar ganda.
- Siswa tercatat di dua sekolah.

Solusi open source:

Python + Pandas untuk duplikasi dasar

Elasticsearch + fuzzy matching untuk duplikasi tingkat lanjut

n8n/NiFi untuk otomatisasi

Hasil akhirnya:

- Data provinsi bersih.
- Data pusat konsisten.
- Tidak ada lagi “angka versi daerah” dan “angka versi pusat”.

Arsitektur Pelaporan Real-Time

Pelaporan tidak boleh lagi berbasis file manual. Harus:

- Terstandar
- Otomatis
- Real-time atau near real-time

Komponen Real-Time Nasional:

1. **API Gateway Nasional**
 - Jalur resmi pertukaran data.
2. **Message Bus (Kafka/RabbitMQ)**
 - Mengalirkan data secara terus-menerus.
3. **ETL otomatis (n8n / NiFi)**
 - Membersihkan dan memvalidasi data.
4. **Data Warehouse Nasional**
 - PostgreSQL/ClickHouse.
5. **Dashboard pimpinan (Superset)**
 - Laporan real-time untuk menteri/gubernur.

Contoh Real-Time:

- Puskesmas upload data → 30 detik → dashboard Gubernur.
- Kecamatan upload data bansos → otomatis masuk ke provinsi.
- Sensor harga pangan pasar → masuk Kafka → dianalisis di pusat.

Semua bisa dibangun tanpa lisensi mahal.

Integrasi Sistem Daerah Warisan (Legacy)

Banyak daerah masih menggunakan:

- Aplikasi lama berbasis PHP jadul.
- Database MySQL lawas tanpa standar.
- File Excel puluhan ribu baris.
- Sistem vendor yang tidak bisa diekspor.

Tidak mungkin memaksa daerah "membuang" sistem lama. Yang harus dilakukan adalah **mengintegrasikan, bukan mengganti**.

Strategi integrasi yang paling realistis:

1. **Buat API wrapper**
 - Bungkus aplikasi lama dengan API baru.
2. **Gunakan ETL (NiFi/n8n)**
 - Ambil data dari database lama, ubah formatnya, kirim ke provinsi/pusat.
3. **Mapping data otomatis**
 - Pakai JSON Schema Nasional sebagai template.
4. **Pindah bertahap, bukan sekaligus**
 - Legacy tetap berjalan sambil data bertahap masuk sistem nasional.

Open-source pendukung integrasi Legacy → Modern:

- Apache NiFi
- n8n
- Logstash
- PostgreSQL FDW
- Talend Open Studio

Prinsip penting:

Integrasi dulu → konsolidasi kemudian → modernisasi terakhir.

Tabel Strategi Pengumpulan Data

Komponen	Masalah yang Dipecahkan	Dampak untuk Pemerintah	Contoh Open Source
Sentralisasi vs Desentralisasi	Data tidak sinkron	Struktur nasional yang efisien	OpenAPI, JSON Schema
Validasi	Format data berantakan	Data bisa diproses otomatis	JSON Validator, Ajv
Verifikasi	Data salah/invalid	Data akurat untuk kebijakan	PostgreSQL, Superset
Deduplication	Data ganda	Satu warga = satu data	Pandas, Elasticsearch
Pelaporan Real-Time	Laporan lambat	Dashboard nasional real-time	Kafka, n8n, Superset
Integrasi Legacy	Aplikasi daerah tidak terhubung	Data daerah ikut standar nasional	NiFi, Logstash, FDW

Pesan Inti

- Pengumpulan data dari daerah adalah fondasi Data Governance Indonesia.
- Kunci utama: **standar JSON**, validasi otomatis, ETL, dan real-time.
- Integrasi tidak boleh memaksa daerah membuang sistem lama — cukup dibungkus API atau ETL.
- Open source seperti **Kafka, n8n, NiFi, Superset, PostgreSQL, Wazuh** sudah cukup untuk membangun sistem nasional yang kuat.
- Dengan strategi yang benar, pusat, provinsi, dan kabupaten bisa memiliki **satu data nasional** tanpa biaya lisensi mahal.

Bab 10 — Mekanisme Data Sharing antar Kementerian & Lembaga

Masalah terbesar Data Governance Indonesia **bukan teknologi**, tetapi **kerjasama antar instansi**. Gejala di lapangan:

- Kementerian memiliki data masing-masing, tidak mau berbagi.
- Aplikasi dibangun terpisah, standar berbeda-beda.
- Lembaga ragu membuka data karena khawatir disalahgunakan.
- Pemda tidak tahu data mana yang boleh dibagi.
- Pusat kesulitan mengumpulkan data cepat untuk kebijakan nasional.

Padahal, tanpa *data sharing*, tidak mungkin ada:

- Satu Data Indonesia
- Dashboard nasional
- AI nasional
- Keputusan berbasis bukti
- Efisiensi anggaran

Karena itu, bagian ini menjelaskan mekanisme kerja sama yang **terukur, aman, dan saling menguntungkan**.

Insentif Berbagi Data — Cara Paling Efektif Menghilangkan Ego-Sektoral

Instansi akan berbagi data **hanya jika ada manfaat nyata**.

Insentif efektif untuk K/L/D:

- **Akses balik ke data nasional**
→ K/L yang berbagi data bisa mengakses dashboard nasional.
- **Prioritas anggaran digital**
→ K/L yang aktif berbagi data mendapat prioritas DAK/ICT.
- **Pengurangan beban pelaporan**
→ Jika data sudah otomatis, tidak perlu kirim laporan manual.
- **Skema penghargaan**
→ Ranking keterbukaan data antar K/L/Provinsi.
- **Audit & akuntabilitas lebih ringan**
→ Jika data sudah masuk nasional, audit jadi lebih mudah.

Insentif berbasis teknologi (open source):

- API Gateway Nasional memberikan **akses terbatas namun otomatis**.
- Dashboard berbasis Superset memberikan **keuntungan visual langsung**.
- PDN menyediakan **hosting gratis untuk data instansi**.

Instansi akan berbagi data **ketika mereka merasakan manfaat langsung**.

Trust Framework Nasional — Membangun Kepercayaan Antar Instansi

Data sharing tidak akan berjalan kalau instansi saling curiga. Karena itu harus ada **trust framework**.

Elemen penting Trust Framework Nasional:

1. **Legal**
 - Ada PKS/SOP yang jelas antara instansi.
 - Ada batasan data yang boleh/tidak boleh dibagi.
2. **Teknis**
 - Akses data dibatasi role-based access control (RBAC).
 - API Gateway Nasional mencatat semua akses (*audit trail*).
 - Data terenkripsi saat transit (WireGuard/OpenVPN).
3. **Keamanan**
 - Wazuh memantau aktivitas server.
 - Suricata mendeteksi serangan.
 - Incident response jelas ketika terjadi pelanggaran.
4. **Kualitas Data**
 - Data harus lewat pipeline validasi–verifikasi–deduplication.
 - JSON Schema Nasional mengontrol standar format.

Open source pendukung Trust Framework:

- **Keycloak** untuk Single Sign-On & RBAC antar K/L
- **Wazuh** untuk keamanan & audit log
- **Kong API Gateway** untuk kontrol akses
- **WireGuard** untuk VPN terenkripsi

Tanpa kepercayaan, tidak ada data sharing.

Tanpa mekanisme, kepercayaan tidak akan pernah terbentuk.

Federated Data Governance — Cara Praktis Menyatukan Data Tanpa Menyatukan Sistem

Indonesia terlalu besar dan kompleks untuk satu sistem tunggal. Maka yang realistis adalah **Federated Data Governance**.

Prinsip Federasi Data:

- Data tetap berada di kementerian/lembaga/pemda masing-masing.
- Tetapi metadata, standar JSON, dan API mengikuti standar nasional.
- Instansi lain tidak mengambil data mentah, namun **mengakses data terpilih** melalui API Gateway.
- Setiap instansi tetap punya otonomi dan kendali.

Keuntungan untuk pemerintah:

- Tidak perlu memaksa semua instansi memakai aplikasi yang sama.
- Tidak memakan biaya migrasi data yang sangat besar.
- Daerah tetap bisa mempertahankan sistem lama (legacy).
- Integrasi lebih cepat dan tidak mengganggu operasional.

Open source untuk federasi data:

- **OpenAPI** untuk standardisasi API.
- **Superset** untuk dashboard gabungan lintas instansi.
- **NiFi** untuk federated data flows.
- **ClickHouse** sebagai federated query engine.

Federated Data Governance = **jalan tengah antara kedaulatan instansi dan interoperabilitas nasional**.

Konsep “Data Exchange Day” — Forum Operasional untuk Koordinasi Data Nasional

Banyak instansi tidak tahu:

- Siapa memegang data apa
- Format data apa yang dipakai
- Akses apa yang diperbolehkan
- Standar terbaru yang harus diikuti

Solusinya: **Data Exchange Day (DED)**.

Apa itu Data Exchange Day?

Forum rutin (bulanan/kuartalan) untuk:

- Sinkronisasi antar K/L dan provinsi
- Menyelesaikan masalah integrasi data
- Menetapkan perubahan schema JSON
- Menyepakati endpoint API baru
- Membahas insiden data dan perbaikan kualitas
- Menyelaraskan kebutuhan kebijakan dengan data

Mengapa efektif?

- Menghilangkan misinformasi.
- Mengurangi saling menyalahkan antara K/L/D.
- Mempercepat integrasi karena seluruh pihak hadir.
- Menjadikan Data Governance sebagai *aktivitas rutin*, bukan proyek.

Contoh materi DED:

- Update schema JSON bansos
- Masalah deduplication NIK
- Integrasi data kesehatan provinsi
- Pemanfaatan Superset untuk dashboard lintas instansi
- Rencana penggunaan Kafka untuk streaming data

Tabel Mekanisme Data Sharing

Komponen	Masalah yang Dipecahkan	Dampak Praktis	Contoh Open Source
Insentif Berbagi Data	Ego sektoral	Instansi mau berbagi	Superset, API gateway
Trust Framework Nasional	Ketidakpercayaan antar K/L	Data sharing aman & legal	Keycloak, Kong, WireGuard
Federated Data Governance	Sistem berbeda-beda	Integrasi tanpa migrasi	NiFi, OpenAPI, ClickHouse
Data Exchange Day	Koordinasi lemah	Sinkronisasi rutin	Superset + JSON Schema as agenda

Pesan Inti

- Data sharing adalah kunci keberhasilan Satu Data Indonesia dan SPBE.
- Insentif, bukan paksaan, adalah cara paling efektif menghilangkan ego-sektoral.
- Trust Framework Nasional membuat data sharing aman dan legal.
- Federated Data Governance memungkinkan integrasi tanpa mematikan sistem daerah.
- Data Exchange Day membuat koordinasi data menjadi *rutinitas*, bukan proyek ad hoc.
- Semua bisa berjalan dengan teknologi open source yang sudah matang dan murah.

BAGIAN V — DATA ANALYTICS, AI, & FORECASTING KEBIJAKAN

Bab 11 — Dasar-Dasar Forecasting untuk Pemerintah

Forecasting adalah “ramalan berbasis data”. Namun bukan ramalan sembarangan. Forecasting menggunakan:

- data historis,
- pola perilaku,
- tren perubahan,
- indikator sosial-ekonomi,

untuk memprediksi apa yang akan terjadi. Pemerintah membutuhkan forecasting agar **tidak hanya bereaksi**, tetapi **bersiap sebelum masalah muncul**.

Tanpa forecasting:

- anggaran sering terlambat,
- intervensi tidak tepat sasaran,
- program gagal dicegah sebelum terjadi,
- keputusan menjadi spekulatif.

Dengan forecasting:

- kebijakan lebih proaktif,
- anggaran lebih hemat,
- target lebih realistis,
- risiko bisa ditangani sebelum membesar.

Forecast Kesejahteraan Rakyat, Ekonomi, dan Pendidikan

Forecasting di tiga area ini sangat penting karena dampaknya langsung ke rakyat:

A. Kesejahteraan Rakyat

Digunakan untuk memprediksi:

- jumlah rumah tangga miskin,
- perubahan pengeluaran rumah tangga,
- kebutuhan bantuan sosial,
- wilayah yang mulai “turun kelas” (kerentanan meningkat).

Open-source tools:

- **Prophet** (Facebook open-source forecasting)
- **Python Pandas + Scikit-Learn**
- **PostgreSQL + TimescaleDB** untuk time-series

Contoh implementasi:

- Prediksi jumlah penerima bansos 6 bulan ke depan.
- Forecast kerentanan desa untuk perencanaan DAK.

B. Ekonomi Daerah

Dipakai untuk memprediksi:

- harga pangan,
- supply–demand komoditas,
- inflasi lokal,
- pendapatan daerah,
- aktivitas pasar dan UMKM.

Open-source pendukung:

- **ClickHouse** untuk analitik cepat.
- **Apache Spark** untuk prediksi ekonomi skala besar.

C. Pendidikan

Forecasting membantu:

- memprediksi angka putus sekolah,
- kebutuhan guru,
- jumlah murid baru,
- beban PPDB,
- kebutuhan ruang kelas.

Contoh implementasi:

- Provinsi dapat menghitung kebutuhan guru hingga 2030.
- Kabupaten dapat melihat desa mana yang rawan dropout.

Early Warning System Berbasis Data

Early Warning System (EWS) adalah sistem peringatan dini. Tujuannya sederhana: **beri alarm sebelum masalah besar terjadi.**

EWS penting untuk:

- stunting
- penyakit menular
- inflasi pangan
- kemiskinan ekstrem
- kegagalan panen
- lonjakan harga barang
- dropout siswa
- banjir atau hotspot kebakaran

Cara kerja EWS berbasis open-source:

1. **Kumpulkan data real-time**
 - Kafka, RabbitMQ (streaming data)
 - n8n, NiFi (ETL otomatis)
2. **Analisis pola abnormal**
 - Spark, Pandas, TensorFlow Lite
 - PostgreSQL + triggers
3. **Tampilkan peringatan**
 - Superset dashboard
 - Notifikasi Telegram/WA otomatis via n8n
4. **Tindak lanjut cepat**
 - Kirim peringatan ke kecamatan/Puskesmas

Contoh “alarm data”:

- Berat badan bayi tidak naik dalam 2 bulan → alarm stunting.
- Harga beras naik 5 hari berturut-turut → alarm inflasi.
- Absensi siswa menurun drastis → alarm dropout.
- Kasus DBD meningkat tiba-tiba → alarm kesehatan.

EWS membuat pemerintah **tidak hanya melihat laporan masa lalu**, tetapi **bergerak berdasarkan indikasi masa depan.**

Value for Money (VfM) Berbasis Data

Value for Money menjawab pertanyaan kritis:

“Apakah anggaran yang dikeluarkan menghasilkan manfaat maksimal?”

VfM fokus pada tiga hal:

A. Economy

Apakah barang/jasa dibeli dengan harga paling efisien? Contoh indikator:

- biaya per murid,
- biaya per balita ditimbang,
- biaya per penerima bansos.

B. Efficiency

Apakah input → output berjalan efisien? Contoh:

- rasio guru–murid,
- jumlah Puskesmas per kecamatan,
- kecepatan distribusi bantuan.

C. Effectiveness

Apakah program mencapai hasilnya? Contoh:

- penurunan stunting,
- peningkatan angka sekolah,
- penurunan kemiskinan.

Open-source pendukung VfM:

- **Metabase / Superset**: dashboard monitoring anggaran vs hasil
- **PostgreSQL**: data penganggaran + realisasi
- **Python (Pandas)**: menghitung efisiensi dan efektivitas
- **Grafana**: monitoring operasional

Manfaat VfM berbasis data:

- Mengurangi program yang tidak berdampak.
- Mengarahkan ulang anggaran ke kegiatan yang memberi manfaat terbesar.
- Menjaga akuntabilitas pemerintah kepada publik.

Tabel Dasar Forecasting & EWS

Komponen	Fungsi Utama	Dampak Praktis	Open-Source Pendukung
Forecast kesejahteraan	Prediksi kemiskinan & bansos	Bantuan tepat sasaran	Prophet, Pandas, PostgreSQL
Forecast ekonomi	Prediksi inflasi, harga pangan	Stabilitas pasar	Spark, ClickHouse
Forecast pendidikan	Prediksi dropout & kebutuhan guru	Intervensi cepat	Pandas, TimescaleDB
EWS berbasis data	Alarm sebelum masalah besar	Respons cepat	Kafka, n8n, Superset
Value for Money	Ukur manfaat anggaran	Efisiensi fiskal	Metabase, Grafana, PostgreSQL

Pesan Inti

- Forecasting membantu pemerintah **melihat masa depan sebelum terlambat**.
- Early Warning System membuat pemerintah lebih cepat merespons masalah rakyat.
- Value for Money memastikan anggaran benar-benar memberikan dampak.
- Semua ini bisa dibangun dengan **open-source murah dan aman** seperti Prophet, Superset, Kafka, PostgreSQL, dan n8n.
- Forecasting bukan hal rumit — ini adalah tool sederhana untuk meningkatkan **kesejahteraan, pendidikan, dan stabilitas ekonomi**.

Bab 12 — Pemanfaatan AI untuk Klasifikasi & Prediksi

AI bukan teknologi futuristik. AI adalah tool untuk membantu pemerintah **membaca pola** dan **mempercepat keputusan**.

Dalam konteks Indonesia:

- Data besar tidak mungkin dianalisis manual.
- Banyak masalah rakyat membutuhkan deteksi cepat.
- Anggaran terbatas harus diarahkan ke tempat yang benar.

AI membantu pemerintah tahu:

- *Siapa yang harus dibantu lebih dulu,*
- *Daerah mana paling rawan,*
- *Program mana yang akan gagal bila tidak diintervensi,*
- *Berapa anggaran yang sebenarnya dibutuhkan,*
- *Wilayah mana yang akan mengalami tekanan ekonomi atau pendidikan.*

Semua itu bisa dicapai dengan **AI open-source murah**, tanpa ketergantungan vendor luar negeri.

Klasifikasi Kelompok Rentan & Wilayah Prioritas

Klasifikasi adalah kemampuan AI untuk **mengelompokkan** masyarakat atau wilayah berdasarkan pola data.

Klasifikasi kelompok rentan

AI dapat mendeteksi:

- rumah tangga yang turun kesejahteraannya,
- balita yang berisiko stunting,
- siswa yang berisiko dropout,
- UMKM yang rawan tutup,
- warga lansia yang membutuhkan intervensi cepat.

Data yang digunakan (contoh):

- NIK
- status pekerjaan
- pengeluaran rumah tangga
- data kesehatan dasar
- catatan Puskesmas
- absensi sekolah
- data survei desa

Klasifikasi wilayah prioritas

AI dapat mengelompokkan wilayah:

- desa rawan gizi
- kecamatan rentan inflasi
- kelurahan rawan kriminalitas
- sekolah berisiko banyak dropout
- zona harga pangan tidak stabil

Hasilnya:

Pimpinan daerah langsung tahu **wilayah mana yang butuh intervensi dulu**, tanpa harus menunggu laporan menumpuk.

Contoh open-source untuk klasifikasi:

- **Scikit-Learn**
- **XGBoost**
- **LightGBM**
- **HuggingFace Transformers** (untuk klasifikasi dokumen)
- **Ollama** untuk klasifikasi teks berbahasa Indonesia secara lokal

Prediksi Anggaran, Dropout, Stunting, UMKM Workload

AI dapat memprediksi kondisi masa depan, sehingga pemerintah bisa bergerak lebih awal.

A. Prediksi Anggaran

AI membantu memperkirakan:

- alokasi bansos,
- biaya kesehatan,
- kebutuhan guru,
- beban Puskesmas,
- anggaran stunting per desa,
- kebutuhan cadangan pangan.

Contoh open-source:

- **Prophet (Facebook)**
- **StatsModels**
- **Spark MLlib** untuk skala nasional

B. Prediksi Dropout Sekolah

AI bisa melihat:

- pola absensi,
- nilai rapor,
- kondisi ekonomi keluarga,
- data transportasi sekolah,
- jarak rumah ke sekolah.

AI memunculkan skor risiko dropout untuk setiap siswa.

Manfaat:

Dinas pendidikan bisa intervensi **sebelum** siswa berhenti sekolah.

C. Prediksi Stunting

AI membaca:

- berat badan
- tinggi badan
- usia
- data imunisasi
- faktor keluarga
- sanitasi

Lalu memprediksi:

- siapa anak yang berisiko mengalami stunting bulan depan,
- desa mana yang perlu pemantauan ekstra.

Ini jauh lebih cepat daripada menunggu angka tahunan.

D. Prediksi Beban UMKM / Workload

AI memprediksi:

- UMKM yang rawan gagal,
- kebutuhan modal,
- kenaikan/penurunan omzet,
- potensi pasar berdasarkan pola konsumsi.

Open-source tools:

- **Python Pandas, Scikit-Learn**
- **ClickHouse** untuk analitik cepat
- **Kafka** untuk data real-time pasar

Studi Kasus GovAI Global — Apa yang Bisa Kita Tiru?

Beberapa negara telah berhasil menerapkan Artificial Intelligence dalam penyelenggaraan pemerintahan (Government AI/GovAI). Berikut contoh nyata yang dapat menjadi referensi langsung bagi Indonesia.

Singapura – Deteksi Dini Kebutuhan Sosial

- AI digunakan untuk memprediksi rumah tangga yang berisiko gagal membayar sewa rumah (ComLink programme oleh MSF).
- Petugas sosial dapat melakukan intervensi dini sebelum terjadi penggusuran atau krisis.
- Teknologi yang dipakai: Python, Apache Spark, dan dashboard Apache Superset.

Inggris – Kesehatan Masyarakat (NHS)

- National Health Service (NHS) menggunakan AI untuk mendeteksi dini risiko penyakit kronis (misalnya diabetes, penyakit jantung).
- Data real-time dari berbagai sumber diproses menggunakan Apache Kafka dan pipeline machine learning.

Tiongkok – Perencanaan Transportasi dan Kota

- Kota-kota besar seperti Shenzhen dan Hangzhou menggunakan AI untuk memprediksi kepadatan penduduk serta beban sistem transportasi umum.
- Sistem berbasis teknologi open-source dengan analisis data real-time untuk optimalisasi rute dan jadwal.

Chile – Pencegahan Putus Sekolah

- Kementerian Pendidikan Chile mengembangkan model AI yang memprediksi siswa berisiko putus sekolah (dropout).
- Intervensi berbasis data (kunjungan guru, bantuan khusus) berhasil meningkatkan angka kelulusan nasional secara signifikan.

Kesimpulan dan Pelajaran Praktis untuk Indonesia

1. Prioritaskan penyelesaian masalah nyata masyarakat, bukan sekadar membangun proyek teknologi.
2. Mulai dari data yang sudah dimiliki pemerintah (tidak perlu menunggu data sempurna).
3. Manfaatkan teknologi open-source untuk menekan biaya pengembangan dan pemeliharaan.
4. AI bukan pengganti pegawai negeri (ASN), melainkan tool bantu agar ASN dapat mengambil keputusan lebih cepat dan tepat sasaran.

Tabel GovAI

Topik	Fungsi Utama	Dampak Cepat	Open-Source Pendukung
Klasifikasi Kelompok Rentan	Mengelompokkan warga/wilayah	Intervensi lebih tepat	Scikit-Learn, XGBoost, Ollama
Prediksi Anggaran	Forecast kebutuhan anggaran	Efisiensi fiskal	Prophet, Spark MLlib
Prediksi Dropout	Deteksi siswa berisiko	Pencegahan dini	Pandas, LightGBM
Prediksi Stunting	Deteksi balita rawan	Intervensi lebih cepat	Scikit-Learn, TensorFlow Lite
Prediksi UMKM Workload	Analisis risiko & omzet	Dukungan UMKM lebih tepat	ClickHouse, Kafka
Studi Kasus GovAI	Contoh global	Inspirasi implementasi cepat	Kombinasi Python + open-source

Pesan Inti

- AI membantu pemerintah *melihat apa yang tidak terlihat oleh laporan biasa*.
- AI bisa mengklasifikasi warga rentan, menentukan wilayah prioritas, dan memprediksi masalah sebelum terjadi.
- Prediksi anggaran, stunting, dropout, dan UMKM kini bisa dilakukan dengan **open-source murah dan aman**.
- GovAI global membuktikan bahwa negara modern bergerak dari “melihat masa lalu” → “mengantisipasi masa depan”.
- Indonesia harus mulai menggunakan AI bukan sebagai proyek, tetapi sebagai **tool kerja harian** pemerintah.

Bab 13 — Arsitektur AI Berjenjang Republik Indonesia

Indonesia adalah negara besar: 514 kabupaten/kota, 34 provinsi, lebih dari 83.000 desa/kelurahan. Tidak mungkin satu model AI nasional menangani semuanya.

Masalah hari ini:

- Data desa tidak langsung bisa dipakai pusat.
- AI pusat terlalu “jauh” dari realitas lapangan.
- Kabupaten butuh AI cepat, bukan AI raksasa.
- Provinsi butuh analisis lintas daerah, bukan data mentah.

Karena itu, Indonesia butuh **arsitektur AI berjenjang**, di mana setiap level pemerintahan menjalankan AI sesuai kapasitas, kebutuhan, dan jenis datanya.

AI Level Desa → Kabupaten → Provinsi → Pusat

AI tidak harus selalu dibuat di Jakarta. AI yang efektif adalah AI yang **mendekat ke data**.

Berikut perannya:

A. AI Level Desa/Kelurahan (AI Mikro – Mendekat ke Warga)

AI di desa berfungsi untuk **melihat masalah harian**:

- mendeteksi balita berisiko stunting,
- memprediksi keluarga rawan miskin,
- membaca keluhan warga dari WhatsApp/Form,
- memudahkan operator desa membuat laporan otomatis.

Contoh open-source untuk desa:

- **Ollama** (LLM lokal),
- **Python + Pandas**,
- **Superset mini dashboard**,
- **n8n** untuk otomasi laporan.

AI desa harus **kecil, cepat, tidak butuh internet kuat**, dan **privasi terjaga**.

B. AI Level Kabupaten/Kota (AI Operasional – Urusan Pelayanan Publik)

AI kabupaten fokus pada **pelayanan publik** dan **manajemen operasional**:

- prediksi dropout siswa,
- prediksi kebutuhan Puskesmas,
- analitik UMKM per kecamatan,
- prediksi beban bansos dan permintaan layanan,
- deteksi anomali data dari desa.

Tools open-source kabupaten:

- **PostgreSQL + TimescaleDB** untuk time-series,
- **Scikit-Learn** untuk prediksi sederhana,
- **Apache Airflow/NiFi** untuk ETL otomatis,
- **Superset** untuk dashboard operasional.

C. AI Level Provinsi (AI Taktis – Koordinasi Lintas Daerah)

Provinsi perlu AI yang melihat **pola besar lintas kabupaten**, misalnya:

- perbandingan indikator antar kab/kota,
- prediksi anggaran sektor kesehatan dan pendidikan,
- mengidentifikasi kabupaten kritis,
- early warning system provinsi (stunting, harga pangan, penyakit).

Provinsi tidak mengelola data mentah, tetapi **mengonsumsi data terstandarisasi** dari kabupaten.

Tools provinsi (open-source):

- **ClickHouse** untuk analitik cepat,
- **Spark** untuk model ML skala besar,
- **Kafka** untuk data streaming kabupaten→provinsi,
- **Grafana/Superset** untuk dashboard pimpinan provinsi.

D. AI Level Pusat (AI Strategis – Untuk Kebijakan Nasional)

AI nasional bekerja pada **data agregat**, bukan data individu. Contoh prediksi nasional:

- kebutuhan fiskal nasional,
- beban bansos nasional,
- peta risiko stunting 1–3 tahun ke depan,
- proyeksi tenaga kerja,
- analitik pendidikan nasional,
- perencanaan UMKM nasional.

AI pusat harus menjawab pertanyaan tingkat kebijakan:

- “**Program apa yang akan berhasil?**”
- “**Anggaran mana yang harus diprioritaskan?**”
- “**Wilayah mana paling kritis tahun depan?**”

Tools nasional (open-source):

- **Kubeflow** (AI pipeline nasional),
- **Spark MLlib** (analitik nasional),
- **Delta Lake/Iceberg** (data lakehouse nasional),
- **HuggingFace models** untuk NLP kebijakan,
- **Wazuh + Suricata** untuk keamanan AI nasional.

AI Operasional, AI Taktis, AI Strategis

Model berjenjang tidak hanya soal level pemerintahan, tetapi juga soal **jenis keputusan**.

A. AI Operasional (Kabupaten/Kota)

Digunakan untuk pekerjaan harian:

- memproses formulir,
- membuat keputusan cepat (misalnya bansos harian),
- klasifikasi surat masuk,
- deteksi anak berisiko stunting,
- memprediksi kapasitas Puskesmas.

AI operasional membantu **ASN bekerja lebih cepat**.

Contoh open-source:

- **Ollama**,
- **n8n**,
- **Metabase**,

- **LightGBM / Scikit-Learn.**

B. AI Taktis (Provinsi)

Digunakan untuk analisis antar kabupaten:

- memprioritaskan kabupaten dengan risiko tertinggi,
- memantau inflasi daerah,
- memetakan beban kesehatan antar wilayah,
- mengidentifikasi kabupaten yang butuh bantuan teknis.

AI taktis membantu pemerintah provinsi **mengkoordinasikan daerah**.

Contoh open-source:

- **Apache Spark,**
- **ClickHouse,**
- **Superset,**
- **Kafka.**

AI Strategis (Nasional)

Digunakan untuk kebijakan jangka panjang:

- prediksi dampak kebijakan nasional,
- proyeksi anggaran 3–5 tahun,
- permodelan skenario (simulasi kebijakan),
- analitik pembangunan nasional.

AI strategis digunakan oleh:

- Bappenas,
- Kemendagri,
- Kemenkeu,
- Kemenkes,
- Kemendikbud,
- BSSN,
- dan kementerian strategis lain.

Open-source nasional:

- **Kubeflow,**
- **Delta Lake / Iceberg,**
- **Spark MLlib,**
- **HuggingFace Transformers.**

Tabel Arsitektur AI Berjenjang

Level	Fungsi AI	Contoh Implementasi	Open Source
Desa / Kelurahan	AI mikro, data harian	deteksi stunting, klasifikasi aduan	Ollama, n8n, Pandas
Kabupaten / Kota	AI operasional	prediksi dropout, beban Puskesmas	Scikit-Learn, Airflow
Provinsi	AI taktis	koordinasi lintas daerah, early warning	Spark, ClickHouse, Kafka
Pusat / Nasional	AI strategis	perencanaan nasional, proyeksi anggaran	Kubeflow, Iceberg, HF models
AI Operasional	tugas harian	klasifikasi dokumen	Ollama, Metabase
AI Taktis	analisis wilayah	prioritas anggaran provinsi	Superset, Spark
AI Strategis	kebijakan nasional	simulasi kebijakan	Kubeflow, Spark MLlib

Pesan Inti

- AI harus dibangun **berjenjang**, bukan dipusatkan di Jakarta saja.
- Desa punya AI sendiri, kabupaten punya AI operasional, provinsi punya AI taktis, pusat punya AI strategis.
- Open-source seperti **Ollama, Spark, ClickHouse, Kubeflow** memungkinkan seluruh jenjang menggunakan AI tanpa biaya mahal.
- Dengan arsitektur ini, Indonesia bisa membangun **GovAI nasional** yang terukur, murah, aman, dan efektif.

Bab 14 — Penyiapan Data Training AI Nasional

AI hanya sebaik kualitas data latihnya.

Jika data buruk, maka hasil prediksi salah.

Jika data bias, maka kebijakan ikut bias.

Jika data tidak lengkap, maka AI “buta”.

Untuk pemerintah, ini sangat kritis karena AI akan digunakan untuk:

- menentukan kelompok rentan,
- memutuskan wilayah prioritas,
- menghitung kebutuhan anggaran,
- membuat peringatan dini,
- mengarahkan bansos dan layanan publik.

Artinya:

Sebelum membuat AI nasional, Indonesia harus membangun fondasi data training yang kuat, bersih, adil, dan representatif.

Membangun *Ground Truth* Nasional

Ground truth adalah **kebenaran dasar** yang dijadikan patokan AI. Pemerintah harus menetapkan data mana yang dianggap *benar* dan standar.

Contoh ground truth nasional:

- NIK dari Dukcapil → identitas valid
- Data sekolah dari Dapodik → pendidikan valid
- Data fasilitas kesehatan dari Kemenkes → kesehatan valid
- Data kemiskinan dari BPS → ekonomi valid
- Peta desa/kelurahan → geospasial valid

Tanpa ground truth, AI akan:

- menghasilkan klasifikasi salah,
- menduplikasi entitas,
- memprioritaskan wilayah yang sebenarnya tidak prioritas,
- membuat kesalahan kebijakan.

Cara membangun ground truth secara bertahap:

1. Tentukan master data utama (identitas, lokasi, sekolah, faskes).
2. Bersihkan data dengan ETL (n8n, NiFi, Airflow).
3. Terapkan *deduplication* (Pandas, Elasticsearch fuzzy matching).
4. Buat standar JSON Schema Nasional.
5. Gunakan *version control* untuk data (DVC, Git-LFS).

Tools open-source pendukung ground truth:

- **PostgreSQL + PostGIS** (master data nasional)
- **ElasticSearch** (deduplication)
- **DVC (Data Version Control)**
- **JSON Schema + OpenAPI**

Labeling oleh ASN & Mahasiswa (Crowdsourcing Nasional)

Labeling adalah proses memberi tanda, kategori, atau penjelasan pada data agar AI bisa “belajar”.

AI tidak bisa mengetahui sendiri apakah:

- seorang anak berisiko dropout,
- desa masuk kategori rawan,
- teks keluhan masyarakat termasuk aduan prioritas,
- laporan Puskesmas termasuk kasus DBD,
- UMKM masuk kategori risiko tinggi.

Itu semua harus dilabeli manusia terlebih dahulu.

Mengapa ASN & mahasiswa penting?

Karena mereka:

- punya konteks lokal,
- mengerti data desanya,
- mengerti kondisi keluarga,
- dekat dengan Puskesmas/sekolah/UMKM,
- mempunyai kapasitas untuk labeling yang akurat.

Model crowdsourcing nasional:

1. **ASN daerah** → labeling data operasional
2. **Mahasiswa (Kampus Merdeka)** → labeling skala besar
3. **Relawan statistik** → validasi acak untuk memastikan kualitas
4. **Pusat (K/L)** → mengaudit dan menyusun quality score

Tools open-source untuk labeling massal:

- **Label Studio**
- **Doccano**
- **TagTog (self-hosted)**

Semua dapat dipasang di server pemda atau provinsi.

Synthetic Data, Data Augmentation, dan Federated Learning

Karena privasi penting, tidak semua data boleh dibagikan ke pusat. Solusinya: **AI modern tidak membutuhkan data mentah** — AI dapat dilatih dengan teknik alternatif yang aman.

A. Synthetic Data

Data buatan yang menyerupai data asli, tetapi **tanpa identitas pribadi**.

Contoh:

- Data kesehatan sintetis
- Data ekonomi sintetis
- Data sekolah sintetis

Tools open-source:

- **SDV (Synthetic Data Vault)**
- **YData Synthetic**
- **CTGAN, TVAE**

Synthetic data menjaga privasi sambil tetap memberikan pola bagi AI.

B. Data Augmentation

Menambah variasi data agar model AI lebih kuat.

Contoh augmentation:

- teks aduan → parafrase otomatis (Ollama)
- gambar peta → diputar/diperbesar
- data time-series → ditambah variasi noise ringan

Augmentation membuat AI lebih tahan terhadap data yang beragam dari daerah.

C. Federated Learning

Data tetap berada di daerah, hanya **model yang bergerak**, bukan datanya.

Langkahnya:

1. Pusat mengirim model awal ke provinsi/kabupaten.
2. Model belajar dari data lokal tanpa memindahkan datanya.
3. Model kembali ke pusat dalam bentuk bobot yang sudah diperbarui.
4. Model nasional jadi semakin pintar, tanpa melanggar privasi.

Open-source untuk federated learning:

- **Flower**
- **TensorFlow Federated (TFF)**
- **OpenFL**

Dengan federated learning, AI nasional dapat dilatih tanpa pernah melihat data sensitif warga.

Audit Bias & Fairness

Ini bagian paling kritis. AI yang tidak diaudit akan berpotensi:

- mendiskriminasi kelompok tertentu,
- salah menandai rumah tangga miskin,
- memperbesar stigma sosial,
- memprioritaskan daerah tertentu secara tidak adil.

Karena itu, pemerintah harus **mengukur bias** pada setiap model AI.

Jenis bias yang harus diaudit:

- **Bias wilayah** → provinsi tertentu terlalu sering diprioritaskan.
- **Bias ekonomi** → keluarga tani/nelayan sering salah diklasifikasi.
- **Bias gender** → bias terhadap perempuan atau lansia.
- **Bias layanan** → satu sekolah/UMKM lebih sering gagal lewat sistem.

Cara audit bias:

1. Pisahkan data berdasarkan kelompok.
2. Uji apakah AI memperlakukan kelompok berbeda tanpa alasan valid.
3. Gunakan fairness metrics.
4. Lakukan perbaikan pada dataset.
5. Rilis laporan audit AI secara berkala.

Tools open-source untuk audit fairness:

- **AI Fairness 360 (IBM)**
- **Fairlearn**
- **What-If Tool (TensorFlow)**

Audit ini wajib agar AI pemerintah **adil, transparan, dan tidak merugikan warga**.

Tabel Penyiapan Data Training AI Nasional

Komponen	Fungsi	Dampak	Open-Source
Ground Truth Nasional	Standar data patokan AI	AI lebih akurat	PostgreSQL, DVC, JSON Schema
Crowdsourcing Labeling	Pelabelan besar-besaran	Data lebih kaya & valid	Label Studio, Doccano
Synthetic Data	Lindungi privasi	Data aman dibagi	SDV, CTGAN
Augmentation	Variasi data	AI lebih tahan & kuat	Ollama, Pandas
Federated Learning	Latih AI tanpa kirim data	Privasi terjaga	Flower, TFF
Audit Bias & Fairness	Memastikan keadilan	AI tidak merugikan warga	AIF360, Fairlearn

Pesan Inti

- AI nasional hanya akan berhasil jika data training-nya kuat.
- Ground truth adalah fondasi.
- Labeling massal oleh ASN & mahasiswa adalah solusi skala nasional.
- Synthetic data & federated learning menjaga privasi dan keamanan.
- Audit fairness wajib agar AI tidak bias dan tidak merugikan rakyat.
- Seluruh proses ini dapat dibangun dengan **open-source murah, aman, dan dapat dikontrol pemerintah**.

Bab 15 — Analitik Berjenjang untuk Pengambilan Keputusan

Indonesia terlalu besar untuk memakai satu jenis dashboard saja. Setiap level pemerintahan memiliki kebutuhan analitik yang **berbeda**, sehingga dashboard harus “berjenjang”:

- Desa & kabupaten → kebutuhan **operasional harian**
- Provinsi → kebutuhan **taktis lintas-wilayah**
- Pusat → kebutuhan **strategis nasional**

Jika dashboard disamakan, maka:

- data membingungkan,
- keputusan lambat,
- intervensi tidak tepat sasaran,
- pimpinan kewalahan informasi.

Karena itu, analitik harus **dibangun sesuai level kewenangan**.

Dashboard Operasional (Kabupaten/Kota)

Fokus: pelayanan publik harian.

Dashboard kabupaten/kota harus sederhana, langsung aksi, dan menampilkan kondisi lapangan.

Jenis analitik yang dibutuhkan kabupaten/kota:

- jumlah bayi stunting hari ini,
- beban Puskesmas mingguan,
- siswa berisiko dropout,
- UMKM rentan tutup,
- progres bansos desa,
- harga pangan harian,
- laporan keluhan masyarakat.

Karakter dashboard operasional:

- real-time / harian,
- visual sederhana,
- fokus per kecamatan/desa,
- bisa dipakai camat, kadis, dan operator.

Contoh open-source:

- **Apache Superset** → visualisasi cepat
- **Metabase** → tanya-jawab data sederhana
- **Grafana** → monitoring real-time
- **PostgreSQL** → data warehouse lokal
- **n8n** → pengiriman data otomatis dari desa

Output utama:

Pengambilan keputusan cepat, misalnya:

- Puskesmas A overload → kirim tenaga tambahan.
- Kecamatan B inflasi → intervensi harga.
- Sekolah C dropout tinggi → kirim pendamping.

Dashboard Taktis (Provinsi)

Provinsi tidak mengurus detail teknis. Provinsi membutuhkan **gambaran antar-kabupaten** untuk koordinasi dan alokasi anggaran.

Jenis analitik provinsi:

- peta risiko stunting antar kabupaten,
- perbandingan inflasi antar wilayah,
- monitoring belanja daerah kab/kota,
- beban layanan kesehatan per kabupaten,
- capaian pendidikan tingkat provinsi,
- early warning system lintas wilayah.

Karakter dashboard taktis:

- agregat kabupaten/kota,
- analisis tren bulanan,
- fokus prioritas wilayah,
- komparasi antar daerah.

Contoh open-source:

- **ClickHouse** → analitik cepat untuk seluruh kab/kota
- **Apache Spark** → analitik skala besar
- **Superset** → peta risiko dan tren
- **Kafka** → streaming data kabupaten → provinsi

Output utama:

Koordinasi lintas daerah, misalnya:

- kabupaten X masuk zona merah stunting → provinsi kirim tim,
- kabupaten Y undervaccinated → alokasi vaksin dipercepat,
- kabupaten Z inflasi tinggi → operasi pasar provinsi.

Dashboard Strategis (Pusat/Nasional)

Dashboard pusat digunakan untuk **kebijakan tingkat nasional**, bukan untuk memonitor desa atau kecamatan.

Jenis analitik di pusat:

- proyeksi anggaran nasional,
- prediksi kebutuhan bansos nasional,
- analitik kemiskinan ekstrem nasional,
- monitoring capaian RPJMN,
- peta risiko nasional (ekonomi, kesehatan, pendidikan),
- tren stunting nasional 3–5 tahun ke depan.

Pusat tidak butuh detail micro-data; pusat butuh **interpretasi strategis**.

Karakter dashboard strategis:

- fokus kebijakan, bukan operasional
- agregasi nasional → provinsi
- tren jangka panjang
- indikator keberhasilan pembangunan

Contoh open-source:

- **Kubeflow** → pipeline AI nasional
- **Iceberg/Delta Lake** → lakehouse nasional
- **Superset (enterprise-scale)** → dashboard menteri
- **Spark MLlib** → prediksi nasional

Output utama:

Kebijakan nasional berbasis bukti:

- revisi anggaran,
- penentuan program prioritas,
- analitik untuk Presiden, Menteri PPN, Mendagri, Kemenkeu.

Mengapa Open Source Sangat Penting?

Karena pemerintah butuh:

- transparansi,
- keamanan,
- kedaulatan data,
- biaya rendah,
- interoperabilitas,
- tidak terkunci vendor.

Dengan open source:

- kabupaten bisa pakai Superset gratis,
- provinsi bisa pakai ClickHouse tanpa lisensi,
- pusat bisa pakai Kubeflow tanpa biaya vendor,
- seluruh pipeline end-to-end bisa dikuasai ASN.

Tabel Analitik Berjenjang Indonesia

Level Dashboard	Fokus	Contoh Analitik	Open-Source
Kabupaten/Kota	Operasional	stunting harian, dropout, UMKM, harga pasar	Superset, Metabase, Grafana, PostgreSQL
Provinsi	Taktis	peta risiko kabupaten, inflasi wilayah, koordinasi	ClickHouse, Spark, Kafka, Superset
Pusat/Nasional	Strategis	proyeksi anggaran, capaian RPJMN, analitik nasional	Kubeflow, Iceberg, Spark MLlib

Pesan Inti

- Analitik harus **berjenjang**, bukan satu dashboard untuk semua.
- Kabupaten fokus operasional → provinsi fokus taktis → pusat fokus strategis.
- Open-source seperti **Superset, Metabase, ClickHouse, Spark, Kubeflow** memungkinkan seluruh jenjang bekerja efisien tanpa biaya lisensi.
- Dengan analitik berjenjang, setiap level pemerintahan memiliki informasi tepat sesuai kewenangannya.
- Inilah fondasi **keputusan cepat, akurat, dan berbasis bukti** untuk Republik Indonesia.

BAGIAN VI — KEAMANAN, PRIVASI, & PERLINDUNGAN DATA PRIBADI

Bab 16 — Keamanan Infrastruktur Data Pemerintah

Mengapa Keamanan Data Pemerintah Sangat Kritis?

Pemerintah menyimpan **data paling sensitif** di Republik ini:

- NIK, KK
- data kesehatan
- data bansos
- data guru dan siswa
- data keuangan publik
- data aset, infrastruktur, dan operasi pemerintahan

Serangan ke sistem pemerintah selalu meningkat. Banyak daerah tidak siap. Masalah tipikal di lapangan:

- password sama selama bertahun-tahun,
- server tidak dipantau,
- akses terlalu luas,
- tidak ada enkripsi,
- sistem lama tidak ditambah,
- data bisa diakses siapa saja,
- tidak ada monitoring insiden nyata.

Karena itu, Bab 16 menekankan bahwa keamanan bukan tugas BSSN saja. **Setiap kabupaten/kota, provinsi, kementerian harus menerapkan arsitektur keamanan modern: Zero Trust.**

Zero Trust Architecture (ZTA)

Zero Trust berarti:

“Jangan percaya siapapun, cek semuanya, verifikasi setiap akses.”

Tidak ada lagi asumsi bahwa:

- perangkat internal aman,
- jaringan kantor aman,
- aplikasi warisan aman,
- user terdaftar otomatis aman.

Prinsip Zero Trust yang wajib untuk pemerintah:

1. **Verifikasi identitas setiap akses**
— bahkan untuk ASN di dalam kantor.
2. **Akses minimal (least privilege)**
— user hanya bisa mengakses data yang benar-benar dibutuhkan.
3. **Micro-segmentation**
— pecah jaringan menjadi zona kecil agar serangan tidak meluas.
4. **Selalu audit & monitor**
— semua aktivitas dicatat.
5. **Tidak ada akses tanpa enkripsi**
— semua komunikasi harus lewat saluran aman.

Contoh Open Source untuk Zero Trust:

- **WireGuard** → VPN terenkripsi antar instansi
- **Keycloak** → autentikasi dan Single Sign-On nasional
- **pfSense/OPNsense** → firewall, segmentasi jaringan
- **Wazuh** → pemantauan ancaman
- **Suricata** → deteksi serangan jaringan

ZTA memastikan bahwa **sekali pun server bocor, data tetap aman** karena setiap komponen saling memverifikasi.

Enkripsi — Fondasi Mutlak Keamanan Nasional

Tanpa enkripsi, data pemerintah **bisa diambil siapa saja** yang berhasil masuk jaringan.

Enkripsi yang wajib digunakan pemerintah:

- **Enkripsi saat transit** (HTTPS, TLS, WireGuard)
- **Enkripsi saat disimpan** (PostgreSQL TDE, LUKS disk encryption)
- **Enkripsi untuk backup**
- **Enkripsi API antar instansi**

Open source untuk enkripsi:

- **OpenSSL**
- **WireGuard VPN**
- **Age/Minisign** untuk backup terenkripsi
- **pgcrypto** untuk kolom sensitif di PostgreSQL

Dengan enkripsi penuh, sekalipun pelaku mencuri database, data tetap tidak bisa dibaca.

IAM (Identity Access Management) — Kendalikan Siapa Boleh Akses Apa

IAM memastikan bahwa **akses pengguna sesuai perannya**, tidak lebih.

Masalah hari ini:

- operator desa bisa melihat data seluruh kabupaten,
- kontraktor IT memiliki akses lebih besar daripada ASN,
- akun lama tidak dinonaktifkan,
- pergantian pegawai tidak disertai pengaturan ulang akses.

IAM yang benar harus:

1. mengatur identitas seluruh ASN,
2. memberi akses berdasarkan jabatan, bukan orang,
3. mencabut akses otomatis ketika pegawai mutasi,
4. punya audit log lengkap.

Open-source IAM:

- **Keycloak** → SSO, OAuth2, OpenID Connect
- **FreeIPA** → manajemen user & perangkat

IAM = **pagar pertama** untuk melindungi data pemerintah.

RBAC (Role-Based Access Control)

RBAC memastikan bahwa:

- bidang pendidikan melihat data pendidikan saja,
- bidang kesehatan hanya melihat data kesehatan,
- kecamatan hanya melihat data kecamatannya,
- operator hanya bisa CRUD sesuai kewenangan.

RBAC sangat mudah diterapkan jika pemerintah sudah menggunakan:

- JSON Schema Nasional
- API Gateway Nasional
- IAM (Keycloak)

RBAC membantu mencegah **akses berlebihan** yang sering menjadi sumber kebocoran data.

Logging — Semua Aktivitas harus Terekam

Sebagian besar kebocoran data tidak terdeteksi karena **tidak ada log** atau log langsung hilang.

Logging membantu:

- mengetahui siapa mengakses data kapan,
- menelusuri kejadian sebelum insiden,
- mengaudit kelalaian operator,
- mengidentifikasi pola serangan.

Open-source logging:

- **Wazuh** → log endpoint
- **Elastic Stack (ELK)** → analisis log skala besar
- **Fluentd / Logstash** → pengumpul log

Tanpa logging, keamanan **hanya asumsi**.

SIEM — Deteksi Serangan secara Real-Time

SIEM (Security Information and Event Management) adalah “Pusat Intelijen Keamanan”.

SIEM mengumpulkan log dari:

- server
- aplikasi
- API
- jaringan
- firewall

Lalu mendeteksi pola serangan secara otomatis.

Contoh serangan yang bisa dideteksi SIEM:

- brute force login
- akses data tidak wajar
- pencurian token API
- percobaan injeksi SQL
- ransomware awal

Open-source SIEM:

- **Wazuh** (SIEM + XDR)
- **OpenSearch Security** → analisis log
- **TheHive** → response insiden

Dengan SIEM yang benar, serangan dapat terdeteksi dalam menit, bukan berbulan-bulan.

Tabel Keamanan Infrastruktur Data Pemerintah

Komponen	Fungsi	Dampak	Open Source
Zero Trust Architecture	Cek semua akses, tidak percaya otomatis	Mengurangi serangan internal & lateral movement	Keycloak, WireGuard, pfSense
Enkripsi	Melindungi data dalam perjalanan & penyimpanan	Data tetap aman meski server bocor	OpenSSL, pgcrypto, LUKS
IAM	Kelola identitas & akses pengguna	Akses lebih aman & terkontrol	Keycloak, FreeIPA
RBAC	Batasi akses sesuai jabatan	Mencegah akses berlebihan	Keycloak, API Gateway
Logging	Rekam semua aktivitas	Audit insiden & pelacakan	Wazuh, ELK
SIEM	Deteksi serangan real-time	Respon cepat ancaman	Wazuh, OpenSearch

Pesan Inti

- Keamanan data pemerintah adalah fondasi negara digital.
- Zero Trust harus menjadi standar nasional.
- Enkripsi, IAM, RBAC, logging, dan SIEM wajib diterapkan di semua level — desa, kabupaten, provinsi, pusat.
- Semua dapat dibangun dengan **open-source**, sehingga murah, aman, dan transparan.
- Keamanan bukan proyek; keamanan adalah **proses harian** yang harus dijalankan oleh seluruh instansi.

Bab 17 — Cyber Security yang Efektif dan Efisien untuk Pemerintah

Realitas Hari Ini: Pemerintah Diserang Setiap Hari, tetapi Banyak Pemda Tidak Siap

Kebocoran data dan ransomware yang menimpa pemerintah biasanya tidak terjadi karena “serangan canggih”, tetapi karena:

- password lemah,
- layanan tidak pernah diperbarui,
- server tidak dipantau,
- database tidak dienkripsi,
- tidak ada firewall,
- tidak ada log,
- tidak ada tim respon insiden.

Akibatnya akun bocor, server diambil alih, dan data masyarakat tersebar di internet.

Karena itu, bagian ini menekankan konsep:

Cyber security harus sederhana, efisien, open-source, dan bisa dijalankan oleh ASN daerah yang SDM-nya terbatas.

Hardening Sistem — Langkah Murah yang Paling Mengurangi Risiko

Hardening berarti memperkuat sistem sebelum diserang. Ini adalah langkah tercepat, termurah, dan paling efektif.

Prinsip Hardening untuk Pemerintah:

1. **Matikan layanan yang tidak perlu**
 - Semakin sedikit layanan aktif, semakin kecil permukaan serangan.
2. **Wajib update semua sistem**
 - Banyak serangan sukses karena patch tidak dipasang.
3. **Gunakan firewall default-deny**
 - Semua port ditutup kecuali yang digunakan.
4. **Wajib MFA (multi-factor authentication)**
 - Mengurangi risiko pembobolan akun.
5. **Enkripsi database & backup**
 - Jika database dicuri, data tetap aman.

6. **Pisahkan server internal dan publik**
 - Jangan satukan database dengan API publik.

Contoh Open-Source untuk Hardening:

- **Lynis** — audit keamanan server Linux
- **OpenSCAP** — compliance baseline keamanan
- **ufw/pfSense/OPNsense** — firewall
- **Fail2ban** — blokir brute force

Hardening yang benar dapat mengurangi **80% risiko serangan** tanpa biaya besar.

Security-in-a-Box untuk Pemda Kecil

Sebagian besar kabupaten/kota tidak punya tim keamanan.

Tidak punya staf khusus.

Tidak punya anggaran besar.

Solusinya adalah **security-in-a-box**, yaitu paket keamanan siap pakai yang murah dan mudah dipasang.

Komponen Security-in-a-Box (Open Source):

1. **Firewall + VPN**
 - pfSense / OPNsense + WireGuard
2. **SIEM kecil**
 - Wazuh (server + agen)
 - Bisa memantau seluruh server OPD
3. **IDS/IPS**
 - Suricata
4. **Monitoring layanan**
 - Grafana + Prometheus
5. **Backup terenkripsi otomatis**
 - restic atau borgbackup
6. **Autentikasi tunggal**
 - Keycloak untuk SSO dan RBAC

Keuntungan untuk pemda kecil:

- tidak perlu teknisi ahli,
- instalasi cepat (2–3 jam),
- monitoring terpusat,
- bisa berjalan di server kelas menengah,

- murah dan bisa mandiri.

Dengan **security-in-a-box**, pemda kecil bisa punya keamanan setara kementerian.

GovCSIRT Nasional — Pusat Respon Insiden untuk Seluruh Pemerintah

CSIRT adalah **Computer Security Incident Response Team**.

Banyak daerah tidak punya CSIRT sendiri. Maka negara harus menyediakan **GovCSIRT Nasional** sebagai pusat koordinasi keamanan.

Fungsi GovCSIRT Nasional:

- menerima laporan insiden dari pemda/kementerian,
- menyediakan panduan langkah cepat,
- membuat database ancaman nasional,
- mengirim alert ke seluruh pemda,
- membantu analisis forensik sederhana,
- memberikan patch dan rekomendasi keamanan,
- melatih tim lokal pemda.

GovCSIRT yang efektif bekerja seperti:

- hotline nasional cybersecurity,
- pusat command and control insiden,
- pusat analitik ancaman nasional,
- repository keamanan nasional.

Open-source pendukung GovCSIRT:

- **TheHive** — platform manajemen insiden
- **Cortex** — analisis artefak otomatis
- **MISP** — berbagi intelijen ancaman
- **Wazuh** — SIEM nasional

GovCSIRT harus mudah diakses oleh seluruh pemda, tanpa birokrasi rumit.

Red Team – Blue Team & Bug Bounty untuk Pemerintah

Keamanan tidak boleh pasif. Harus ada latihan dan pengujian berkala.

Red Team

Tim yang **mensimulasikan serangan**:

- uji brute force,
- penetrasi API,

- eksploitasi server lama,
- simulasi ransomware.

Blue Team

Tim yang **bertahan dan merespons**:

- memantau log,
- mendeteksi serangan,
- melakukan isolasi server,
- mengaktifkan backup,
- memblokir IP penyerang.

Pelatihan red-blue harus berjalan minimal 2x setahun.

Bug Bounty untuk Pemerintah

Cara paling efisien untuk menemukan celah keamanan:
ajak komunitas hacker etis menemukan bug dan laporkan.

Manfaat:

- biaya rendah,
- menemukan celah lebih cepat daripada audit vendor,
- membangun budaya keamanan yang terbuka.

Open-source untuk red-blue team:

- **Kali Linux** (red team)
- **Metasploit** (penetration testing)
- **OpenVAS** (vulnerability scanning)
- **Wazuh** (blue team monitoring)
- **URBackup/restic** (blue team recovery)

Dengan bug bounty & red-blue team, pemerintah bisa **aktif** menjaga keamanan, bukan hanya menunggu serangan.

Tabel Cyber Security Efektif & Efisien

Komponen	Fungsi	Dampak Cepat	Open Source
Hardening System	Mengurangi permukaan serangan	Risiko turun drastis	Lynis, pfSense, Fail2ban
Security-in-a-Box	Keamanan siap pakai untuk pemda	Proteksi cepat & murah	Wazuh, Suricata, WireGuard
GovCSIRT Nasional	Koordinasi insiden seluruh pemerintah	Respon cepat & terpusat	TheHive, MISP, Cortex
Red Team–Blue Team	Uji kekuatan & kesiapan	Deteksi celah sebelum diserang	Kali, Metasploit, Wazuh
Bug Bounty	Deteksi celah oleh komunitas	Lebih cepat & efisien	Platform self-hosted + audit OS

Pesan Inti

- Keamanan data pemerintah harus efisien, sederhana, dan kuat.
- Hardening + security-in-a-box adalah solusi cepat untuk pemda kecil.
- GovCSIRT Nasional wajib menjadi pusat respon insiden.
- Red-blue team & bug bounty memastikan keamanan terus meningkat.
- Semua dapat dibangun dengan **open-source** tanpa biaya lisensi mahal.

Bab 18 — Perlindungan Data Pribadi sesuai UU PDP

UU PDP: Bukan Dokumen Hukum, Tetapi “Perisai Negara Digital”

UU PDP adalah fondasi agar pemerintah mengelola data pribadi warga secara aman, etis, dan bertanggung jawab. Masalah saat ini:

- Banyak instansi menyimpan data pribadi tanpa pengamanan.
- Server pemda masih terbuka ke internet tanpa firewall.
- Data seperti NIK, KK, nomor telepon beredar bebas.
- Kebocoran data sering tidak terdeteksi dan tidak dilaporkan.

UU PDP memberikan aturan jelas:

data pribadi harus dijaga, diproses dengan izin, digunakan secara minimal, dan dilindungi dengan teknis yang memadai.

Untuk pemerintah, implementasinya harus praktis dan murah → **gunakan open-source yang sudah matang.**

Prinsip Utama UU PDP yang Wajib Dipahami ASN

UU PDP menekankan 8 prinsip utama. Yang paling penting untuk operasional pemerintah:

A. Lawfulness & Purpose Limitation — Data digunakan untuk tujuan jelas

Setiap pengumpulan data harus punya tujuan spesifik.
Dilarang mengumpulkan “sebanyak mungkin” data tanpa alasan.

B. Data Minimization — Hanya ambil yang diperlukan

Contoh: Untuk survei stunting, tidak perlu mengambil nomor rekening.

C. Accuracy — Data harus akurat dan terbaru

Jika salah, pemerintah wajib memperbaiki.

D. Storage Limitation — Tidak boleh menyimpan data selamanya

Harus ada masa retensi.

E. Integrity & Confidentiality — Data harus dilindungi

Gunakan enkripsi, firewall, IAM, RBAC.

F. Accountability — Pemerintah bertanggung jawab atas pelanggaran

Ada sanksi administratif dan pidana jika lalai.

Intinya:

Pemerintah wajib memastikan seluruh sistem mematuhi prinsip ini secara otomatis, bukan manual.

Masking, Hashing, Tokenisasi — Teknik Perlindungan Data yang Wajib Dipakai

UU PDP mewajibkan **pengamanan data berbasis risiko**.

Tiga teknik ini wajib dipakai di semua sistem pemerintah.

A. Masking — Menyembunyikan sebagian data

Dipakai ketika data perlu ditampilkan tetapi tidak boleh utuh.

Contoh masking:

- NIK → 3175*****12
- Nomor telepon → 08*****331
- Alamat → ditampilkan hanya level desa

Open-source tools:

- **pgcrypto** di PostgreSQL
- **Metabase/Superset masking policies**
- **Python Pandas** masking function

B. Hashing — Mengubah data menjadi kode permanen

Data asli tidak bisa dikembalikan.

Cocok untuk:

- password ASN
- ID unik warga
- identitas yang tidak perlu konsumsi publik

Open-source tools:

- **bcrypt, scrypt, Argon2**
- **OpenSSL** hashing

C. Tokenisasi — Mengganti data asli dengan token sementara

Token bisa dipakai sistem lain tanpa membuka data asli.

Contoh:

API kabupaten membutuhkan nomor identitas, tetapi pusat memberikan “token”, bukan NIK asli.

Open-source tools:

- **Vault (HashiCorp)** versi open-source
- **OpenTokenization libraries**

Tokenisasi adalah teknik paling aman untuk pertukaran data antar instansi.

Incident Response & Breach Handling — Harus Ada SOP Nasional

Sebagian besar kebocoran data di Indonesia **tidak terdeteksi**, tidak dilaporkan, atau ditangani terlalu lambat.

UU PDP wajibkan pemerintah:

- mendeteksi insiden cepat,
- menginformasikan pihak terdampak,
- memperbaiki sistem dalam waktu tertentu.

Langkah incident response untuk pemerintah:

1. **Deteksi cepat**
gunakan Wazuh, Suricata, OpenSearch.
2. **Isolasi sistem terdampak**
matikan akses publik sementara.
3. **Analisis sumber insiden**
cek log, file access, API token.
4. **Perbaikan teknis**
patch, rollback, atau hardening.
5. **Laporan insiden**
ke GovCSIRT Nasional (Bab 17).
6. **Pemberitahuan publik**
sesuai ketentuan UU PDP.
7. **Pencegahan ulang**
audit konfigurasi dan akses.

Open-source pendukung incident response:

- **Wazuh** → SIEM + XDR
- **TheHive** → manajemen insiden
- **Cortex** → analisis artefak
- **ELK/OpenSearch** → analitik log

Dengan pipeline ini, pemda/kementerian bisa meningkatkan kecepatan respons secara drastis.

Klasifikasi Data Sensitif & Rahasia Negara

Pemerintah harus mengklasifikasi data berdasarkan tingkat sensitivitasnya. Ini adalah syarat dalam UU PDP **dan** pedoman keamanan nasional.

Kategori Data:

A. Data Pribadi Umum

- nama lengkap
- jenis kelamin
- alamat umum

Akses hanya untuk keperluan layanan.

B. Data Pribadi Sensitif (kategori UU PDP)

- kesehatan
- biometrik
- catatan pendidikan lengkap
- keuangan pribadi
- data populasi lengkap
- lokasi presisi
- agama, kepercayaan
- data anak

Harus dilindungi dengan:

- enkripsi kuat
- RBAC ketat
- audit lengkap
- tidak boleh disimpan tanpa retensi

C. Data Rahasia Negara

- konstelasi keamanan
- intelijen siber
- data infrastruktur strategis
- topologi PDN
- konfigurasi firewall pusat
- vulnerability reports
- data keuangan strategis pemerintah

Akses sangat terbatas.

Tidak boleh disimpan di cloud komersial.

Open-source untuk manajemen klasifikasi:

- **Keycloak RBAC**
- **PostgreSQL row-level security (RLS)**
- **Vault** untuk penyimpanan rahasia
- **MinIO encrypted storage**

Tabel Perlindungan Data Pribadi

Komponen	Fungsi	Dampak	Open Source
Prinsip UU PDP	landasan hukum & etika	proses data lebih aman & terkontrol	—
Masking	sembunyikan sebagian data	kurangi risiko eksposur	pgcrypto, Pandas
Hashing	amankan data permanen	cegah pembacaan data sensitif	bcrypt, Argon2
Tokenisasi	tukar data tanpa buka data asli	aman untuk API antar instansi	Vault, opentoken
Incident Response	tangani kebocoran cepat	kurangi dampak insiden	Wazuh, TheHive
Klasifikasi Data	batasi akses berdasarkan risiko	keamanan lebih ketat & terukur	Keycloak, RLS

Pesan Inti

- UU PDP mewajibkan pemerintah melindungi data warga dengan standar tinggi.
- Masking, hashing, tokenisasi, enkripsi, IAM, dan logging harus menjadi default.
- Insiden harus ditangani cepat dengan prosedur modern dan open-source seperti Wazuh & TheHive.
- Data sensitif dan data rahasia negara harus diklasifikasikan jelas dan diberi perlindungan berlapis.
- Perlindungan data bukan proyek IT—ini adalah **mandat negara dan hak warga Republik Indonesia**.

BAGIAN VII — SDM, KAPASITAS, & KURIKULUM NASIONAL

Bab 19 — Kompetensi SDM Pemerintah di Era Data

SDM adalah Masalah Utama. Teknologi Bukan Penghambat.

Saat ini banyak pemda dan kementerian mengalami masalah yang sama:

- data ada, tapi tidak dibaca,
- dashboard ada, tapi tidak digunakan,
- rekomendasi AI muncul, tapi tidak diikuti,
- API tersedia, tapi tidak ada yang paham,
- keamanan data penting, tapi tidak ada yang menjaga.

Karena itu, Bab 19 menekan satu hal penting:

Transformasi digital pemerintah hanya berhasil jika ASN paham data, bukan hanya memakai aplikasi.

Kompetensi SDM jauh lebih penting daripada tool.

Data Literacy ASN — Pondasi Pemerintah Modern

Data literacy berarti **kemampuan ASN untuk membaca, memahami, dan menggunakan data dalam pekerjaan hariannya.**

Ini bukan kemampuan teknis tinggi.

Bukan coding.

Bukan statistik rumit.

Data literacy = ASN tahu cara menggunakan data untuk membuat keputusan.

Kompetensi dasar Data Literacy ASN:

- paham arti indikator (angka kemiskinan, stunting, inflasi daerah)
- bisa membaca grafik dengan benar
- bisa membedakan data valid dan tidak valid
- paham pentingnya kualitas data
- tahu bahaya data ganda / data kotor
- paham kenapa retensi & privasi penting
- mampu menyimpulkan tren sederhana

Tools open-source untuk pembelajaran ASN:

- **Superset** → belajar baca dashboard
- **Metabase** → belajar tanya data tanpa coding
- **Google Colab / Jupyter** → pembelajaran ringan data
- **n8n** → mengenalkan konsep workflow data otomatis

Tujuan utama: ASN tidak hanya menjadi operator, tetapi menjadi **pemikir berbasis data**.

Membaca Dashboard, Insight, dan Rekomendasi Kebijakan

Tidak semua pemimpin butuh coding. Tetapi semua pemimpin butuh **membaca insight**.

Kompetensi yang diperlukan:

1. **Interpretasi Dashboard**
 - memahami arti warna merah, kuning, hijau
 - membaca tren 3–12 bulan
 - memahami peta risiko
 - menghubungkan indikator antar sektor
2. **Membaca Insight**
 - apa yang artinya “anomali data”?
 - apa makna kenaikan mendadak?
 - apakah ini masalah data atau masalah lapangan?
3. **Mengambil Tindakan dari Data**
 - jika stunting naik → intervensi gizi
 - jika dropout tinggi → pendamping sekolah
 - jika harga pangan naik → operasi pasar
4. **Membaca Rekomendasi Kebijakan dari AI**

AI memberikan rekomendasi, tetapi ASN harus:

 - mengecek akurasi,
 - mengecek sumber datanya,
 - menilai apakah rekomendasi masuk akal,
 - mengaitkannya dengan kebijakan daerah.

Tools open-source pendukung:

- **Apache Superset** → dashboard pimpinan
- **Metabase** → query sederhana
- **Grafana** → indikator real-time
- **ClickHouse** → analitik cepat

Tujuan: pimpinan bisa berkata,

“Saya membuat keputusan berdasarkan data, bukan opini.”

Pemahaman API & Keamanan Data — Wajib untuk Semua Wali Data

Tidak semua ASN perlu menulis API, tetapi **semua Wali Data dan pejabat struktural harus memahami konsep dasarnya.**

API adalah “jembatan pengantar data” antar instansi.

Pemahaman API yang wajib untuk pejabat data:

- data tidak boleh dikirim manual via WA/Excel
- API wajib menggunakan format standar (JSON Nasional)
- data yang lewat API otomatis tervalidasi
- API harus punya batas akses (rate limit / RBAC)
- API harus lewat jalur aman (HTTPS, VPN, WireGuard)

Apa risiko jika tidak paham API?

- data bocor
- data dipakai instansi lain tanpa izin
- data tidak sinkron
- kesalahan fatal dalam laporan nasional
- gagal memenuhi UU PDP

Pemahaman keamanan data yang harus dimiliki:

- apa itu data sensitif?
- apa yang boleh dibagi dan tidak boleh dibagi?
- apa itu masking, hashing, enkripsi?
- bagaimana menangani insiden kebocoran?
- bagaimana audit log digunakan?

Tools open-source untuk memahami API dan keamanan:

- **Kong API Gateway** → melihat alur API
- **Keycloak** → SSO dan kontrol akses
- **WireGuard** → jalur aman antar instansi
- **Wazuh** → melihat log keamanan
- **Postman (OSS)** → mencoba API

ASN tidak harus ahli teknis.

Tetapi ASN harus **tahu konsep**, agar bisa melakukan pengawasan.

Tabel Kompetensi SDM Pemerintah

Kompetensi	Fungsi	Dampak	Open Source Pendukung
Data Literacy ASN	Memahami data & tren	ASN mampu analisis mandiri	Superset, Metabase
Membaca Dashboard & Insight	Interpretasi kondisi lapangan	Keputusan cepat & tepat	Superset, Grafana
Rekomendasi Kebijakan dari AI	Menerjemahkan insight ke aksi	Kebijakan berbasis bukti	Ollama + Superset
Pemahaman API	Jaga standar pertukaran data	Data lintas instansi konsisten	Kong, Postman, JSON Schema
Keamanan Data	Lindungi data pribadi & rahasia negara	Patuh UU PDP & aman dari kebocoran	Wazuh, WireGuard, Keycloak

Pesan Inti

- SDM adalah fondasi utama Data Governance Indonesia.
- ASN tidak wajib bisa coding, tetapi wajib bisa membaca data.
- Dashboard, insight, dan rekomendasi AI harus menjadi tool harian birokrat.
- Wali Data dan pejabat struktural wajib memahami API dan keamanan data.
- Semua kompetensi dapat dilatih dengan **open-source murah dan mudah digunakan**.

Bab 20 — Kurikulum Kampus untuk Mendukung Data Governance RI

Indonesia Butuh SDM yang Bisa Bangun Sistem Data, Bukan Hanya Pengguna Aplikasi

Masalah terbesar Data Governance Indonesia bukan hardware, bukan software, tetapi **kekurangan SDM yang paham data**:

- kampus tidak mengajarkan standar data nasional,
- ASN tidak diajarkan membaca dashboard,
- mahasiswa tidak diajarkan ETL, API, JSON Schema,
- lulusan IT tidak bisa mengelola data pemerintah,
- keamanan siber belum diajarkan dengan konteks pemerintahan.

Karena itu, kurikulum kampus harus mengikuti kebutuhan **Data Governance RI**, bukan hanya kebutuhan industri umum.

Kurikulum harus **praktis, open-source, murah, dan langsung bisa diterapkan di pmda/kementerian**.

Program Data Engineering & AI

(materi lengkap tersedia gratis di: opencourse.itts.ac.id)

Data Engineering adalah tulang punggung Data Governance. Tanpa Data Engineer, tidak ada ETL, tidak ada data warehouse, tidak ada pipeline AI.

Kompetensi inti Data Engineering untuk Indonesia:

- memahami JSON Nasional / JSON Schema
- memahami ETL dari desa → kabupaten → provinsi → pusat
- membangun data warehouse dengan PostgreSQL
- membangun pipeline data dengan n8n, Airflow, atau NiFi
- memproses data besar dengan Spark
- mengintegrasikan API antar instansi
- mengelola data lake & lakehouse
- menyiapkan data training AI

Open source yang wajib ada dalam kurikulum:

- **PostgreSQL** (DW nasional)
- **n8n, Apache NiFi, Airflow** (ETL/ELT)
- **Kafka** (real-time streaming)
- **ClickHouse** (analitik cepat provinsi/pusat)
- **Ollama, HuggingFace, Spark MLlib** (AI)

Kurikulum ini harus membuat mahasiswa **langsung siap kerja di pemda/kementerian**.

Keamanan Siber & Privasi (UU PDP)

(materi tersedia gratis di: **opencourse.itts.ac.id**)

Kampus harus mengajarkan keamanan yang relevan dengan konteks pemerintah. Bukan hanya teori, tetapi:

Kompetensi keamanan yang dibutuhkan pemerintah:

- Zero Trust Architecture
- enkripsi data
- pengelolaan API aman
- IAM & RBAC
- forensik sederhana
- SIEM berbasis open-source
- deteksi intrusi
- penanganan insiden kebocoran (UU PDP)
- masking, hashing, tokenisasi

Open-source yang wajib masuk kurikulum:

- **Wazuh** (SIEM nasional)
- **Suricata** (IDS/IPS)
- **pfSense/OPNsense** (firewall)
- **WireGuard** (VPN antar instansi)
- **Keycloak** (IAM nasional)
- **OpenSSL, pgcrypto**

Mahasiswa harus memahami bahwa keamanan data pemerintah bukan opsional — ini kewajiban negara.

Open Source for Government

(materi tersedia di komunitas **Sahabat Ubuntu Indonesia**)

Pemerintah butuh SDM yang bisa bekerja **tanpa ketergantungan vendor**, karena:

- anggaran terbatas,
- daerah banyak yang kecil,
- aplikasi komersial mahal,
- kebutuhan interoperabilitas tinggi,
- kedaulatan data penting.

Kompetensi open-source untuk pemerintah:

- instalasi & konfigurasi server Ubuntu LTS
- manajemen PostgreSQL
- deployment aplikasi open-source pemerintah
- membuat dashboard dengan Superset/Metabase
- membuat workflow dengan n8n
- memahami container (Docker/Podman)
- manajemen monitoring dengan Grafana/Prometheus
- keamanan open-source

Open-source wajib dalam kurikulum pemerintah:

- **Ubuntu LTS**
- **PostgreSQL**
- **Superset, Metabase**
- **n8n, NiFi**
- **Grafana, Prometheus**
- **Wazuh, Suricata**

Kampus harus melatih mahasiswa untuk bisa **membangun layanan pemerintah dari nol**, bukan hanya belajar membuat aplikasi startup.

Micro-credential untuk ASN

Tidak semua ASN perlu kuliah panjang. Tetapi ASN harus punya **micro-credential singkat**, 6–12 jam, langsung bisa dipakai.

Micro-credential prioritas untuk ASN:

A. Data Literacy (Wajib)

- membaca dashboard
- memahami tren
- mengenali data tidak valid
- dasar-dasar UU PDP

B. Dasar ETL

- konsep data pipeline
- memvalidasi JSON
- upload data standar

C. Dasar API & Pertukaran Data

- cara kerja API
- apa yang boleh/tidak boleh dibagi
- keamanan API dasar

D. Keamanan Data Dasar

- password policy
- privasi data
- mengenali phishing
- penanganan insiden awal

Open-source yang cocok untuk micro-credential:

- **Metabase** (baca dashboard)
- **Superset** (analitik sederhana)
- **Postman** (cek API)
- **Keycloak** (akses & SSO)
- **Wazuh** (lihat insiden dasar)

Micro-credential membuat ASN **naik kelas** tanpa mengganggu pekerjaan harian.

Tabel Kurikulum untuk Data Governance RI

Komponen	Tujuan	Dampak	Open Source
Data Engineering & AI	bangun pipeline data nasional	data bersih & siap AI	PostgreSQL, n8n, NiFi, Spark
Keamanan Siber & Privasi	patuhi UU PDP & Zero Trust	keamanan meningkat	Wazuh, Suricata, Keycloak
Open Source for Government	kemandirian teknologi daerah	murah & efisien	Ubuntu, Superset, Grafana
Micro-credential ASN	tingkatkan kemampuan cepat	keputusan lebih berbasis data	Metabase, Postman

Pesan Inti

- Data Governance hanya berhasil jika kampus menghasilkan SDM yang *mengerti data*, bukan sekadar pengguna aplikasi.
- Kurikulum harus fokus pada Data Engineering, AI, keamanan, dan open-source.
- ASN butuh micro-credential cepat agar bisa mengikuti transformasi digital.
- Semua kompetensi bisa diajarkan dengan **open-source murah dan mudah** yang sudah tersedia di:
 - **opencourse.itts.ac.id**
 - **Sahabat Ubuntu Indonesia**

Bab 21 — Model Pelatihan Nasional Bertingkat

Kenapa Indonesia Butuh Pelatihan Bertingkat?

Karena 514 kabupaten/kota, 34 provinsi, dan puluhan kementerian **tidak mungkin dilatih dengan model yang sama**.

Masalah nyata di lapangan:

- kabupaten kekurangan SDM yang paham data,
- provinsi butuh koordinator data lintas kabupaten,
- kementerian perlu analisis strategis,
- tidak semua daerah punya anggaran besar,
- guru, operator, dan ASN lapangan belum paham standar data nasional.

Solusinya: **model pelatihan nasional bertingkat**, sederhana, murah, dan bisa disebarluaskan cepat.

Pelatihan Level Kabupaten, Provinsi, Kementerian

Pelatihan harus **dipecah sesuai fungsi**, bukan dibuat seragam.

A. Pelatihan Level Kabupaten/Kota (Level Operasional)

Fokus: kemampuan teknis dasar untuk menjalankan Data Governance harian.

Kompetensi yang diajarkan:

- input data standar (JSON Nasional)
- validasi, verifikasi, deduplication
- penggunaan dashboard Superset/Metabase
- ETL sederhana dengan **n8n**
- pengamanan dasar (password, VPN, IAM)
- penanganan insiden awal

Open-source yang diajarkan:

- n8n
- Superset, Metabase
- PostgreSQL dasar
- Wazuh (alert dasar)
- WireGuard

Kabupaten tidak perlu pemrogram tinggi — cukup **operator yang paham data**.

B. Pelatihan Level Provinsi (Level Taktis)

Fokus: koordinasi lintas kabupaten.

Kompetensi provinsi:

- analitik lintas kabupaten
- early warning system provinsi
- supervisi kualitas data
- federated data management
- API lintas kabupaten–provinsi
- Spark/ClickHouse dasar

Open-source yang diajarkan:

- ClickHouse
- Apache Spark / Kafka
- Superset tingkat lanjut
- Kong API Gateway

Provinsi bertindak sebagai “**koordinator data daerah**”.

C. Pelatihan Level Kementerian (Level Strategis)

Fokus: kebijakan nasional berbasis data.

Kompetensi kementerian:

- analitik nasional
- forecasting 1–5 tahun
- kebijakan berbasis data
- menyusun standar JSON nasional
- arsitektur AI nasional
- audit fairness & bias
- federated learning

Open-source yang diajarkan:

- Kubeflow
- Spark MLlib
- Delta Lake / Iceberg
- HuggingFace
- TheHive, Wazuh untuk keamanan nasional

Kementerian harus mampu menjadi **penentu arah** Data Governance.

Train-the-Trainer Nasional (TtT)

Training seluruh ASN secara langsung **tidak mungkin**.

Jika 5% saja dilatih langsung, transformasi tidak akan terjadi.

Maka solusi terbaik adalah **Train-the-Trainer Nasional (TtT)**.

Prinsip Train-the-Trainer:

1. **Kementerian/Provinsi melatih pelatih inti (Trainer Inti Nasional)**
2. Pelatih inti melatih pelatih di provinsi (Provincial Trainer)
3. Trainer provinsi melatih kabupaten/kota
4. Kabupaten melatih operator desa/UPTD

Model ini **menggandakan kapasitas pelatihan** secara eksponensial.

Kompetensi yang diberikan kepada Trainer Nasional:

- memahami arsitektur data nasional
- bisa mengajar Data Engineering dasar
- bisa mengajar penggunaan dashboard
- bisa mengajar keamanan dasar (Zero Trust)
- bisa mengajar ETL dan API
- mampu mentor proyek data di daerah

Open-source untuk TtT:

- platform pembelajaran: Moodle / Canvas
- repository materi: GitHub / GitLab
- demo sistem: Superset, n8n, Wazuh
- modul praktik: notebook Jupyter

Hasil akhirnya: Indonesia memiliki **ribuan trainer lokal** yang mampu membangun ekosistem data daerah.

Government Data Community — Komunitas Praktisi Pemerintah

Kunci sukses Data Governance di dunia adalah **komunitas**. Tanpa komunitas, transformasi digital akan hilang ketika pimpinan berganti.

Indonesia membutuhkan **Government Data Community** nasional.

Tujuan komunitas:

- berbagi best-practice antar daerah
- berbagi kesalahan dan solusi
- berbagi schema JSON dan API
- berbagi ETL workflow n8n

- saling membantu ketika sistem bermasalah
- replikasi cepat inovasi kabupaten ke kabupaten lain
- wadah kolaborasi non-birokratis

Bentuk aktivitas:

- forum bulanan (online)
- hackathon data pemerintah
- workshop open-source
- rilis template JSON Nasional
- pertemuan fisik tahunan: "Government Data Summit"

Platform komunitas (open-source & gratis):

- **Discourse** → forum
- **Mattermost / Rocket.Chat** → komunikasi harian
- **GitLab/GitHub** → berbagi kode & schema
- **Nextcloud** → berbagi file kebijakan
- **Superset gallery** → berbagi dashboard

Government Data Community mempercepat penyebaran kemampuan **tanpa anggaran besar**.

Tabel Model Pelatihan Nasional Bertingkat

Level	Fokus	Output Utama	Open Source Pendukung
Kabupaten/Kota	Operasional	operator data kuat	n8n, Superset, WireGuard
Provinsi	Taktis	koordinator data daerah	ClickHouse, Kafka
Kementerian	Strategis	kebijakan nasional	Kubeflow, Iceberg
Train-the-Trainer	Perbanyak kapasitas	ribuan trainer lokal	Moodle, GitHub, Jupyter
Government Data Community	Berbagi & kolaborasi	inovasi menyebar cepat	Discourse, Rocket.Chat

Pesan Inti

- Indonesia tidak butuh pelatihan mahal—yang dibutuhkan adalah **pelatihan bertingkat** sesuai fungsi.
- Kabupaten belajar operasional, provinsi belajar taktis, kementerian belajar strategis.
- Train-the-Trainer Nasional membuat kapasitas meluas cepat.
- Komunitas praktisi pemerintah mempercepat inovasi tanpa birokrasi.
- Semua pelatihan dapat dilakukan dengan **open-source murah, aman, dan mudah dipelajari**.

BAGIAN VIII — IMPLEMENTASI, ROADMAP, & STUDI KASUS

Bab 22 — Roadmap 10 Tahun Data Governance Nasional

Roadmap 10 Tahun: Perlu agar Indonesia Tidak Berjalan Tanpa Arah

Selama ini banyak inisiatif data pemerintah berjalan sendiri-sendiri.

Ada dashboard, tetapi tidak ada data.

Ada aplikasi, tetapi tidak interoperable.

Ada API, tetapi tidak standar.

Ada data, tetapi tidak bisa dipakai AI.

Karena itu, dibutuhkan **roadmap 10 tahun yang konsisten** dan dapat diikuti oleh seluruh kementerian, provinsi, dan kabupaten.

Roadmap ini sederhana, murah, dan dibangun dengan **open-source** agar semua daerah bisa ikut tanpa hambatan dana.

TAHUN 1–2 — Standardisasi & Interoperability (Fondasi)

Fase ini adalah kunci. Jika gagal, transformasi tidak jalan.

Fokusnya bukan teknologi canggih.

Fokusnya **rapikan fondasi data nasional**.

Prioritas nasional:

- JSON Schema Nasional untuk seluruh sektor
- Kamus data nasional (metadata terstandar)
- Data Catalog nasional & daerah
- API Gateway nasional
- Format pertukaran data antar lembaga
- Validasi & deduplikasi standar kabupaten–provinsi–pusat
- Pelatihan Data Literacy ASN (micro credential)

Deliverable nyata:

- Dashboard nasional memakai data standar
- Pipeline ETL nasional untuk kesehatan, pendidikan, bansos
- Penurunan error data lintas instansi
- Daerah mulai mengirim data dalam format seragam

Open-source yang dipakai:

- **Kong / API7** → API Gateway Nasional
- **PostgreSQL** → standar DW nasional
- **Superset / Metabase** → dashboard nasional
- **n8n / Apache NiFi** → ETL antar instansi
- **CKAN** → Data Catalog Nasional
- **JSON Schema Validator**

Tujuan fase: Semua data pemerintah bisa “berbicara bahasa yang sama”.

TAHUN 3–4 — Konsolidasi Provinsi & GovCloud

Setelah standar terbentuk, tahap berikutnya adalah **sentralisasi ringan** dengan struktur bertingkat.

Prioritas provinsi:

- pusat konsolidasi data kabupaten/kota
- validasi & verifikasi lintas kabupaten
- data pipeline real-time ke pusat
- penguatan keamanan provinsi
- cluster server provinsi untuk sektor strategis (pendidikan, kesehatan, bansos)

Prioritas nasional:

- implementasi **GovCloud**: hybrid + multi cloud
- Pusat Data Nasional terstandar
- segmentasi data sensitif & rahasia negara
- integrasi dashboard nasional lintas sektor
- SIEM nasional untuk keamanan data pemerintah

Open-source yang dipakai:

- **OpenStack / Harvester** → GovCloud open-source
- **MinIO** → object storage nasional
- **ClickHouse** → analitik provinsi & nasional
- **Grafana + Loki** → monitoring nasional
- **Wazuh** → SIEM nasional
- **WireGuard** → koneksi aman kabupaten–provinsi–pusat

Tujuan fase: Data dari 514 kab/kota dapat dikonsolidasikan otomatis lewat provinsi dengan biaya rendah.

TAHUN 5–7 — AI Nasional (AI untuk Semua Sektor)

Pada fase ini, data sudah stabil. Saatnya membangun AI untuk membantu pengambilan keputusan.

Prioritas AI Nasional:

- model AI untuk prediksi stunting
- model AI prediksi kemiskinan & kesejahteraan
- model pendidikan: prediksi dropout, kualitas sekolah
- prediksi UMKM: omzet, risiko bangkrut
- model risiko bencana daerah
- integrasi AI dengan dashboard nasional
- AutoML untuk pemerintah daerah
- federated learning: data tetap aman, model tetap belajar

Open-source untuk AI Nasional:

- **Ollama** → AI lokal di server pemda
- **HuggingFace** → model nasional dapat dikembangkan sendiri
- **Spark MLlib** → AI skala besar
- **Kubeflow** → manajemen AI pipeline nasional
- **Label Studio** → labeling & dataset nasional
- **Delta Lake / Apache Iceberg** → lakehouse nasional

Tujuan fase:

AI menjadi **tool harian** bagi pemerintah.

Tidak perlu vendor mahal.

Tidak perlu AI di cloud luar negeri.

AI bisa berjalan di server pemerintah sendiri.

TAHUN 8–10 — Predictive Government (Pemerintah Prediktif)

Pada fase akhir, Indonesia bertransformasi dari:

“Menunggu masalah terjadi” → “Memprediksi sebelum masalah muncul.”

Ciri pemerintah prediktif:

- stunting terdeteksi sebelum terjadi
- anggaran daerah diarahkan otomatis ke sektor yang butuh
- peringatan dini pendidikan, kesehatan, dan ekonomi
- bansos presisi → tidak lagi salah sasaran
- data terkonsolidasi real-time dari seluruh daerah
- AI memberikan rekomendasi kebijakan, bukan operator manual
- pemimpin daerah melihat risiko 6–12 bulan ke depan
- pusat bisa memonitor kinerja daerah secara objektif

Open-source pendukung fase ini:

- **Prophet, ARIMA, Neural Forecast** → forecasting
- **Grafana + Alerting** → early warning system
- **Elastic / OpenSearch** → analitik geospasial
- **Apache Pinot** → analitik real-time nasional
- **Ollama + RAG Nasional** → AI Q&A kebijakan publik

Tujuan fase:

Indonesia memiliki **Pemerintah Prediktif**, yang mampu menurunkan biaya, meningkatkan layanan, dan meningkatkan kesejahteraan rakyat berbasis data.

TABEL Roadmap 10 Tahun Data Governance Nasional

Tahun	Fokus Utama	Hasil Nyata	Open Source Pendukung
1–2	Standardisasi & interoperabilitas	JSON Schema, API Nasional, DW awal	PostgreSQL, n8n, Kong, CKAN
3–4	Konsolidasi Provinsi & GovCloud	Data lintas kabupaten, cloud nasional	OpenStack, MinIO, Wazuh
5–7	AI Nasional	Model AI sektor strategis	Ollama, HuggingFace, Spark, Kubeflow
8–10	Predictive Government	Early warning & prediksi nasional	Grafana, OpenSearch, Pinot

Pesan Inti

- Roadmap 10 tahun sangat penting agar Data Governance Indonesia tidak lagi sporadis.
- Tahun 1–2 adalah fase paling kritis: **standarisasi data nasional**.
- Tahun 3–4: provinsi menjadi pusat konsolidasi dengan GovCloud open-source.
- Tahun 5–7: AI Nasional mulai berjalan di seluruh sektor.
- Tahun 8–10: Indonesia menjadi **Pemerintah Prediktif** yang bekerja berdasarkan prediksi, bukan reaktif.
- Semuanya bisa dibangun **tanpa biaya lisensi besar** — cukup open-source.

Bab 23 — Indikator Keberhasilan & Audit Sistem Data Nasional

Bagian ini menjawab satu pertanyaan mendasar:

“Bagaimana pemerintah tahu bahwa Data Governance berjalan, membawa dampak, dan menghemat anggaran negara?”

Audit tidak boleh hanya formalitas.

Audit harus menghasilkan **angka**, **skor**, dan **rekomendasi nyata**.

Tiga indikator utama yang ditetapkan dokumen adalah:

1. **Indeks Maturitas Data Daerah**
2. **Indeks Ketahanan Siber Pemerintah**
3. **Penghematan anggaran dari Data Governance**

Ini adalah tiga indikator yang dapat diukur oleh seluruh kabupaten, provinsi, dan kementerian.

Indeks Maturitas Data Daerah

Indeks ini mengukur **seberapa siap daerah dalam mengelola data**, dari level paling dasar (manual) hingga terintegrasi penuh (AI readiness).

Komponen penilaiannya:

- **Standarisasi data:** apakah daerah sudah memakai JSON Nasional?
- **Kualitas data:** duplikasi, missing data, inkonsistensi.
- **Interoperabilitas:** apakah data kabupaten mengalir otomatis ke provinsi?
- **Data catalog:** apakah aset data terdokumentasi?
- **ETL daerah:** apakah n8n/airflow atau workflow lain sudah dipakai?
- **Dashboard operasional:** apakah data digunakan pimpinan?
- **Keamanan dasar:** IAM, enkripsi, password policy.

Contoh open-source pendukung penilaian:

- **Superset** → audit dashboard
- **Metabase** → cek kualitas data lewat query
- **n8n** → cek pipeline data
- **PostgreSQL** → audit struktur data
- **CKAN** → data catalog daerah
- **Wazuh** → audit konfigurasi keamanan dasar

Indeks maturitas ini menjadi dasar **penghargaan, pendampingan, dan alokasi anggaran data**.

Indeks Ketahanan Siber Pemerintah

Ketahanan siber harus diukur secara rutin, bukan setelah serangan terjadi.

Komponen penilaiannya:

1. **Hardening system**
 - apakah port terbuka tidak perlu ditutup?
2. **Sistem deteksi ancaman**
 - apakah IDS/IPS aktif?
3. **SIEM**
 - apakah log dikumpulkan dan dianalisis?
4. **Response time**
 - berapa menit daerah mendeteksi insiden?
5. **Backup & recovery**
 - apakah backup terenkripsi dan rutin diuji?
6. **VPN & Zero Trust**
 - apakah semua akses internal lewat jalur aman?
7. **Incident reporting**
 - apakah daerah melapor ke GovCSIRT Nasional?

Open-source untuk mengukur ketahanan siber:

- **Wazuh** → SIEM, XDR, audit konfigurasi
- **Suricata** → deteksi intrusi
- **pfSense/OPNsense** → audit firewall
- **WireGuard** → audit VPN
- **OpenVAS** → vulnerability assessment
- **TheHive** → pencatatan insiden

Indeks ini mengukur apakah daerah **tahan diserang**, bukan sekadar punya antivirus.

Penghematan Anggaran dari Data Governance

Data Governance yang benar menghasilkan **penghematan anggaran nyata**, terutama karena:

A. Efisiensi Operasional

- ETL otomatis → mengurangi pekerjaan manual.
- Interoperabilitas → tidak perlu entri data ulang.
- Dashboard → pimpinan cepat ambil keputusan.

B. Pengurangan Kecurangan & Ketidaktepatan

- bansos tidak salah sasaran → hemat miliaran per tahun,
- stunting lebih akurat → intervensi tepat,
- data pendidikan bersih → bantuan tepat sekolah.

C. Penghematan Lisensi

Open-source mengurangi biaya besar:

- PostgreSQL menggantikan database mahal,
- Superset menggantikan BI komersial,
- WireGuard menggantikan VPN lisensi,
- Wazuh menggantikan SIEM ratusan juta per tahun.

Open-source yang menghasilkan penghematan langsung:

- **PostgreSQL, MariaDB, ClickHouse** → DW nasional
- **Superset, Metabase** → dashboard gratis
- **n8n, NiFi** → ETL tanpa lisensi
- **Grafana** → monitoring gratis
- **Wazuh** → SIEM tanpa biaya lisensi
- **MinIO** → object storage murah untuk big data

Penghematan ini **bisa dihitung**, bukan sekadar asumsi. Hasil audit akan menunjukkan daerah mana yang paling efisien.

Tabel Indikator & Audit Data Nasional

Indikator	Makna	Cara Audit	Open Source Pendukung
Indeks Maturitas Data Daerah	kesiapan daerah mengelola data	cek standar, ETL, dashboard	Superset, n8n, CKAN
Indeks Ketahanan Siber	kemampuan bertahan dari serangan	audit firewall, SIEM, IDS	Wazuh, Suricata, pfSense
Penghematan Anggaran	efisiensi dari data governance	hitung efisiensi & lisensi	PostgreSQL, Grafana, MinIO

Pesan Inti

- Keberhasilan Data Governance **harus diukur**, bukan diduga.
- Tiga indikator utama nasional: maturitas data daerah, ketahanan siber, dan penghematan anggaran.
- Semua indikator bisa diaudit murah dan cepat dengan **open-source**.
- Audit bukan formalitas—audit menentukan **arah anggaran dan prioritas nasional**.

Bab 24 — Studi Kasus Pemerintah Republik Indonesia

Bagian ini akan memberikan gambaran konkret:

“Jika Data Governance diterapkan dengan benar, apa dampaknya bagi program pemerintah?”

Setiap sektor memiliki masalah klasik yang sama:

- data tidak sinkron,
- data ganda,
- laporan lambat,
- tidak ada standar,
- tidak ada API,
- dashboard tidak dipakai,
- dan akhirnya **kebijakan meleset dari target**.

Dengan data yang rapi, interoperable, dan dianalisis dengan open-source, lima sektor ini dapat berubah total.

Bansos — Dari Salah Sasaran ke Presisi Nasional

Masalah hari ini:

- data ganda antar kementerian,
- data lama tidak diperbarui,
- desa → kecamatan → kabupaten sering inkonsisten,
- penerima tidak memenuhi kriteria,
- anggaran bocor dan tidak efektif.

Dengan Data Governance:

- Data penduduk terstandar melalui JSON Nasional
- Deduplication otomatis (PostgreSQL + Elasticsearch fuzzy matching)
- Integrasi kabupaten → provinsi → pusat via API
- Dashboard presisi penerima bantuan
- Deteksi otomatis rumah tangga rentan menggunakan AI lokal (Ollama/Scikit-Learn)

Open-source pendukung:

- **n8n** → ETL bansos daerah → pusat
- **PostgreSQL** → cleansing dan dedup
- **Superset** → dashboard verifikasi
- **Kafka** → streaming update bansos real-time

Dampaknya:

Bansos tepat sasaran, pemborosan hilang, dan penghematan miliaran per kabupaten per tahun.

Pendidikan — Deteksi Dini Dropout & Kualitas Sekolah

Masalah hari ini:

- Dapodik tidak sinkron dengan data daerah,
- sulit deteksi siswa berisiko dropout,
- tidak ada peringatan dini kondisi sekolah,
- laporan lambat dan manual.

Dengan Data Governance:

- Data sekolah terstandarisasi
- Absensi, nilai, dan status ekonomi digabungkan
- AI memprediksi siswa berisiko dropout
- Dashboard risiko per sekolah dan kecamatan
- Provinsi dapat mengirim pendamping lebih cepat

Open-source pendukung:

- **ClickHouse** → analitik cepat antar sekolah
- **Superset** → peta risiko dropout
- **Ollama + Python** → prediksi dropout lokal
- **Airflow** → pipeline harian Dapodik → provinsi

Dampaknya:

Dropout dapat dicegah **sebelum terjadi**, bukan setelah laporan akhir tahun.

Kesehatan — Stunting, Puskesmas, dan Early Warning System

Masalah hari ini:

- data balita, Puskesmas, dan posyandu terpisah,
- tidak ada integrasi kabupaten-provinsi-pusat,
- intervensi datang terlambat,
- data gizi tidak akurat.

Dengan Data Governance:

- Data balita terstandar JSON Nasional
- Validasi kesehatan otomatis dengan n8n
- Prediksi stunting oleh model AI nasional
- Dashboard kesehatan lintas kabupaten
- Early Warning System provinsi

Open-source pendukung:

- **PostgreSQL + TimescaleDB** → data kesehatan harian
- **Grafana** → monitoring Puskesmas

- **Spark MLlib** → prediksi stunting skala provinsi
- **Superset** → dashboard kesehatan

Dampaknya:

Intervensi gizi bisa diberikan **bulan sebelumnya**, bukan setelah stunting muncul.

UMKM — Prediksi Risiko dan Dukungan Tepat Sasaran

Masalah hari ini:

- data UMKM tersebar di berbagai instansi,
- tidak ada dashboard omzet, risiko, atau permasalahan,
- tidak ada prediksi UMKM yang akan tutup,
- kebijakan bantuan UMKM tidak tepat sasaran.

Dengan Data Governance:

- integrasi data perizinan, NIB, pajak, dan laporan daerah
- segmentasi UMKM berdasarkan omzet dan risiko
- prediksi UMKM yang membutuhkan modal
- dashboard perekonomian per kecamatan
- prioritas intervensi otomatis

Open-source pendukung:

- **ClickHouse** → analitik omzet dan transaksi
- **Metabase** → insight cepat untuk dinas UMKM
- **Kafka + n8n** → aliran data perizinan
- **Scikit-Learn** → prediksi risiko UMKM

Dampaknya:

Bantuan UMKM menjadi presisi, bukan merata ke semua tanpa data.

Fiskal Daerah — Perencanaan Anggaran Berbasis Data

Masalah hari ini:

- anggaran disusun berdasarkan pola lama, bukan prediksi,
- tidak ada integrasi data ekonomi daerah,
- keputusan fiskal tidak berbasis risiko atau tren,
- evaluasi sering terlambat.

Dengan Data Governance:

- penggabungan data belanja, PAD, populasi, dan inflasi
- forecasting 1–3 tahun (kesehatan, pendidikan, bansos)
- dashboard fiskal daerah

- skenario kebijakan otomatis (simulasi APBD)
- rekomendasi alokasi berbasis data

Open-source pendukung:

- **Prophet / ARIMA** → forecasting fiskal
- **Superset** → dashboard anggaran
- **PostgreSQL** → data warehouse fiskal
- **Grafana** → alert fiskal mingguan

Dampaknya:

APBD menjadi lebih efisien, tepat sasaran, dan berbasis bukti.

Tabel Studi Kasus Pemerintah RI

Sektor	Masalah Utama	Solusi Data Governance	Open Source
Bansos	salah sasaran, data ganda	ETL standar, dedup, dashboard	n8n, Superset, PostgreSQL
Pendidikan	sulit prediksi dropout	model AI, analitik lintas sekolah	ClickHouse, Superset, Ollama
Kesehatan	stunting terlambat diketahui	predictive model, EWS	Spark, Grafana, TimescaleDB
UMKM	intervensi tidak tepat	segmentasi risiko, prediksi omzet	Metabase, Scikit-Learn
Fiskal Daerah	anggaran tidak berbasis data	forecasting, simulasi kebijakan	Prophet, PostgreSQL, Superset

Pesan Inti

- Data Governance bukan teori. Dampaknya **langsung terasa** pada bansos, pendidikan, kesehatan, UMKM, dan fiskal daerah.
- Semua studi kasus dapat dijalankan dengan open-source murah dan stabil.
- Dengan data yang rapi dan AI lokal, pemerintah bisa **lebih cepat, lebih presisi, dan lebih hemat anggaran**.

Penutup & Rekomendasi Nasional

Bab terakhir ini menjawab satu pertanyaan yang paling penting:

“Bagaimana Indonesia bisa memastikan Data Governance berjalan 10 tahun ke depan tanpa bergantung pada figur, tanpa menunggu proyek besar, dan tanpa biaya lisensi yang mahal?”

Intinya sederhana:

Bangun fondasi, latih SDM, pakai open-source, dan buat mekanisme audit yang konsisten. Jika empat hal ini dilakukan, Data Governance Indonesia akan otomatis bergerak, meskipun terjadi pergantian pemimpin atau reformasi birokrasi.

Kesimpulan Nasional: Data = Infrastruktur Negara

Dokumen ini menunjukkan hal yang sangat jelas:

- **Data sudah setara dengan jalan, listrik, air, dan jaringan internet.**
- Tanpa data yang rapi, semua program pemerintah berjalan dengan tebakan.
- Tanpa interoperabilitas, anggaran terbuang.
- Tanpa standardisasi, AI tidak mungkin berjalan.
- Tanpa keamanan, data bocor dan negara rugi besar.

Indonesia membutuhkan **sistem data nasional** yang kuat, bukan aplikasi yang banyak. Data Governance bukan proyek. Data Governance adalah **cara kerja negara modern**.

Empat Prinsip Utama yang Harus Dipertahankan Indonesia

1. Standarisasi nasional adalah wajib, bukan opsional.

Semua data pemerintah harus mengikuti:

- JSON Nasional
 - metadata dan kamus data nasional
 - mekanisme API nasional
- Ini adalah dasar interoperabilitas.

2. Infrastruktur data harus open-source dan bisa dikelola daerah.

Tanpa open-source:

- biaya tinggi,
- ketergantungan vendor,
- daerah kecil tidak mampu ikut.

3. SDM pemerintah adalah fondasi yang harus dibangun terus-menerus.

Tanpa Data Engineer, Analis Data, Wali Data, dan DPO yang terlatih, sistem tidak akan jalan.

4. Audit berkala adalah kunci keberlanjutan.

Apa yang tidak diaudit akan berhenti.
Audit menjaga disiplin data seluruh daerah.

Rekomendasi Nasional

A. Untuk Presiden & Kementerian Koordinator

- Tetapkan Data Governance sebagai **agenda prioritas nasional**.
- Pastikan semua kementerian memakai standar JSON Nasional.
- Dorong penggunaan **open-source nasional** untuk mengurangi ketergantungan.

B. Untuk Kementerian Teknis (Kemenkes, Kemendikbud, Kemensos, UMKM, dll.)

- Buat **API standar** untuk kabupaten–provinsi–pusat.
- Wajibkan dashboard Superset/Metabase untuk pimpinan.
- Terapkan ETL otomatis dengan **n8n** di setiap program prioritas.

C. Untuk Provinsi

- Jadi **pusat konsolidasi data** kabupaten/kota.
- Bangun GovCloud provinsi berbasis OpenStack/Harvester.
- Terapkan SIEM provinsi dengan **Wazuh**.

D. Untuk Kabupaten/Kota

- Gunakan open-source (PostgreSQL, n8n, Superset).
- Bersihkan data secara berkala (validasi, dedup, audit).
- Latih operator desa & UPTD (micro credential).

E. Untuk Perguruan Tinggi

- Terapkan kurikulum Data Engineering, AI, dan Keamanan Siber.
- Berkolaborasi dengan pemda untuk proyek data nyata.
- Sediakan lab open-source (Ubuntu, PostgreSQL, Grafana).

F. Untuk BSSN & GovCSIRT

- Bangun indeks ketahanan siber nasional.
- Wajibkan Wazuh/Suricata sebagai baseline keamanan instansi.

- Latih red team & blue team daerah.

Peran Open-Source dalam Kedaulatan Digital Indonesia

Open-source adalah fondasi yang memungkinkan:

- **biaya rendah,**
- **pelatihan cepat,**
- **daerah kecil ikut transformasi,**
- **kedaulatan data nasional,**
- **AI lokal berjalan di server pemerintah,**
- **interoperabilitas antar instansi tanpa vendor lock-in.**

Contoh open-source yang menjadi tulang punggung:

- PostgreSQL (DW nasional)
- n8n / NiFi (ETL)
- Superset & Metabase (dashboard)
- Keycloak (IAM)
- Wazuh & Suricata (keamanan)
- Ollama (AI lokal)
- MinIO (object storage)
- ClickHouse (analitik provinsi/pusat)

Dengan open-source, Indonesia bisa menghemat **triliunan rupiah per tahun**, terutama pada lisensi database, dashboard, firewall, SIEM, dan AI.

Pesan Akhir: Data Governance sebagai Fondasi Indonesia Emas 2045

Jika Indonesia ingin mencapai **Indonesia Emas 2045**, tiga hal harus menjadi prioritas nasional:

1. **Data yang rapi**
2. **SDM yang paham data**
3. **Keamanan data yang kuat**

Tanpa tiga hal ini:

- bansos tetap salah sasaran,
- stunting sulit diturunkan,
- pendidikan tidak presisi,
- UMKM sulit naik kelas,
- kebijakan fiskal tidak akurat,
- AI Nasional tidak punya “bensin”.

Tetapi jika Data Governance diterapkan:

- anggaran lebih efisien,

- kebijakan lebih tepat,
- layanan publik meningkat,
- AI menjadi tool bantu negara,
- Indonesia lebih kompetitif,
- dan rakyat merasakan langsung manfaatnya.

Data Governance bukan dokumen.

Data Governance adalah **cara kerja baru Republik Indonesia**.

Tabel Rekomendasi Nasional

Area	Rekomendasi	Teknologi Open Source
Standarisasi	JSON Nasional, Kamus Data	JSON Schema Validator
Infrastruktur	GovCloud + Data Lake	OpenStack, MinIO
ETL Nasional	otomatisasi lintas instansi	n8n, Apache NiFi
Dashboard	analitik untuk pimpinan	Superset, Metabase
Keamanan	SIEM & IDS nasional	Wazuh, Suricata
AI Nasional	AI lokal sektor publik	Ollama, Spark MLlib
SDM	micro credential ASN	Moodle, Jupyter
Audit	maturitas & siber	Grafana, OpenVAS

Pesan Singkat

- Data Governance harus menjadi **fondasi pembangunan nasional**.
- Gunakan open-source untuk efisiensi dan kedaulatan.
- Bangun SDM, bukan proyek aplikasi.
- Audit data harus rutin.
- Dengan data yang rapi, Indonesia bisa membuat **kebijakan lebih presisi dan lebih adil**.

LAMPIRAN — JSON Schema Nasional

Lampiran ini adalah bagian paling penting dari keseluruhan buku. Karena disinilah “bahasa bersama” antar kementerian, provinsi, dan kabupaten ditetapkan secara teknis dan operasional.

Selama ini, instansi:

- memakai format data sendiri-sendiri,
- struktur berbeda,
- nama kolom tidak sinkron,
- format tanggal berbeda,
- nomor identitas tidak konsisten,
- unit daerah membuat format lokal yang tidak bisa dibaca pusat.

Akibatnya:

- data **tidak dapat digabung**,
- analitik nasional tidak valid,
- AI Nasional tidak bisa dilatih,
- interoperabilitas gagal,
- kebijakan tidak presisi.

JSON Schema Nasional adalah solusi langsung dan permanen.

Apa Itu JSON Schema Nasional

Ini adalah **format baku** untuk jenis data pemerintah yang paling kritis. Format ini menetapkan:

- nama field (kolom)
- tipe data (string, number, boolean)
- format data (NIK, tanggal, koordinat)
- aturan validasi wajib
- struktur objek dan relasi dasar
- atribut minimal dan opsional

Dengan schema ini, semua data dari desa, kabupaten, provinsi, kementerian **otomatis kompatibel**.

Tidak perlu lagi data mapping manual.

Tidak perlu lagi rekonsiliasi.

Tidak perlu lagi pembersihan besar-besaran.

Mengapa JSON Schema Nasional Sangat Penting?

Karena **standarisasi** adalah fondasi Data Governance.

Tanpa schema yang seragam:

- n8n dan Airflow tidak bisa memproses data
- PostgreSQL tidak bisa melakukan validasi
- Superset tidak bisa menampilkan dashboard nasional
- Kafka tidak bisa men-stream data antar instansi
- AI tidak bisa dilatih karena format kacau

Dengan schema nasional, seluruh pipeline pemerintah menjadi lebih stabil, akurat, dan efisien.

Struktur Umum JSON Schema Nasional

Schema nasional memiliki pola umum:

1. **Identitas dasar**
 - NIK / NIP / ID sekolah / ID UMKM
2. **Lokasi**
 - kode desa, kode kecamatan, kode kabupaten, kode provinsi
3. **Waktu**
 - tanggal pelaporan, timestamp
4. **Atribut inti sektor**
 - pendidikan, kesehatan, bansos, UMKM, fiskal
5. **Metadata pelaporan**
 - sumber data, versi schema, tingkat validasi, petugas entri

Format yang digunakan tegas, konsisten, dan mudah dipahami operator di daerah.

Contoh JSON Schema Nasional (Kategori Besar)

Berikut kategori besar JSON Schema yang ada di lampiran:

A. Bansos

1. Rumah Tangga Miskin
2. Individu Penerima Bantuan
3. Penyaluran Bantuan
4. Verifikasi & Validasi Lapangan

B. Pendidikan

- 5. Data Siswa
- 6. Data Sekolah
- 7. Data Guru
- 8. Risiko Dropout
- 9. Nilai & Kehadiran

C. Kesehatan

- 10. Data Balita
- 11. Status Gizi
- 12. Puskesmas & Layanan
- 13. Vaksinasi
- 14. Early Warning Kesehatan

D. UMKM

- 15. Data UMKM
- 16. Transaksi & Omzet
- 17. Risiko UMKM
- 18. Bantuan Modal
- 19. Data Legalitas & NIB

E. Fiskal Daerah

- 20. Belanja Daerah
- 21. Pendapatan Asli Daerah (PAD)
- 22. Pembiayaan
- 23. Proyeksi Fiskal
- 24. Realisasi Anggaran

F. Metadata

- 25. Metadata Pelaporan Pemerintah

Setiap schema memiliki definisi field standar, validasi, dan aturan interoperabilitas nasional.

Manfaat Praktis untuk Pemerintah

Untuk Kabupaten/Kota

- operator desa tidak perlu menebak format data
- laporan langsung konsisten
- lebih sedikit koreksi dari provinsi

Untuk Provinsi

- validasi otomatis
- integrasi 514 kabupaten menjadi lebih mudah
- dashboard provinsi tidak perlu pembersihan manual

Untuk Kementerian

- data nasional bisa dianalisis tanpa pre-processing besar
- AI sektor kesehatan/pendidikan bisa dilatih langsung
- pelaporan cepat, akurat, dan dapat diverifikasi

Untuk BSSN / GovCSIRT

- struktur data memudahkan audit keamanan
- data sensitif bisa diberi aturan masking/hashing

Contoh Implementasi Open Source

JSON Schema Nasional langsung dapat digunakan di tool open-source:

Validasi Data

- **ajv** (JavaScript)
- **pyjsonschema** (Python)
- **jsonschema-cli**

ETL & Pipeline

- **n8n**
- **Apache NiFi**
- **Airflow**

Database

- **PostgreSQL + pg_validate_json**
- **MongoDB Schema Validation**

API Gateway

- **Kong** → validasi schema sebelum request diterima
- **API7** → enforce schema pada semua API kementerian

Data Catalog

- **CKAN** → menyimpan schema & dokumentasi

AI Training

- Schema memastikan dataset rapi sehingga AI dapat dilatih tanpa cleaning besar.

Tabel JSON Schema Nasional

Kategori	Contoh Schema	Manfaat	Tool Open Source
Bansos	data rumah tangga, penyaluran	tepat sasaran	n8n, PostgreSQL
Pendidikan	siswa, guru, kehadiran	deteksi dropout	ClickHouse, Superset
Kesehatan	balita, gizi, vaksin	prediksi stunting	Spark, Grafana
UMKM	data usaha, transaksi	prediksi risiko	Metabase, Scikit-Learn
Fiskal	PAD, belanja	anggaran presisi	Superset, Prophet
Metadata	struktur pelaporan	audit & interoperabilitas	CKAN

Pesan Inti

- 25 JSON Schema Nasional adalah **tulang punggung interoperabilitas RI**.
- Tanpa schema ini, data tidak kompatibel dan AI Nasional tidak mungkin berjalan.
- Dengan schema yang baku, seluruh pipeline data pemerintah dapat berjalan otomatis dari desa → kabupaten → provinsi → pusat.
- Open-source membuat implementasinya **murah, cepat, dan aman**.

LAMPIRAN — Template JSON Schema

Siap Pak. Saya buat **template teknis 25 JSON Schema** dalam bentuk **versi siap pakai**, dengan **nama dan struktur** yang mudah dipakai pemda/kementerian.

Konvensi umum yang saya pakai:

- Semua schema pakai:
`"$schema": "https://json-schema.org/draft/2020-12/schema"`
- Ada blok **wilayah** dan **pelaporan** di hampir semua schema.
- ID schema saya contohkan di domain fiktif: `https://data.go.id/schema/...` (bisa Bapak ganti).
- Bisa divalidasi pakai **ajv**, **python-jsonschema**, **Kong plugin**, atau bahkan di **PostgreSQL**.

Agar tidak terlalu berat, saya buat **field inti** saja (bisa Bapak extend sendiri).

0. Blok Standar (bisa di-copy ke semua schema)

Silakan copy blok ini ke masing-masing schema:

```
{
  "wilayah": {
    "type": "object",
    "properties": {
      "kode_prov": { "type": "string", "pattern": "^([0-9]{2})$" },
      "kode_kab": { "type": "string", "pattern": "^([0-9]{4})$" },
      "kode_kec": { "type": "string", "pattern": "^([0-9]{6})$" },
      "kode_desa": { "type": "string", "pattern": "^([0-9]{10})$" }
    },
    "required": ["kode_prov", "kode_kab", "kode_kec", "kode_desa"],
    "additionalProperties": false
  },
  "pelaporan": {
    "type": "object",
    "properties": {
      "tanggal_lapor": { "type": "string", "format": "date" },
      "periode": { "type": "string" },
      "sumber_data": { "type": "string" },
      "petugas_id": { "type": "string" },
      "schema_version": { "type": "string" }
    },
    "required": ["tanggal_lapor", "sumber_data"],
    "additionalProperties": false
  }
}
```

Di bawah ini 25 schema sektor-sektor.

A. BANSOS

1) bansos_rumah_tangga.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/bansos_rumah_tangga.json",
  "title": "Rumah Tangga Bansos",
  "type": "object",
  "properties": {
    "rumah_tangga_id": { "type": "string" },
    "nik_kepala": { "type": "string", "pattern": "^[0-9]{16}$" },
    "nama_kepala": { "type": "string", "minLength": 1 },
    "jumlah_anggota": { "type": "integer", "minimum": 1 },
    "alamat": { "type": "string" },
    "status_miskin": {
      "type": "string",
      "enum": ["ekstrem", "miskin", "rentan", "tidak_miskin"]
    },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "rumah_tangga_id",
    "nik_kepala",
    "nama_kepala",
    "jumlah_anggota",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": {
    "wilayah": {
      "type": "object",
      "properties": {
        "kode_prov": { "type": "string", "pattern": "^[0-9]{2}$" },
        "kode_kab": { "type": "string", "pattern": "^[0-9]{4}$" },
        "kode_kec": { "type": "string", "pattern": "^[0-9]{6}$" },
        "kode_desa": { "type": "string", "pattern": "^[0-9]{10}$" }
      },
      "required": ["kode_prov", "kode_kab", "kode_kec", "kode_desa"],
      "additionalProperties": false
    },
    "pelaporan": {
      "type": "object",
      "properties": {
        "tanggal_lapor": { "type": "string", "format": "date" },
        "periode": { "type": "string" },
        "sumber_data": { "type": "string" },
        "petugas_id": { "type": "string" },
        "schema_version": { "type": "string" }
      },
      "required": ["tanggal_lapor", "sumber_data"],
    }
  }
}
```

```

        "additionalProperties": false
    }
}
}

```

2) bansos_individu.json

```

{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/bansos_individu.json",
  "title": "Individu Penerima Bansos",
  "type": "object",
  "properties": {
    "penerima_id": { "type": "string" },
    "rumah_tangga_id": { "type": "string" },
    "nik": { "type": "string", "pattern": "^[0-9]{16}$" },
    "nama": { "type": "string" },
    "tanggal_lahir": { "type": "string", "format": "date" },
    "jenis_kelamin": { "type": "string", "enum": ["L", "P"] },
    "kategori_bansos": {
      "type": "string",
      "enum": ["PKH", "BPNT", "BLT", "Lainnya"]
    },
    "aktif": { "type": "boolean" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "penerima_id",
    "rumah_tangga_id",
    "nik",
    "nama",
    "kategori_bansos",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}

```

3) bansos_penyaluran.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/bansos_penyaluran.json",
  "title": "Penyaluran Bansos",
  "type": "object",
  "properties": {
    "transaksi_id": { "type": "string" },
    "penerima_id": { "type": "string" },
    "program": {
      "type": "string",
      "enum": ["PKH", "BPNT", "BLT", "Lainnya"]
    },
    "periode_bulan": { "type": "string", "pattern": "^[0-9]{4}-[0-9]{2}$" },
    "nilai": { "type": "number", "minimum": 0 },
    "metode_salur": { "type": "string", "enum": ["tunai", "rekening",
"e-wallet"] },
    "status": { "type": "string", "enum": ["berhasil", "gagal", "pending"]
  },
  "wilayah": { "$ref": "#/$defs/wilayah" },
  "pelaporan": { "$ref": "#/$defs/pelaporan" }
},
  "required": [
    "transaksi_id",
    "penerima_id",
    "program",
    "periode_bulan",
    "nilai",
    "status",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

4) bansos_verifikasi.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/bansos_verifikasi.json",
  "title": "Verifikasi & Validasi Bansos",
  "type": "object",
  "properties": {
    "vv_id": { "type": "string" },
    "penerima_id": { "type": "string" },
    "hasil_verifikasi": {
      "type": "string",
      "enum": ["layak", "tidak_layak", "perlu_kunjungan"]
    },
    "tanggal_verifikasi": { "type": "string", "format": "date" },
    "petugas_nik": { "type": "string", "pattern": "^[0-9]{16}$" },
    "catatan": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "vv_id",
    "penerima_id",
    "hasil_verifikasi",
    "tanggal_verifikasi",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

B. PENDIDIKAN

5) edu_siswa.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/edu_siswa.json",
  "title": "Data Siswa",
  "type": "object",
  "properties": {
    "siswa_id": { "type": "string" },
    "nisn": { "type": "string", "pattern": "^([0-9]{10})$" },
    "nik": { "type": "string", "pattern": "^([0-9]{16})$" },
    "nama": { "type": "string" },
    "jenis_kelamin": { "type": "string", "enum": ["L", "P"] },
    "tanggal_lahir": { "type": "string", "format": "date" },
    "sekolah_id": { "type": "string" },
    "kelas": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": ["siswa_id", "nisn", "nama", "sekolah_id", "wilayah",
    "pelaporan"],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

6) edu_sekolah.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/edu_sekolah.json",
  "title": "Data Sekolah",
  "type": "object",
  "properties": {
    "sekolah_id": { "type": "string" },
    "npsn": { "type": "string" },
    "nama_sekolah": { "type": "string" },
    "jenjang": { "type": "string", "enum": ["SD", "SMP", "SMA", "SMK",
    "PAUD"] },
    "status": { "type": "string", "enum": ["negeri", "swasta"] },
    "alamat": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": ["sekolah_id", "nama_sekolah", "jenjang", "wilayah",
    "pelaporan"],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

7) edu_guru.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/edu_guru.json",
  "title": "Data Guru",
  "type": "object",
  "properties": {
    "guru_id": { "type": "string" },
    "nuptk": { "type": "string" },
    "nik": { "type": "string", "pattern": "^[0-9]{16}$" },
    "nama": { "type": "string" },
    "sekolah_id": { "type": "string" },
    "status": { "type": "string", "enum": ["PNS", "PPPK", "Honorer"] },
  },
  "mapel_utama": { "type": "string" },
  "wilayah": { "$ref": "#/$defs/wilayah" },
  "pelaporan": { "$ref": "#/$defs/pelaporan" },
},
  "required": ["guru_id", "nik", "nama", "sekolah_id", "status",
"wilayah", "pelaporan"],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

8) edu_dropout_risk.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/edu_dropout_risk.json",
  "title": "Risiko Dropout Siswa",
  "type": "object",
  "properties": {
    "record_id": { "type": "string" },
    "siswa_id": { "type": "string" },
    "periode_bulan": { "type": "string", "pattern":
"^[0-9]{4}-[0-9]{2}$" },
    "skor_risiko": { "type": "number", "minimum": 0, "maximum": 1 },
    "kategori_risiko": {
      "type": "string",
      "enum": ["rendah", "sedang", "tinggi"]
    },
    "sumber_model": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "record_id",
    "siswa_id",
    "periode_bulan",
    "skor_risiko",
    "kategori_risiko",
    "wilayah",

```

```

        "pelaporan"
    ],
    "additionalProperties": false,
    "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}

```

9) edu_kehadiran.json

```

{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/edu_kehadiran.json",
  "title": "Kehadiran Siswa",
  "type": "object",
  "properties": {
    "record_id": { "type": "string" },
    "siswa_id": { "type": "string" },
    "tanggal": { "type": "string", "format": "date" },
    "status_hadir": {
      "type": "string",
      "enum": ["hadir", "izin", "sakit", "alpa"]
    },
    "sekolah_id": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "record_id",
    "siswa_id",
    "tanggal",
    "status_hadir",
    "sekolah_id",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}

```

C. KESEHATAN

10) health_balita.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/health_balita.json",
  "title": "Data Balita",
  "type": "object",
  "properties": {
    "balita_id": { "type": "string" },
    "nik": { "type": "string", "pattern": "^[0-9]{16}$" },
    "nama": { "type": "string" },
    "tanggal_lahir": { "type": "string", "format": "date" },
    "jenis_kelamin": { "type": "string", "enum": ["L", "P"] },
    "ibu_nik": { "type": "string", "pattern": "^[0-9]{16}$" },
    "rumah_tangga_id": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": ["balita_id", "nama", "tanggal_lahir", "wilayah",
    "pelaporan"],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

11) health_status_gizi.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/health_status_gizi.json",
  "title": "Status Gizi Balita",
  "type": "object",
  "properties": {
    "record_id": { "type": "string" },
    "balita_id": { "type": "string" },
    "tanggal_ukur": { "type": "string", "format": "date" },
    "berat": { "type": "number", "minimum": 0 },
    "tinggi": { "type": "number", "minimum": 0 },
    "zscore_tbu": { "type": "number" },
    "kategori_gizi": {
      "type": "string",
      "enum": ["normal", "stunting", "gizi_kurang", "gizi_buruk"]
    },
    "puskesmas_id": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "record_id",
    "balita_id",
    "tanggal_ukur",

```

```

        "berat",
        "tinggi",
        "kategori_gizi",
        "wilayah",
        "pelaporan"
    ],
    "additionalProperties": false,
    "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}

```

12) health_faskes.json

```

{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/health_faskes.json",
  "title": "Fasilitas Kesehatan",
  "type": "object",
  "properties": {
    "faskes_id": { "type": "string" },
    "kode_faskes": { "type": "string" },
    "nama_faskes": { "type": "string" },
    "jenis": { "type": "string", "enum": ["puskesmas", "rumah_sakit",
"klinik"] },
    "kelas_rs": { "type": "string" },
    "alamat": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": ["faskes_id", "nama_faskes", "jenis", "wilayah",
"pelaporan"],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}

```

13) health_vaksinasi.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/health_vaksinasi.json",
  "title": "Data Vaksinasi",
  "type": "object",
  "properties": {
    "record_id": { "type": "string" },
    "nik": { "type": "string", "pattern": "^[0-9]{16}$" },
    "nama": { "type": "string" },
    "jenis_vaksin": { "type": "string" },
    "dosis_ke": { "type": "integer", "minimum": 1 },
    "tanggal_vaksin": { "type": "string", "format": "date" },
    "faskes_id": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "record_id",
    "nik",
    "jenis_vaksin",
    "dosis_ke",
    "tanggal_vaksin",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

14) health_ews_case.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/health_ews_case.json",
  "title": "Kasus Early Warning Kesehatan",
  "type": "object",
  "properties": {
    "kasus_id": { "type": "string" },
    "jenis_penyakit": { "type": "string" },
    "tanggal_lapor": { "type": "string", "format": "date" },
    "jumlah_kasus": { "type": "integer", "minimum": 0 },
    "status_ews": { "type": "string", "enum": ["hijau", "kuning",
"merah"] },
    "faskes_id": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "kasus_id",
    "jenis_penyakit",
    "tanggal_lapor",

```

```
    "jumlah_kasus",
    "status_ews",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

D. UMKM

15) `umkm_usaha.json`

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/umkm_usaha.json",
  "title": "Data UMKM",
  "type": "object",
  "properties": {
    "umkm_id": { "type": "string" },
    "nib": { "type": "string" },
    "nama_usaha": { "type": "string" },
    "sektor": { "type": "string" },
    "bentuk_badan": { "type": "string" },
    "npwp": { "type": "string" },
    "jumlah_pekerja": { "type": "integer", "minimum": 0 },
    "omzet_tahunan": { "type": "number", "minimum": 0 },
    "kategori_umkm": {
      "type": "string",
      "enum": ["mikro", "kecil", "menengah"]
    },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": ["umkm_id", "nama_usaha", "kategori_umkm", "wilayah",
    "pelaporan"],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

16) umkm_transaksi.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/umkm_transaksi.json",
  "title": "Transaksi UMKM",
  "type": "object",
  "properties": {
    "transaksi_id": { "type": "string" },
    "umkm_id": { "type": "string" },
    "tanggal": { "type": "string", "format": "date" },
    "jenis_transaksi": { "type": "string", "enum": ["penjualan",
"pembelian"] },
    "nilai": { "type": "number", "minimum": 0 },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "transaksi_id",
    "umkm_id",
    "tanggal",
    "jenis_transaksi",
    "nilai",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

17) umkm_risiko.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/umkm_risiko.json",
  "title": "Risiko UMKM",
  "type": "object",
  "properties": {
    "record_id": { "type": "string" },
    "umkm_id": { "type": "string" },
    "periode_bulan": { "type": "string", "pattern":
      "^[0-9]{4}-[0-9]{2}$" },
    "skor_risiko": { "type": "number", "minimum": 0, "maximum": 1 },
    "kategori_risiko": {
      "type": "string",
      "enum": ["rendah", "sedang", "tinggi"]
    },
    "sumber_model": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "record_id",
    "umkm_id",
    "periode_bulan",
    "skor_risiko",
    "kategori_risiko",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

18) umkm_bantuan.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/umkm_bantuan.json",
  "title": "Bantuan UMKM",
  "type": "object",
  "properties": {
    "bantuan_id": { "type": "string" },
    "umkm_id": { "type": "string" },
    "jenis_bantuan": { "type": "string" },
    "tanggal_salur": { "type": "string", "format": "date" },
    "nilai": { "type": "number", "minimum": 0 },
    "program": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "bantuan_id",
    "umkm_id",
    "jenis_bantuan",
    "tanggal_salur",
    "nilai",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

19) umkm_legalitas.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/umkm_legalitas.json",
  "title": "Legalitas UMKM",
  "type": "object",
  "properties": {
    "record_id": { "type": "string" },
    "umkm_id": { "type": "string" },
    "nib": { "type": "string" },
    "npwp": { "type": "string" },
    "izin_lain": { "type": "string" },
    "status_formal": { "type": "string", "enum": ["formal",
"non_formal"] },
    "tanggal_update": { "type": "string", "format": "date" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "record_id",
    "umkm_id",
    "status_formal",
    "tanggal_update",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

E. FISKAL DAERAH

20) `fiskal_belanja_daerah.json`

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/fiskal_belanja_daerah.json",
  "title": "Belanja Daerah",
  "type": "object",
  "properties": {
    "belanja_id": { "type": "string" },
    "tahun": { "type": "integer", "minimum": 2000 },
    "kode_akun": { "type": "string" },
    "nama_akun": { "type": "string" },
    "program": { "type": "string" },
    "kegiatan": { "type": "string" },
    "pagu": { "type": "number", "minimum": 0 },
    "realisasi": { "type": "number", "minimum": 0 },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "belanja_id",
    "tahun",
    "kode_akun",
    "nama_akun",
    "pagu",
    "realisasi",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

21) fiskal_pad.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/fiskal_pad.json",
  "title": "Pendapatan Asli Daerah (PAD)",
  "type": "object",
  "properties": {
    "pad_id": { "type": "string" },
    "tahun": { "type": "integer", "minimum": 2000 },
    "jenis_pad": { "type": "string" },
    "target": { "type": "number", "minimum": 0 },
    "realisasi": { "type": "number", "minimum": 0 },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "pad_id",
    "tahun",
    "jenis_pad",
    "target",
    "realisasi",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

22) fiskal_pembiayaan.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/fiskal_pembiayaan.json",
  "title": "Pembiayaan Daerah",
  "type": "object",
  "properties": {
    "pembiayaan_id": { "type": "string" },
    "tahun": { "type": "integer", "minimum": 2000 },
    "jenis_pembiayaan": { "type": "string" },
    "penerimaan": { "type": "number", "minimum": 0 },
    "pengeluaran": { "type": "number", "minimum": 0 },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "pembiayaan_id",
    "tahun",
    "jenis_pembiayaan",
    "penerimaan",
    "pengeluaran",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

23) fiskal_proyeksi.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/fiskal_proyeksi.json",
  "title": "Proyeksi Fiskal",
  "type": "object",
  "properties": {
    "proyeksi_id": { "type": "string" },
    "tahun": { "type": "integer", "minimum": 2000 },
    "skenario": { "type": "string", "enum": ["optimis", "moderate",
"pesimis"] },
    "belanja_proyeksi": { "type": "number", "minimum": 0 },
    "pendapatan_proyeksi": { "type": "number", "minimum": 0 },
    "defisit_proyeksi": { "type": "number" },
    "sumber_model": { "type": "string" },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "proyeksi_id",
    "tahun",
    "skenario",
    "belanja_proyeksi",
    "pendapatan_proyeksi",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

24) fiskal_realisasi_bulanan.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/fiskal_realisasi_bulanan.json",
  "title": "Realisasi Fiskal Bulanan",
  "type": "object",
  "properties": {
    "record_id": { "type": "string" },
    "tahun": { "type": "integer", "minimum": 2000 },
    "bulan": { "type": "integer", "minimum": 1, "maximum": 12 },
    "total_pendapatan": { "type": "number", "minimum": 0 },
    "total_belanja": { "type": "number", "minimum": 0 },
    "total_pembiayaan": { "type": "number", "minimum": 0 },
    "wilayah": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "record_id",
    "tahun",
    "bulan",
    "total_pendapatan",
    "total_belanja",
    "wilayah",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

F. METADATA & PELAPORAN

25) metadata_pelaporan.json

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "$id": "https://data.go.id/schema/metadata_pelaporan.json",
  "title": "Metadata Pelaporan Pemerintah",
  "type": "object",
  "properties": {
    "metadata_id": { "type": "string" },
    "nama_dataset": { "type": "string" },
    "deskripsi": { "type": "string" },
    "versi_schema": { "type": "string" },
    "schema_ref": { "type": "string", "format": "uri" },
    "instansi_pemilik": { "type": "string" },
    "frekuensi_pelaporan": {
      "type": "string",
      "enum": ["harian", "mingguan", "bulanan", "tahunan",
"insidental"]
    },
    "tingkat_agr": {
      "type": "string",
      "enum": ["individu", "rumah_tangga", "sekolah", "faskes",
"umkm", "wilayah"]
    },
    "kontak_pic": { "type": "string" },
    "wilayah_cakupan": { "$ref": "#/$defs/wilayah" },
    "pelaporan": { "$ref": "#/$defs/pelaporan" }
  },
  "required": [
    "metadata_id",
    "nama_dataset",
    "versi_schema",
    "schema_ref",
    "instansi_pemilik",
    "frekuensi_pelaporan",
    "pelaporan"
  ],
  "additionalProperties": false,
  "$defs": { "$ref": "bansos_rumah_tangga.json#/$defs" }
}
```

Ringkas: Cara Pakai di Open Source

- **Validasi di API Gateway (Kong / APISIX):**
→ gunakan plugin JSON Schema untuk request/response checking.
- **Validasi di ETL (n8n, NiFi, Airflow):**
→ buat node custom yang memanggil validator (ajv, python-jsonschema).
- **Validasi di DB (PostgreSQL):**
→ simpan payload JSONB dan tambahkan check constraint dengan fungsi PL/pgSQL yang memanggil validator atau minimal cek pattern dasar.

LAMPIRAN: Template Peraturan Daerah Data Governance

LAMPIRAN ini penting untuk **kepala daerah, biro hukum, Bappeda, Diskominfo, BPS daerah, dan Wali Data**. Di sini bukan lagi bicara konsep, tapi **tool politik & hukum**: *Template Peraturan Daerah (Perda) Data Governance* yang bisa langsung dipakai sebagai dasar gerak di daerah.

Mengapa Perda Data Governance Dibutuhkan?

Selama ini banyak pemda sudah bicara:

- “*Satu Data Daerah*”
- “*Digitalisasi pelayanan*”
- “*Smart city, smart village*”

Tapi tanpa **Perda**, semua itu hanya:

- jadi proyek sementara,
- tergantung semangat kepala dinas,
- mudah hilang saat pejabat ganti,
- sulit memaksa OPD lain ikut standar.

Perda Data Governance di lampiran ini dirancang sebagai:

- payung hukum daerah untuk **data, AI, dan keamanan**,
- pegangan bersama Bappeda, Diskominfo, BPS, Inspektorat, BPKAD, OPD sektor,
- penghubung antara kebijakan nasional (Satu Data, SPBE, UU PDP) dengan kondisi lokal.

Intinya:

Tanpa Perda, Data Governance hanya “imbauan”

Dengan Perda, Data Governance menjadi **kewajiban**.

Isi Utama Template Perda Data Governance (Disederhanakan)

Template di lampiran menyusun Perda dalam **struktur yang sangat praktis**, biasanya berisi bagian-bagian seperti:

Ketentuan Umum

- definisi *Data Daerah*, *Wali Data*, *Produsen Data*, *Data Pribadi*, *Open Data*
- rujukan ke **Perpres Satu Data Indonesia**, **SPBE**, dan **UU PDP**

Tujuan & Ruang Lingkup

- memperjelas bahwa Data Governance dipakai untuk:
 - perencanaan & penganggaran,
 - pelayanan publik,
 - pengawasan & evaluasi,
 - pemanfaatan AI & analitik.

Kelembagaan Data Daerah

- pembentukan **Dewan Data Daerah** atau forum sejenis,
- penetapan **Wali Data** per OPD,
- tugas Diskominfo/Bappeda sebagai koordinator data.

Standar Data & Interoperabilitas

- kewajiban OPD memakai **format baku (JSON Schema Nasional/Daerah)**,
- kewajiban setiap sistem baru → **punya API**,
- larangan membangun aplikasi silo tertutup.

Infrastruktur & Keamanan

- kewajiban penempatan data di **Pusat Data Daerah / GovCloud** yang memenuhi standar,
- pengamanan minimal: enkripsi, VPN, IAM, logging, SIEM,
- selaras dengan pedoman BSSN dan UU PDP.

Pendanaan & SDM

- data & infrastruktur digital **masuk dalam prioritas anggaran** daerah,
- dukungan pelatihan ASN di bidang data, keamanan, dan open source.

Ketentuan Pengawasan & Sanksi

- audit berkala data & keamanan,
- kewajiban perbaikan jika melanggar standar,
- sanksi administratif internal bagi OPD yang tidak patuh.

Template ini dibuat agar pemda **tidak mulai dari nol** saat menyusun Perdanya.

Penekanan Penggunaan Open Source di Dalam Perda

Salah satu poin kuat dari template ini:

Perda tidak hanya bicara “*pakai teknologi*”, tapi secara halus **mendorong penggunaan open source** untuk:

- menghemat anggaran,
- menghindari *vendor lock-in*,
- memperkuat kedaulatan data.

Contoh rumusan norma (dalam semangat template):

- “*Pemerintah Daerah mengutamakan penggunaan perangkat lunak open source untuk pengelolaan data, sepanjang memenuhi standar keamanan, keandalan, dan interoperabilitas.*”
- “*Setiap pengadaan sistem informasi harus mempertimbangkan interoperabilitas dan kompatibilitas dengan standar terbuka (open standard) dan/atau open source.*”

Contoh implementasi open source yang bisa “di-backup” oleh Perda:

- **Database:** PostgreSQL, MariaDB → gudang data daerah
- **Dashboard:** Apache Superset, Metabase → laporan ke kepala daerah
- **ETL/Integrasi:** n8n, Apache NiFi → tarik data dari OPD ke *data lake* daerah
- **Cloud & Storage:** OpenStack/Harvester, MinIO → GovCloud lokal
- **Keamanan:** Wazuh (SIEM), Suricata (IDS), pfSense/WireGuard (firewall & VPN)
- **Kolaborasi:** Nextcloud, GitLab CE → berbagi dokumen & kode antar OPD

Perda tidak perlu menyebut nama software satu per satu, cukup:

- menegaskan **prinsip open source & open standard**,
- mensyaratkan “*tidak tergantung vendor tunggal*”,
- membuka ruang bagi pemda memakai solusi murah dan mandiri.

Bagaimana Pemda Menggunakan Template Ini?

Template di lampiran dimaksudkan sebagai **draft awal**.

Bukan harus copy-paste mentah, tapi disesuaikan:

Langkah praktis:

1. **Biro Hukum + Diskominfo + Bappeda duduk bersama**
 - bahas struktur template,
 - sesuaikan dengan nama OPD, Perda lain, RPJMD.
2. **Masukkan konteks lokal**
 - prioritas daerah: bansos, kesehatan, pendidikan, UMKM, fiskal, dll.
 - kapasitas infrastruktur yang sudah ada.
3. **Sinkronkan dengan regulasi nasional**
 - Satu Data Indonesia, SPBE, UU PDP, BSSN, dan pedoman data sektoral.
4. **Bahas dengan DPRD**
 - jelaskan bahwa Perda ini bukan proyek IT, **tetapi instrumen efisiensi anggaran & tata kelola**.
 - tunjukkan contoh potensi penghematan melalui open source.
5. **Angkat ke dokumen turunan**
 - setelah Perda, susun Perkada/SK Kepala Daerah, SOP teknis, dan rencana aksi 1–3 tahun (roadmap daerah).

Poin Kritis yang Harus Dijaga dalam Perda

Agar Perda Data Governance **tidak menjadi macan kertas**, template ini pada dasarnya mendorong beberapa hal kritis:

- **Setiap sistem baru wajib interoperable**
→ ada API, ikut JSON Schema, terhubung dengan data daerah.
- **Setiap OPD wajib punya Wali Data**
→ tidak ada lagi data “milik pribadi dinas”.
- **Audit data & keamanan menjadi rutinitas, bukan insiden**
→ minimal tahunan, dengan indikator jelas (kualitas data, ketahanan siber, efisiensi).
- **Pelatihan ASN & penggunaan open source masuk ke dokumen resmi**
→ bukan hanya wacana, tapi ada dasar hukum untuk anggaran.
- **Perda mengaitkan Data Governance dengan efisiensi anggaran & pelayanan publik**
→ sehingga dukungan politik lebih kuat.

Pesan Inti untuk Kepala Daerah & DPRD

- Template Perda ini bukan dokumen teknis biasa.
Ini adalah “**kontrak sosial dan hukum**” antara pemerintah daerah dan warga, bahwa **data akan dikelola dengan baik, aman, dan dipakai untuk meningkatkan kesejahteraan**.
- Dengan menyusun Perda berbasis template ini, daerah:
 - punya dasar hukum kuat untuk memaksa standar data,
 - bisa mendorong open source untuk hemat anggaran,
 - bisa menyiapkan pondasi AI dan GovCloud lokal,
 - bisa memastikan program bansos, kesehatan, pendidikan, UMKM, dan fiskal berjalan **lebih presisi**.

LAMPIRAN — Tool Open Source + Panduan Deploy

Lampiran ini adalah salah satu bagian paling penting dalam dokumen. Ini adalah **daftar final** tool open-source yang dapat dipakai pemda, provinsi, dan kementerian **tanpa biaya lisensi**, **tanpa vendor lock-in**, dan **dengan kemampuan yang setara—bahkan lebih baik—dibanding aplikasi komersial**.

Lampiran ini menjawab masalah klasik pemerintah:

- anggaran IT terbatas,
- daerah kecil tidak punya vendor,
- aplikasi tidak interoperable,
- sulit deploy,
- tidak ada standar,
- sering tergantung vendor tunggal,
- keamanan lemah.

Dengan 50 tools ini, pemerintah daerah sebenarnya bisa membangun:

- **GovCloud lokal**,
- **Data Warehouse nasional**,
- **ETL lintas OPD**,
- **Dashboard pimpinan**,
- **AI lokal**,
- **Keamanan data**,
- **Early Warning System**,
- dengan biaya sangat rendah.

Semua tools juga dapat di-deploy menggunakan:

- **Ubuntu Server LTS**,
- **Docker**,
- **Docker Compose**,
- atau **Kubernetes** (jika skala besar).

Filosofi Tools Open Source

Filosofi dari lampiran ini sederhana:

*“Pemda harus bisa membangun sistem sendiri.
Tanpa menunggu vendor.
Tanpa biaya besar.
Tanpa tergantung pihak luar.”*

Tools yang dipilih memenuhi 5 kriteria:

1. **Open-source murni**
2. **Stabil & terbukti dipakai pemerintah dunia**
3. **Ringan & mudah dipasang**
4. **Kompatibel dengan Data Governance Nasional**
5. **Bisa berjalan di server kecil maupun GovCloud**

Kategori Tools

Lampiran mengelompokkan tools dalam beberapa kategori praktis:

A. Database dan Data Warehouse

Tools untuk menyimpan data standar nasional.

- **PostgreSQL** → DW nasional
- **MariaDB**
- **ClickHouse** → analitik cepat
- **DuckDB** → OLAP ringan
- **TimescaleDB** → time-series kesehatan / IoT
- **SQLite** → offline desa/kelurahan

Contoh deploy:

```
docker run -d --name postgres \
  -e POSTGRES_PASSWORD=rahasia \
  -p 5432:5432 postgres:15
```

B. ETL / Integrasi Data

Untuk menarik data dari OPD → kab → prov → pusat.

- **n8n** → no-code ETL
- **Apache NiFi**
- **Airflow**
- **Meltano**
- **Singer.io** → data pipelines standar
- **Logstash**

Contoh deploy n8n:

```
docker run -d -p 5678:5678 n8nio/n8n
```

C. Dashboard & BI untuk Pimpinan

Untuk menampilkan analitik kepada bupati / gubernur.

- **Apache Superset**
- **Metabase**
- **Grafana**
- **Redash**

Contoh deploy Superset:

```
docker-compose up -d
```

D. Big Data & Streaming

Untuk data skala besar lintas daerah.

- **Apache Spark**
- **Hadoop**
- **Apache Flink**
- **Kafka**
- **Pinot** (real-time analytics)

E. AI & Machine Learning

Untuk AI Nasional dan AI Daerah.

- **Ollama** → LLM lokal
- **HuggingFace** → model open-source
- **TensorFlow / PyTorch**
- **MLflow** → manajemen model
- **Kubeflow**

Contoh deploy Ollama:

```
curl https://ollama.ai/install.sh | sh
ollama run llama3
```

F. Data Lake & Storage

Untuk penyimpanan data besar, aman, dan murah.

- **MinIO** → pengganti S3
- **Ceph**
- **OpenEBS**
- **HDFS**

G. API Gateway & Interoperabilitas

Untuk membangun Sistem Pertukaran Data Nasional.

- **Kong Gateway**
- **APISIX**
- **KrakenD**
- **Traefik**

Contoh kong:

```
docker run -d --name kong \
  -p 8000:8000 kong:latest
```

H. Cybersecurity

Open-source untuk keamanan nasional.

- **Wazuh** → SIEM nasional
- **Suricata** → IDS/IPS
- **pfSense / OPNsense** → firewall
- **Keycloak** → IAM
- **WireGuard** → VPN antar instansi
- **OpenVAS** → vulnerability scan
- **TheHive** → incident response
- **Velociraptor** → forensik digital

I. DevOps & Automation

Agar sistem pemerintah bisa berjalan stabil.

- **Docker**
- **Kubernetes (k3s / k8s)**
- **Ansible**
- **GitLab CE**
- **Harbor Registry**

J. Kolaborasi & Arsip Dokumen

Untuk dokumen OPD, lintas kabupaten, dan provinsi.

- **Nextcloud**
- **ONLYOFFICE**
- **Mattermost / Rocket.Chat**
- **OpenProject**
- **Jitsi** → video meeting

10 Tools Paling Penting Bagi Kabupaten/Kota

Lampiran ini secara halus menganjurkan 10 tools ini berikut karena:

- mudah,
- ringan,
- biaya rendah,
- langsung berdampak.

Top 10 tools:

1. **PostgreSQL**
2. **Superset**
3. **Metabase**
4. **n8n**
5. **MinIO**
6. **Wazuh**
7. **WireGuard**
8. **Keycloak**
9. **Nextcloud**
10. **Ollama**

Kabupaten yang memakai 10 tools ini **sudah memenuhi 70% kebutuhan Data Governance**.

Panduan Deploy Singkat (Untuk Daerah)

Lampiran menyediakan panduan praktis berbasis **Ubuntu 22.04/24.04 + Docker**:

Langkah cepat 5 menit:

1) Install Docker

```
apt update && apt install docker.io docker-compose -y
```

2) Deploy PostgreSQL

```
docker run -d --name postgres \  
-e POSTGRES_PASSWORD=xxxx \  
-p 5432:5432 postgres:15
```

3) Deploy n8n

```
docker run -d -p 5678:5678 n8nio/n8n
```

4) Deploy Superset

```
docker-compose -f superset.yml up -d
```

5) Deploy Wazuh (keamanan)

```
docker-compose -f wazuh.yml up -d
```

Dalam 30 menit, pemda sudah bisa punya:

- data warehouse,
- ETL,
- dashboard pimpinan,
- AI lokal,
- keamanan dasar,
- penyimpanan data.

Semua tanpa lisensi.

Tabel Open Source Tools & Fungsinya

Kategori	Tools	Fungsi Utama
Database	PostgreSQL, MariaDB, ClickHouse	DW & analitik
Dashboard	Superset, Metabase	laporan pimpinan
ETL	n8n, NiFi, Airflow	integrasi OPD
Big Data	Spark, Kafka, Pinot	streaming & analitik
AI	Ollama, HF, Kubeflow	AI nasional
Storage	MinIO, Ceph	data lake
API	Kong, APISIX	pertukaran data
Cybersecurity	Wazuh, Suricata	SIEM & IDS
VPN & IAM	WireGuard, Keycloak	keamanan akses
DevOps	Docker, GitLab CE	otomatisasi
Kolaborasi	Nextcloud, Jitsi	dokumen & meeting

Pesan Inti

- Pemerintah **tidak butuh software mahal** untuk membangun Data Governance.
- tools open-source di lampiran sudah **lebih dari cukup** untuk membangun sistem data setingkat nasional.
- Semua tools bisa di-deploy **cepat, murah, dan aman**.
- Daerah kecil pun bisa membangun GovCloud sendiri.
- Open-source adalah **jalan tercepat menuju kedaulatan data Indonesia**.

LAMPIRAN — Draft Perpres Data Governance & GovCloud

Ditulis agar **birokrat, perumus kebijakan, Bappenas, KemenPANRB, Kominfo, BSSN, BPKAD, dan pengelola data daerah** langsung mengerti *apa esensinya, apa yang harus diwajibkan, dan apa implementasinya*.

Lampiran ini adalah pondasi hukum yang sangat penting. Jika Perda adalah pondasi daerah, maka **Perpres Data Governance & GovCloud** adalah **fondasi negara**.

Tanpa Perpres:

- data antar kementerian tidak terhubung,
- GovCloud tidak punya standar,
- interoperabilitas tidak wajib,
- keamanan tidak konsisten,
- AI Nasional tidak punya sumber data,
- anggaran negara terus bocor karena duplikasi sistem.

Dengan Perpres ini:

- seluruh kementerian, provinsi, dan kabupaten **wajib mengikuti standar tunggal**,
- pembangunan data nasional **tidak bergantung pada pemimpin atau proyek**,
- interoperabilitas menjadi **mandat**, bukan imbauan,
- Indonesia bisa membangun **GovCloud yang aman, hemat, dan souverain**.

Tujuan Utama Draft Perpres (Disederhanakan)

Perpres ini dirancang untuk menjawab 5 masalah klasik:

- **Data pemerintah tersebar dan tidak sinkron.**
Perpres menegaskan **Single Source of Truth** nasional.
- **Tidak ada standar pertukaran data lintas lembaga.**
Perpres mewajibkan penggunaan **JSON Schema Nasional** yang telah ditetapkan di lampiran terpisah.
- **Infrastruktur data tidak merata dan bergantung vendor.**
Perpres mengatur **GovCloud Nasional** berbasis **open-source**.
- **Keamanan tidak konsisten antar instansi.**
Perpres menetapkan **baseline keamanan** nasional: IAM, enkripsi, SIEM, VPN.
- **AI Nasional tidak mungkin berjalan karena data kacau.**
Perpres menetapkan bahwa semua data pemerintah harus machine-readable dan AI-ready.

Struktur Draft Perpres (Disederhanakan)

Draft Perpres di lampiran memiliki bagian-bagian berikut:

A. Ketentuan Umum

- definisi: Data Pemerintah, Interoperabilitas, GovCloud, Wali Data Nasional
- penegasan prinsip: integritas, keamanan, interoperabilitas

B. Kebijakan Data Nasional

- semua instansi wajib mengikuti standar nasional
- data prioritas ditetapkan oleh **Dewan Data Nasional**
- tidak boleh membangun sistem silo

C. Arsitektur Data Nasional

- mewajibkan penggunaan:
 - Data Lake Nasional
 - Data Warehouse Nasional
 - Data Exchange Nasional

D. GovCloud Nasional

- GovCloud adalah *infrastruktur negara*
- pengelolaan dapat hybrid:
 - PDN, PDNS, GovCloud provinsi
- mewajibkan open-source: OpenStack, Harvester, MinIO

E. Keamanan Siber

- baseline keamanan BSSN menjadi **mandatory**
- SIEM nasional wajib (Wazuh)
- VPN antar instansi: WireGuard
- IAM nasional: Keycloak

F. Standar Data dan JSON Nasional

- setiap data yang dikirim antar instansi → **harus pakai JSON Schema Nasional**
- setiap sistem baru → wajib menyediakan **API terbuka**
- setiap aplikasi pemerintah → harus *interoperable-ready*

G. Pemanfaatan Data & AI Nasional

- data harus AI-ready
- data bersih, rapi, terstandar
- prioritas sektor: kesehatan, bansos, pendidikan, fiskal

H. Pengawasan & Audit

- audit kualitas data nasional
- audit keamanan siber
- audit GovCloud
- audit interoperabilitas

Prinsip “Open Source First” dalam Draft Perpres

mendorong open source sebagai pilar kedaulatan data nasional.

Rumusan norma yang seirama dengan dokumen:

- *“Pemerintah menerapkan prinsip Open Standards dan mengutamakan Open Source untuk mendukung interoperabilitas dan efisiensi anggaran.”*
- *“Pemanfaatan GovCloud menggunakan platform open-source yang memenuhi standar keamanan nasional.”*

Tools open-source yang menjadi referensi:

Infrastruktur GovCloud

- OpenStack
- Harvester
- Proxmox
- MinIO (pengganti S3)

Data Governance

- PostgreSQL
- n8n
- Kafka
- CKAN
- Superset / Metabase

Keamanan Nasional

- Wazuh (SIEM)
- Suricata (IDS/IPS)
- WireGuard (VPN)
- Keycloak (IAM)

Perpres ini memberikan **justifikasi hukum** untuk menggunakan open-source di seluruh pemerintahan.

Dampak Praktis Perpres Ini bagi Pemerintah

Untuk Presiden dan Menteri

- kebijakan berbasis data menjadi standar
- semua kementerian wajib ikut alur interoperabilitas

Untuk Provinsi

- memiliki dasar hukum kuat untuk membangun GovCloud daerah
- bisa memaksa kabupaten mengikuti standar schema

Untuk Kabupaten/Kota

- tidak perlu membangun aplikasi sendiri-sendiri
- cukup mengirim data ke provinsi/pusat dengan JSON Schema Nasional

Untuk BSSN

- baseline keamanan menjadi nasional, bukan sektoral
- memudahkan audit daerah

Untuk K/L sektor

- data lintas kementerian otomatis “nyambung”
- dashboard nasional dapat dibangun cepat

Contoh Poin-Poin Kritis yang Perpres Wajib Tegaskan

Berikut ringkasan tajam yang mudah dipahami pimpinan:

- **Semua aplikasi pemerintah wajib punya API**
- **Semua data lintas instansi wajib pakai JSON Schema Nasional**
- **Semua instansi wajib masuk GovCloud—bukan server sendiri-sendiri**
- **Standar keamanan BSSN berlaku otomatis**
- **Open-source diutamakan**
- **AI Nasional hanya memakai data bersih, terstandar, dan tersinkron**
- **Audit data menjadi keharusan—bukan pilihan**

Dengan poin-poin ini, Perpres menjadi *tool transformasi negara*, bukan dokumen simbolis.

Tabel Ringkas — Isi Draft Perpres

Komponen	Isi Inti	Teknologi Open Source Pendukung
Kebijakan Data Nasional	standar, interoperabilitas, JSON nasional	JSON Schema, CKAN
GovCloud Nasional	cloud pemerintah berbasis open-source	OpenStack, Harvester, MinIO
Data Exchange Nasional	API wajib, schema wajib	Kong, APISIX
Data Warehouse Nasional	konsolidasi data lintas sektor	PostgreSQL, ClickHouse
Keamanan Siber	baseline BSSN, SIEM nasional	Wazuh, Suricata, VPN WireGuard
IAM Nasional	identitas & akses	Keycloak
AI Nasional	data AI-ready	Ollama, HF, Spark

Pesan Inti

- Draft Perpres ini adalah **mesin penyelaras negara**.
- Dengan Perpres, data antar instansi **harus** nyambung.
- Dengan GovCloud nasional, negara **memiliki kedaulatan digital**.
- Dengan open-source, pemerintah **hemat triliunan rupiah** dan bebas vendor.
- Dengan standar JSON nasional, AI Nasional akhirnya **punya bensin yang benar**.

LAMPIRAN — Contoh Dashboard Nasional

Dashboard Nasional adalah tool utama untuk memastikan pemerintah bekerja **berbasis data**, bukan asumsi. Lampiran ini menunjukkan **contoh konkret** bagaimana data dari kabupaten → provinsi → kementerian bisa divisualisasikan secara seragam, cepat, dan mudah dibaca pimpinan.

Tujuan utama dashboard nasional:

- mempermudah Presiden/Menteri/Gubernur/Bupati memahami situasi terkini,
- mengurangi rapat panjang hanya untuk “klarifikasi data”,
- mempercepat keputusan,
- memastikan anggaran tepat sasaran,
- memberi sistem *early warning* dari pusat ke daerah.

Dashboard yang baik **tidak rumit**.

Ia harus sederhana dan langsung menunjukkan **angka penting**, bukan hiasan visual.

Prinsip Dashboard Nasional yang Baik

Bapak menekankan bahwa dashboard harus:

- **Sederhana**
1 halaman cukup.
- **Fokus pada 5–10 indikator penting**
Bukan ratusan grafik.
- **Real-time atau harian**
Agar pimpinan tidak mengambil keputusan dengan data lama.
- **Lintas sektor**
Bansos, pendidikan, kesehatan, UMKM, fiskal → satu layar.
- **Dibangun dengan open-source**
Supaya daerah bisa menduplikasi tanpa biaya lisensi.

Contoh Kategori Dashboard Nasional

Lampiran menyediakan beberapa contoh dashboard nasional yang bisa ditiru. Penjelasan berikut disederhanakan agar mudah dipahami oleh OPD.

A. Dashboard Bansos Nasional

Indikator inti:

- jumlah keluarga miskin terverifikasi
- penerima bansos aktif
- tingkat duplikasi data
- kecepatan penyaluran

- error rate data kabupaten
- peta sebaran rumah tangga rentan

Contoh open-source:

Superset → heatmap penerima bantuan per kecamatan

n8n → pipeline validasi harian

PostgreSQL → dedup (soundex, fuzzy match)

B. Dashboard Pendidikan Nasional

Indikator inti:

- risiko dropout per sekolah
- kehadiran siswa harian
- rasio guru-siswa
- tingkat literasi numerasi
- kondisi sarpras
- data afirmasi (PIP) tepat sasaran

Contoh open-source:

ClickHouse → analitik ribuan sekolah

Superset → peta risiko dropout

Ollama/Scikit-learn → prediksi dropout

C. Dashboard Kesehatan Nasional

Indikator inti:

- status gizi balita
- potensi stunting
- kapasitas Puskesmas
- rawat jalan & rawat inap harian
- cakupan imunisasi
- early warning penyakit menular

Contoh open-source:

Grafana → monitoring Puskesmas

TimescaleDB → data kesehatan real-time

Spark MLlib → prediksi stunting

D. Dashboard UMKM Nasional

Indikator inti:

- jumlah UMKM aktif
- omzet rata-rata
- tingkat gagal usaha
- permohonan bantuan
- kredit/pembiayaan
- pemetaan sektor (kuliner, craft, jasa)

Contoh open-source:

Metabase → insight untuk Kemenkop

ClickHouse → analitik omzet

Kafka → streaming transaksi

E. Dashboard Fiskal Nasional

Indikator inti:

- realisasi belanja
- realisasi PAD
- rasio belanja wajib
- neraca transfer ke daerah
- proyeksi fiskal 1–3 tahun
- alert penyimpangan pengeluaran

Contoh open-source:

Superset → waterfall chart APBN/APBD

Prophet → forecasting belanja

PostgreSQL → data warehouse fiskal

Mengapa Dashboard Nasional Penting untuk Birokrat?

Karena birokrat perlu tool yang:

- tidak rumit,
- tidak butuh pelatihan teknis,
- bisa dibuka di ponsel,
- laporan bisa keluar 5–10 detik,
- dan langsung menjawab pertanyaan:

“Apa masalah utama hari ini? Daerah mana perlu intervensi?”

Dashboard bukan “hiasan digital”.

Dashboard adalah **tool kendali negara**.

Cara Cepat Membangun Dashboard Nasional (Pola Open Source)

Langkah 1 — Data Masuk

n8n / NiFi menarik data dari:

- kabupaten/kota
- provinsi
- instansi pusat
- server daerah
- API daring

Langkah 2 — Penyimpanan

PostgreSQL / ClickHouse sebagai Data Warehouse.

Langkah 3 — Query Analitik

Disusun oleh:

- analis data OPD
- Bappeda
- Dinas teknis

Langkah 4 — Visualisasi

Superset / Metabase:

- grafik garis, batang, peta, ringkasan KPI
- dashboard 1 halaman
- filter per provinsi/kabupaten

Langkah 5 — Monitoring

Grafana untuk data real-time (contoh: kesehatan, IoT, Puskesmas).

Contoh Hasil Dashboard (disederhanakan)

Contoh “Dashboard Nasional – Satu Layar”:

- **Bansos:** 92% tepat sasaran, 108.234 data ganda terdeteksi
- **Pendidikan:** 1.832 siswa berisiko dropout (high-risk)
- **Kesehatan:** 12 kabupaten potensi stunting >25%
- **UMKM:** 14.200 UMKM berisiko tutup (tren menurun 3 bulan)
- **Fiskal:** realisasi 68%, defisit 2,3%

Satu layar.

Satu tampilan.

Satu sumber kebenaran.

(Single Source of Truth)

Inilah yang negara butuhkan.

Tabel Ringkas Dashboard Nasional

Sektor	Indikator Kunci	Tools Open Source	Manfaat
Bansos	duplikasi, penerima aktif, peta risiko	Superset, PostgreSQL, n8n	tepat sasaran
Pendidikan	dropout, kehadiran, sarpras	ClickHouse, Superset, Ollama	intervensi cepat
Kesehatan	gizi, imunisasi, outbreak	Grafana, Spark, TimescaleDB	early warning
UMKM	omzet, risiko, bantuan	Metabase, ClickHouse	bantuan lebih presisi
Fiskal	belanja, PAD, realisasi	Superset, Prophet	anggaran efisien

Pesan Inti

- Dashboard Nasional adalah **instrumen pengendalian negara**, bukan dekorasi digital.
- Harus sederhana, cepat, dan berbasis data standar nasional.
- Dengan open-source, pemerintah dapat membangun dashboard nasional **tanpa biaya mahal, tanpa vendor, dan dengan kedaulatan penuh**.
- Dashboard adalah pintu masuk menuju **decision-making berbasis data** di seluruh Indonesia.

Glossary

A

API (Application Programming Interface)

Gerbang komunikasi antar aplikasi agar bisa saling bertukar data secara otomatis.

API Gateway Nasional

Gerbang pusat untuk seluruh akses data pemerintah. Menghapus API liar dan memastikan keamanan serta standar.

Apache Kafka

Message bus open source untuk streaming data real-time skala nasional.

Apache Spark

Mesin analitik cepat untuk big data dan AI.

Audit Trail

Jejak siapa mengakses data, kapan, dan apa yang dilakukan. Penting untuk keamanan dan akuntabilitas.

B

Bansos (Bantuan Sosial)

Program bantuan pemerintah yang membutuhkan data kependudukan dan kemiskinan yang sangat akurat.

Batch Processing

Pemrosesan data secara berkala (harian/mingguan). Cocok untuk laporan rutin.

Big Data

Data berukuran besar dan cepat, memerlukan tool khusus (Hadoop, Spark, Kafka).

C

CDO (Chief Data Officer)

Pimpinan strategi data di K/L. Bertanggung jawab atas kualitas, integrasi, dan pemanfaatan data.

Ceph

Penyimpanan terdistribusi open-source untuk PDN/PDD.

ClickHouse

Database analitik super cepat untuk dashboard nasional.

Cloud Hybrid / Multi-Cloud

Penggunaan PDN + cloud lain secara bersama agar sistem lebih fleksibel dan tahan gangguan.

CIO (Chief Information Officer)

Pimpinan teknologi; memastikan infrastruktur mendukung Data Governance.

Crowdsourcing Data Labeling

Pelibatan ASN/mahasiswa untuk memberi label pada data training AI.

D

Data Catalog

Daftar seluruh dataset pemerintah lengkap dengan metadata dan pemiliknya.

Data Deduplication

Menghapus data ganda (contoh: NIK ganda, balita ganda, siswa ganda).

Data Lake

Penyimpanan data mentah dan besar. Cocok untuk AI, CCTV, IoT.

Data Lakehouse

Gabungan Data Lake + Data Warehouse.

Data Mesh

Pembagian domain data di dalam K/L besar agar tidak menumpuk di pusat.

Data Owner

Pejabat yang bertanggung jawab penuh atas akurasi dataset.

Data Steward

Petugas harian yang membersihkan, memvalidasi, dan menjaga kualitas data.

Data Warehouse

Tempat data terstruktur, rapi, dan siap untuk laporan resmi.

Dewan Data Nasional

Pengarah kebijakan data seluruh K/L agar memakai standar yang sama.

DPO (Data Protection Officer)

Penanggung jawab kepatuhan terhadap UU PDP dan privasi masyarakat.

E**Early Warning System (EWS)**

Sistem peringatan dini berbasis data real-time (contoh: stunting, DBD, inflasi, dropout).

Encryption (Enkripsi)

Pengamanan data agar tidak bisa dibaca pihak yang tidak berwenang.

ETL/ELT (Extract–Transform–Load)

Proses mengambil, membersihkan, dan mengirim data dari daerah ke pusat.

ESB (Enterprise Service Bus)

Penghubung sistem lama (legacy) dengan sistem modern.

F**Federated Data Governance**

Data tetap berada di K/L, tetapi berbagi melalui standard API dan metadata nasional.

Forecasting

Prediksi berbasis data historis (kemiskinan, dropout, stunting, ekonomi daerah).

G

GovCloud

Cloud nasional untuk menyimpan dan mengelola data pemerintah secara aman.

GitLab CE

Repository open-source untuk dokumentasi dan pengelolaan domain data.

GPU (Graphics Processing Unit)

Diperlukan untuk AI dan deep learning pemerintah.

Garbage In = Garbage Out

Data buruk menghasilkan keputusan buruk.

H

Hadoop

Sistem penyimpanan big data berskala besar.

Harbor

Container registry open-source untuk Government Cloud.

I

IAM (Identity & Access Management)

Sistem untuk mengatur siapa boleh mengakses data apa.

Interoperabilitas

Kemampuan sistem pemerintah untuk saling bertukar data dengan standar sama.

Istio / Linkerd

Service mesh untuk microservices pemerintah.

J

JSON (JavaScript Object Notation)

Format standar pertukaran data nasional.

JSON Schema Nasional

Standar struktur data pemerintah yang dipakai kabupaten → provinsi → pusat.

K

Kafka

Jalan tol data real-time nasional.

Kamus Data

Definisi setiap kolom/variabel agar tidak membingungkan antar K/L.

Keycloak

SSO dan RBAC open-source untuk autentikasi antar kementerian.

Kubernetes

Orkestrator aplikasi modern microservices pemerintah.

L

Lakehouse

Sistem terpadu data mentah + data terstruktur.

Legacy System

Aplikasi lama di daerah yang harus diintegrasikan, bukan dibuang.

LLM (Large Language Model)

Model AI seperti Llama, Mistral, Phi yang dapat berjalan lokal via Ollama.

M

Machine Learning

Metode AI untuk klasifikasi, prediksi, dan rekomendasi kebijakan.

Message Bus

Sistem pengiriman data real-time (Kafka, RabbitMQ).

Metabase

Dashboard sederhana open-source untuk pemda.

Metadata

Informasi tentang dataset: sumber, tanggal, definisi, kualitas.

Microservices

Aplikasi kecil yang berjalan terpisah tetapi saling terhubung.

N

N8N

ETL open-source mudah dipakai oleh pemda.

NiFi

ETL kuat untuk pengiriman data berskala besar.

NPU (Neural Processing Unit)

Prosesor khusus AI untuk edge computing (misalnya di Puskesmas/desa).

O

Ollama

Platform LLM lokal untuk menjalankan AI pemerintah tanpa internet.

OpenAPI

Standar dokumentasi dan struktur API pemerintah.

OpenStack

Fondasi private cloud untuk PDN dan PDD.

OpenVINO

Optimasi AI untuk perangkat kecil dan edge.

P

PDN (Pusat Data Nasional)

Lokasi penyimpanan data strategis negara.

PDD (Pusat Data Daerah)

Node PDN di provinsi/kab/kota.

PKS (Perjanjian Kerja Sama) Data

Dokumen legal antar-instansi untuk berbagi data secara aman.

PostgreSQL

Database utama open-source untuk pemerintah.

Prometheus & Grafana

Monitoring dan visualisasi server pemerintah.

R

RBAC (Role-Based Access Control)

Pembatasan akses data berdasarkan peran pejabat/ASN.

Real-Time Data

Data yang mengalir langsung saat kejadian (kesehatan, harga pangan, bencana).

Roadmap Data Governance

Rencana 10 tahun untuk standarisasi → GovCloud → AI nasional → predictive government.

S

Schema Versioning

Versi skema data agar perubahan tidak merusak sistem daerah.

Single Source of Truth (SSOT)

Semua instansi memakai satu sumber data nasional.

SPBE (Sistem Pemerintahan Berbasis Elektronik)

Kerangka transformasi digital nasional.

Superset (Apache Superset)

Dashboard open-source untuk provinsi dan kementerian.

Suricata

IDS/IPS open-source untuk keamanan jaringan pemerintah.

T

Talend

ETL open-source untuk integrasi aplikasi lama.

TensorFlow Lite

AI ringan untuk perangkat kecil (edge AI).

TimescaleDB

Ekstensi PostgreSQL untuk data time-series.

Trust Framework Nasional

Struktur legal–teknis–keamanan untuk memastikan data sharing aman.

V

Value for Money (VfM)

Analisis efektivitas anggaran berdasarkan output dan dampak.

Version Control

Pengaturan perubahan skema, API, dan kode secara terkontrol (GitLab CE).

W

Wali Data

Pihak yang mengkoordinasikan data antar unit sesuai Perpres SDI.

Wazuh

Platform keamanan open-source untuk monitoring PDN/PDD.

WireGuard

VPN modern untuk pertukaran data terenkripsi antar K/L/D.

Referensi

Bannister, F., & Connolly, R. (2014). The trouble with transparency: A critical review of openness in e-government. *Policy & Internet*, 3*(1), 1–30.
<https://doi.org/10.2202/1944-2866.1076>

DAMA International. (2017). *DAMA-DMBOK: Data Management Body of Knowledge* (2nd ed.). Technics Publications.

Davenport, T. H., & Harris, J. G. (2017). *Competing on analytics: Updated, with a new introduction: The new science of winning*. Harvard Business Review Press.

European Commission. (2020). *European strategy for data*.
https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

Gartner. (2024). *Gartner glossary: Data governance*.
<https://www.gartner.com/en/information-technology/glossary/data-governance>

Government of Indonesia. (2019). Peraturan Presiden Republik Indonesia Nomor 39 Tahun 2019 tentang Satu Data Indonesia.

Government of Indonesia. (2020). Peraturan Presiden Republik Indonesia Nomor 95 Tahun 2018 tentang Sistem Pemerintahan Berbasis Elektronik (SPBE).

Government of Indonesia. (2022). Undang-Undang Republik Indonesia Nomor 27 Tahun 2022 tentang Perlindungan Data Pribadi.

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, 29*(4), 258–268.
<https://doi.org/10.1080/10580530.2012.716740>

Klievink, B., Bharosa, N., & Tan, Y.-H. (2016). The collaborative realization of public values and business goals: Governance and infrastructure of public–private information platforms. *Government Information Quarterly*, 33*(1), 67–79. <https://doi.org/10.1016/j.giq.2015.12.002>

Kubler, S., Robert, J., Derigent, W., Voisin, A., & Le Traon, Y. (2017). Open data and big data platform for smart cities: The European approach. *Procedia Computer Science*, 98*, 289–296.

Loshin, D. (2019). *The practitioner's guide to data quality improvement*. Morgan Kaufmann.

OECD. (2019). *The path to becoming a data-driven public sector*. OECD Publishing.
<https://doi.org/10.1787/059814a7-en>

Open Data Institute. (2023). *Data institutions: Building blocks for trusted data ecosystems*.
<https://theodi.org/>

Purbo, O. W. (2025). *Data infrastructure & data governance untuk Republik Indonesia: Fondasi negara cerdas 2045*. OnnoCenter & Institut Teknologi Tangerang Selatan.

Ubaldi, B. (2013). Open government data: Towards empirical analysis of open government data initiatives. *OECD Working Papers on Public Governance*, No. 22. OECD Publishing. <https://doi.org/10.1787/5k46bj4f03s7-en>

United Nations. (2022). *United Nations e-government survey 2022: The future of digital government*. Department of Economic and Social Affairs.

Welle Donker, F., & van Loenen, B. (2017). How to assess the success of national spatial data infrastructures? *Computers, Environment and Urban Systems*, 61*, 39–51. <https://doi.org/10.1016/j.compenvurbsys.2016.08.003>

World Bank. (2021). *World development report 2021: Data for better lives*. World Bank. <https://doi.org/10.1596/978-1-4648-1600-0>

Yang, T.-M., Lo, J., & Shiang, J. (2015). To open or not to open? Determinants of open government data. *Journal of Information Technology & Politics*, 12*(4), 372–391. <https://doi.org/10.1080/19331681.2015.1086426>

Tentang Penulis



Onno W. Purbo, saat ini bertugas sebagai rektor di Institut Teknologi Tangerang Selatan (ITTS). Onno memperoleh gelar Ph.D bidang Electrical Engineering dari University of Waterloo, Canada, adalah seorang copyleft, educator dan ICT evangelist. Dia sudah mempublikasikan 50++ buku, termasuk free ICT ebook untuk sekolah tahun 2008. Beberapa buku terakhirnya adalah "Internet-TCP/IP: Konsep Dan Implementasi", 2018; "Sistem Operasi, Konsep Dan Membuat Linux OpenWRT Dan ROM Android", 2019; "IPv6 Untuk Mendukung Operasi Jaringan Dan Domain Name System", 2019; "Kubernetes untuk Pemula", 2024 dan "Membuat Operator Seluler 5G Sendiri", 2024. Dia memimpin sambungan pertama Internet di Institut Teknologi Bandung, tahun 1993-2000, dan menggunakannya untuk membuat jaringan Internet

pendidikan yang pertama di Indonesia. Dia membebaskan frekuensi WiFi, memperkenalkan RT/RW-net, antenna Wajanbolic dan jaringan selular OpenBTS dan private 5G sendiri. Dia memimpin jaringan telepon pertama di atas Internet, VoIP Merdeka, yang kemudian hari dikenal sebagai VoIP Rakyat berbasis SIP dan menggunakan kode area +62520 dan +62521. Dia saat ini aktif memperkenalkan e-Learning, dan menjalankan server e-Learning di <http://lms.onnocenter.or.id/moodle/> 66,000++ siswa / mahasiswa dan <https://opencourse.itts.ac.id> 38.000+ mahasiswa secara gratis. Email: onno@indo.net.id