

Data Improvement Life Cycle untuk Meningkatkan Kualitas Data: Studi Kasus Data Survey Kesehatan Mental

¹Agus Yulianto, ²Firmansyah,
^{1,2}Universitas Bina Sarana Informatika
Jakarta, Indonesia

¹agus.aag@bsi.ac.id, ²firmansyah.fmh@bsi.ac.id

*Penulis Korespondensi

Diajukan : 14/03/2025

Diterima : 27/03/2025

Dipublikasi : 03/04/2025

ABSTRAK

Meningkatkan efisiensi dalam pengambilan keputusan bisnis melalui optimalisasi kinerja data warehouse dengan implementasi data mart juga dapat membuat akurasi data yang digunakan untuk analisis bisnis dengan mengimplementasikan mekanisme pembersihan dan validasi data yang efektif pada data mart, kemudian menyederhanakan proses analisis data dengan menyediakan data yang relevan dan terstruktur dalam data mart dan menurunkan biaya operasional terkait dengan pengelolaan data warehouse melalui optimalisasi penggunaan sumber daya, Model star schema dan model snowflake dalam arsitektur data warehouse merupakan model yang tidak terpisahkan ketika membangun sebuah data warehouse. Performa masing-masing model perlu diukur untuk melihat model yang performanya lebih baik dan dapat diadopsi dalam pengembangan data warehouse. Dalam penelitian ini juga ditambahkan data mart untuk dibandingkan dengan star schema dan snowflake dikarenakan data mart merupakan konsep yang dapat diadopsi dalam data warehouse. Tiga model yaitu star schema, snowflake dan data mart diuji performanya dengan tiga query yang digunakan yaitu query performance, query complexity dan business intelligence performance, pengukuran metrik kinerja dilakukan dengan query performance, seberapa cepat query dapat dijalankan dan menghasilkan hasil. Query complexity melihat seberapa kompleks dapat dijalankan dengan efisien dan business intelligence performanse untuk mendukung analisis bisnis dan pengambilan keputusan, namun, peningkatan kompleksitas query pada snowflake schema dapat mempengaruhi kinerja secara keseluruhan. Penelitian ini menyimpulkan bahwa pemilihan model data mart yang tepat sangat penting untuk mencapai optimalisasi kinerja data warehouse.

Kata Kunci: *Business Intelligence, Data, Database, Data Warehouse, Data Mart, Dimensional Modeling, Snowflake, Etl, Star Schema*

I. PENDAHULUAN

Kualitas data mengacu pada tingkat akurasi, kelengkapan, konsistensi, dan keandalan data dalam suatu kumpulan data. Agar data dianggap berkualitas tinggi, data tersebut harus bebas dari kesalahan, duplikasi, dan inkonsistensi. Data berkualitas tinggi sangat penting untuk pengambilan keputusan yang terinformasi, analisis yang akurat, dan keberhasilan inisiatif berbasis data. Sebagai contoh, sebuah perusahaan ritel yang mengumpulkan data pelanggan seperti riwayat pembelian, informasi kontak, dan preferensi, dapat mengalami kesulitan menargetkan kampanye pemasaran yang dipersonalisasi jika datanya tidak konsisten atau tidak

lengkap. Dengan menerapkan pemeriksaan kualitas data dan secara rutin memperbarui serta membersihkan database mereka, perusahaan dapat memastikan bahwa data mereka akurat dan andal untuk mendukung keputusan bisnis strategis.

Kualitas data memainkan peran penting dalam pengambilan keputusan, karena akurasi dan keandalan data secara langsung memengaruhi hasil dari keputusan yang dibuat. Ketika data berkualitas tinggi, pembuat keputusan dapat merasa yakin terhadap informasi yang mereka gunakan sebagai dasar keputusan. Hal ini mengarah pada proses pengambilan keputusan yang lebih efektif dan sukses. Sebaliknya, data berkualitas rendah dapat menyebabkan keputusan yang tidak efisien, berdampak pada tingkat kepercayaan terhadap data hingga berdampak pada kepuasan customer (Prasad, 2024). Rendahnya kualitas data dipengaruhi beberapa hal seperti data yang tersebar, kompleksitas integrasi data, prosedur dan kebijakan data, keamanan data dan kesalahan manusia (Pansara et al., 2023).

Kerangka kerja Data Management Body of Knowledge memberikan panduan komprehensif bagi organisasi yang ingin meningkatkan kualitas data mereka. Dengan mengikuti prinsip dan praktik terbaik yang diuraikan dalam kerangka ini, organisasi dapat membangun proses tata kelola data yang kuat, manajemen kualitas data yang terorganisir (Ismail et al., 2024), serta tanggung jawab pengelolaan data (data stewardship). Implementasi membantu mengelola manajemen data yang standar, meningkatkan kualitas data dan memastikan keamanan serta kerahasiaan data. Penerapan juga memungkinkan organisasi untuk meningkatkan validitas data, meningkatkan akurasi dan konsistensi data (Rahmawati et al., 2023).

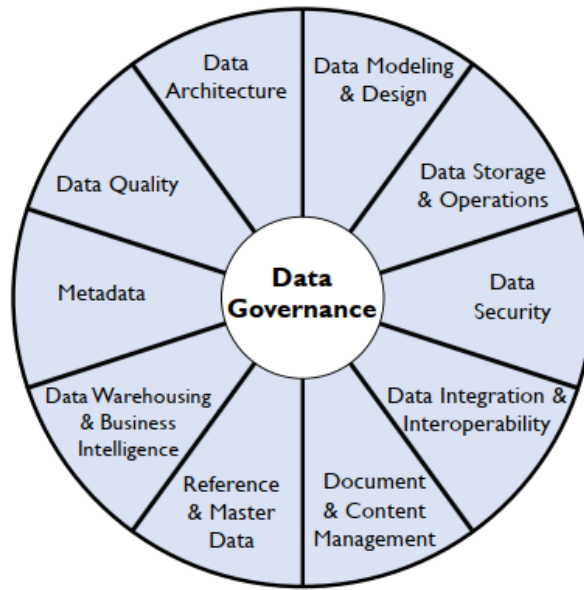
II. STUDI LITERATUR

1. Penelitian Terkait

Penelitian kualitas data dalam bidang kesehatan seperti yang ditulis oleh Bernardi bahwa kualitas data dipengaruhi oleh beberapa hal seperti kebijakan yang standar, penerapan kerangka kerja manajemen data bahkan untuk memastikan kualitas data perlu dilakukan pengawasan (Bernardi et al., 2023). Kualitas data yang baik juga dipengaruhi factor lain seperti kelengkapan data, ketepatan input data bahkan tingkat pengetahuan dari tenaga administrasi di fasilitas kesehatan (Solomon et al., 2021). Dalam penelitian lain, penerapan kerangka kerja kebijakan data dapat diterapkan pada organisasi kesehatan public sehingga dapat menjamin kualitas data (Paramita, 2023). Dalam penelitian lain ditemukan bahwa faktor-faktor manusia juga menjadi penentu kualitas data seperti lama bekerja dan juga tingkat pendidikan, namun ada juga factor lain di dalam organisasi seperti seberapa sering pertemuan rutin untuk membahas data (Morris et al., 2024). Masih dalam penelitian yang lain dalam bidang kesehatan, peningkatan kualitas data yang berkelanjutan untuk mengelola data administrasi kesehatan diteliti menggunakan kerangka kerja konseptual berdasarkan 4 kategori yaitu karakteristik data cacat, isu proses kerja dan manusia, tantangan untuk perbaikan kualitas data berkelanjutan dan peluang untuk perbaikan kualitas data berkelanjutan (Zhang et al., 2024).

2. DAMA-DMBOK

DAMA-DMBOK merupakan kependekan dari Data Management-Data Management Body of Knowledge yang dikembangkan oleh DAMA International yang dibuat sebagai kerangka kerja manajemen data. Manajemen data adalah pengembangan, eksekusi dan supervisi terhadap semua perencanaan, kebijakan, program dan berbagai praktik yang menyajikan, mengendalikan, melindungi serta meningkatkan nilai dari data dan informasi selama siklus hidupnya (DAMA, 2024). Data dan informasi secara konsep memiliki makna yang berbeda, dimana data merupakan sekumpulan fakta dalam bentuk angka, teks ataupun bentuk yang lain dan bersifat mentah sedangkan informasi merupakan hasil ekstraksi dari data yang sudah memiliki makna dan dapat dijadikan sebagai pendukung keputusan. Kerangka kerja DAMA-DMBOK memiliki 10 area pengetahuan yang digambarkan dalam bentuk roda bernama DAMA Wheel dibawah ini :



Gambar 1 DAMA-DMBOK Framework

Untuk mencapai tujuan dari penerapan manajemen data berbasis DAMA-DMBOK, maka diperlukan faktor yang dapat menopang tujuan tersebut. Ada 3 faktor yang mendukung terciptanya tujuan implementasi DAMA di dalam organisasi yaitu manusia, proses dan teknologi yang digambarkan dalam hexagonal dibawah ini :



Gambar 2 DAMA Environmental Factor Hexagon

Dari 3 faktor di atas, dibagi lagi menjadi faktor yg lebih kecil seperti faktor manusia itu ditentukan oleh peran, tanggung jawab dan kultur. Faktor proses ditentukan oleh aktivitas dan tehnik yang dipakai kemudian faktor teknologi ditentukan oleh alat yang digunakan dan hasil akhir atau keluaran yang dihasilkan dari teknologi.

3. Data Quality

Kualitas data memiliki banyak dimensi yang mempengaruhinya seperti akurasi, kelengkapan, konsistensi, ketepatan waktu dan relevansi. Artinya jika semua dimensi memenuhi maka dapat dikatakan data berkualitas (Wang et al., 2024). Akurasi suatu data dapat dicapai salah satunya dengan perencanaan yang tepat dalam pembuatan data untuk memastikan bahwa data yang direkam dan dikumpulkan mencerminkan nilai yang sebenarnya (Hassenstein & Vanella, 2022). Dalam penelitian disebutkan bahwa pengguna data mengatakan data yang akurat itu penting untuk membuat keputusan (Khong et al., 2023). Sebaliknya kualitas data yang buruk dapat membuat ketidapuasan customer, meningkatnya biaya operasional, rendahnya kepuasan karyawan bahkan menurunnya kultur organisasi dan tingkat kepercayaan (Prasad, 2024).

Kualitas data berhubungan dengan sejauh mana data dapat memenuhi spesifikasi yang ditetapkan oleh organisasi dengan karakteristik seperti kebenaran, kelengkapan dan konsistensi (Miller et al., 2024). Misalnya kita menyebut bahwa konsistensi merupakan dimensi yang paling utama untuk menentukan bahwa data berkualitas, namun mungkin pada konsumen data yang lain bisa saja menyebut bahwa konsistensi bukan hal utama yang menentukan kualitas data. Hal ini terjadi dikarenakan setiap pengguna data punya perspektif yang berbeda untuk melihat dimensi dari kualitas data. Pada beberapa penelitian yang terkait dengan kualitas data kesehatan, dimensi kualitas data seperti kelengkapan, konsistensi kebenaran, keunikan, stabilitas, ketepatan waktu, konteks data, kepercayaan dan keterwakilan merupakan dimensi yang mempengaruhi kualitas data (Declerck et al., 2024). Dimensi kualitas data terdapat beberapa perbedaan seperti yang didefinisikan oleh DAMA Netherland ada 12 dimensi kualitas data yang sangat penting yaitu accuracy, availibility, clarity, completeness, consistency, punctuality, timeliness, traceability, uniqueness dan validity (Askham ; Nicola et al., 2013). Dimensi kualitas data yang didefinisikan oleh DAMA-UK berbeda yaitu ada 6 dimensi inti kualitas data yaitu completeness, uniqueness, timeliness, validity, accuracy dan consistency.



Gambar 3 Data Quality Dimensions

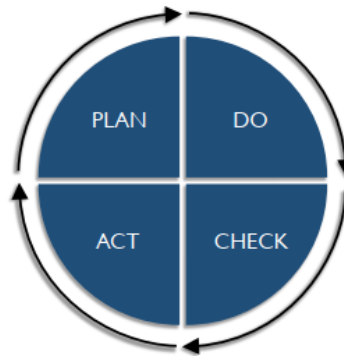
Completeness berarti kelengkapan data yang diinput, Uniqueness berarti tidak ada data yang duplikat atau bermakna ganda. Validity merupakan data yang dikonfirmasi valid sesuai dengan yang didefinisikan misalnya satuan barang, kode pos atau data lain yang dapat dikonfirmasi validitasnya. Timeliness adalah kondisi dimana kondisi data diinput sesuai pada waktunya, tidak mundur waktu dan juga tidak maju namun tepat pada waktunya. Accuracy adalah tingkat kebenaran suatu data untuk mendekati kondisi sebenarnya sehingga semakin akurat suatu data

maka semakin sesuai dengan kondisi sebenarnya. Consistency berarti jika data dibandingkan dengan data yang lain maka data akan tetap sama, tidak ada perbedaan pada tiap data sehingga konsisten.

III. METODE

A. Data Improvement Life Cycle

Untuk dapat meningkatkan Peningkatan kualitas data Data quality improvement merujuk kepada disiplin berkesinambungan untuk fokus meningkatkan akurasi, kelengkapan dan reliabilitas data. Untuk meningkatkan kualitas data yang berkelanjutan, maka dibutuhkan siklus di dalam proses peningkatan kualitas data seperti yang dijabarkan oleh Shewart dan Deming menggunakan konsep Plan-Do-Check-Action (PDCA).



Gambar 4 Plan-Do-Check-Action

Pada dasarnya, siklus PDCA dapat diterapkan pada proses non data, namun dalam penelitian ini untuk peningkatan kualitas data yang berkelanjutan maka akan menggunakan siklus PDCA ini.

1. Plan

Dalam tahap ini ada beberapa hal yang harus dipersiapkan sebelum melakukan penilaian terhadap kualitas data yaitu menentukan kerangka kerja dan alat yang digunakan. Untuk kerangka kerja kita menggunakan 6 dimensi kualitas data yang sudah dijelaskan di atas yaitu completeness, uniqueness, timeliness, validity accuracy dan consistency. Alat yang digunakan untuk mengukur kualitas data adalah Data Quality Metric (DQ Metric). DQ Metric merupakan tabel metrik dimana tabel ini mengukur kualitas data berdasarkan 2 variabel yaitu aturan bisnis dan dimensi kualitas data yang kemudian menghasilkan indikator kualitas data.

2. Do

Pada tahap ini, proses dibagi menjadi 3 tahap, yaitu tahap penilaian kualitas data, analisa akar masalah dan tindakan korektif-preventif. Tahap ke-1, berdasarkan DQ Metric data dianalisis kemudian hasilnya dimasukkan ke dalam template DQ Metric pada setiap poin. Pada tahap ke-2, jika pada setiap point ditemukan indikator yang tidak mencapai standar, maka harus ditemukan akar masalahnya agar pada proses berikutnya tidak terjadi penurunan kualitas data. Setiap akar masalah terdiri dari dua aspek yaitu teknis dan non teknis. Aspek non teknis berkaitan dengan proses kerja yang terjadi data saat mulai input data hingga data selesai diinput, sedangkan aspek teknis terkait kepada perangkat lunak atau teknologi. Tahap ke-3, jika data masih dapat diperbaiki maka segera diperbaiki misalnya data kode pos pada customer banyak kosong atau kota pada customer banyak kosong. Berikutnya setelah korektif, maka tindakan preventif dituliskan dalam template DQ Metric untuk meningkatkan kualitas data.

3. Check

Setelah tindakan korektif-preventif dilakukan, maka data harus dimonitor secara rutin oleh tim tertentu untuk mengevaluasi kualitas data pasca proses penilaian kualitas data. Apabila pada tahap ini masih ditemukan data yang kualitasnya belum sesuai harapan maka harus dilanjutkan ke tahap Action. Proses monitor kualitas data dapat menggunakan business intelligence (BI) seperti PowerBI, Clicksense, Tableau atau alat BI yang lain.

4. Act

Jika pada tahap Check masih ditemukan isu kualitas data atau ada proses yang perlu ditingkatkan agar kualitas data meningkat maka perlu mulai dari tahap awal lagi yaitu Plan.

B. Data Quality Metric

Data Quality Metric menggunakan tabel yang di dalamnya terdapat 2 variabel untuk mengukur kualitas data yaitu dimensi kualitas data dan aturan proses bisnis. Adapun contoh dari template DQ Metric seperti di bawah ini yang bersumber dari DAMA-DMBOK :

Dimension and Business Rule	Measure	Metrics	Status Indicator
Completeness Business Rule 1: Population of field is mandatory	Count the number of records where data is populated, compare to the total number of records	Divide the obtained number of records where data is populated by the total number of records in the table or database and multiply it by 100 to get to percentage complete	Unacceptable: Below 80% populated Above 20% not populated
Example 1: Postal Code must be populated in the address table	Count populated: 700,000 Count not populated: 300,000 Total count: 1,000,000	Positive measure: $700,000/1,000,000 * 100 = 70\%$ populated Negative measure: $300,000/1,000,000 * 100 = 30\%$ not populated	Example result: Unacceptable
Uniqueness Business Rule 2: There should be only one record per entity instance in a table	Count the number of duplicate records identified; report on the percentage of records that represent duplicates	Divide the number of duplicate records by the total number of records in the table or database and multiply it by 100	Unacceptable: Above 0%
Example 2: There should be one and only one current row per postal code on the Postal Codes master list	Count of duplicates: 1,000 Total Count: 1,000,000	$10,000/1,000,000 * 100 = 1.0\%$ of postal codes are present on more than one current row	Example result: Unacceptable
Timeliness Business Rule 3: Records must arrive within a scheduled timeframe	Count the number of records failing to arrive on time from a data service for business transactions to be completed	Divide the number of incomplete transactions by the total number of attempted transactions in a time period and multiply by 100	Unacceptable: Below 99% completed on time Above 1% not completed on

Gambar 5 Data Quality (DQ) Metric

Template di atas merupakan template sederhana yang kemudian kami kembangkan menjadi template di bawah ini yang disebut sebagai dokumen Data Quality Assessment :

Table 1 DQ Metric Template

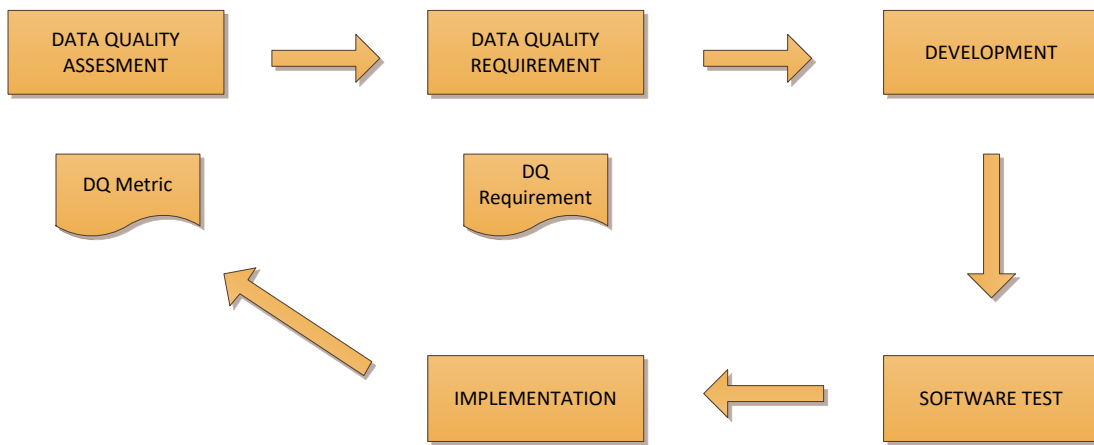
Dimension	Business Rule	Measure	Metric	Standard	Status
1. Completeness	1.1. Nama kota harus terisi	Kota yang diisi : 700 Kota yang tidak diisi : 300	$300/700 * 100 = 58\%$	100%	Unacceptable
.....

--	--	--	--	--	--

Kolom Dimension adalah untuk diisi dengan 6 dimensi kualitas data sedangkan kolom Business Rule merupakan aturan bisnis atau standar yang ditetapkan di dalam data misalnya nama kota harus terisi lengkap atau kode pos harus terisi lengkap. Kolom Measure untuk menampilkan data yang sesuai dengan business rule dan data yang tidak sesuai dengan business rule kemudian angkanya dimasukkan ke dalam kolom Metric dan dijumlahkan sesuai dengan kebutuhan seperti persentase, rasio atau lainnya. Hasil pada kolom Metric ditampilkan dalam bentuk label yang dimasukkan pada kolom Status misalnya acceptable atau unacceptable.

C. Data Quality Development Life Cycle

Siklus yang digunakan dalam penelitian ini menggunakan siklus pengembangan system berbasis kualitas data seperti di bawah ini :



Gambar 6 DQ Development Life Cycle

IV. HASIL DAN PEMBAHASAN

Data bersumber dari Open Sourcing Mental Health dimana data tersebut berisi survey kesehatan mental yang dilaksanakan oleh OSMI. Dari keseluruhan kolom yang ada, kami mengambil kolom dengan random sampling yang dimasukkan ke dalam Business Rule seperti dokumen Data Quality Assesment di bawah ini :

Table 2 Data Quality Assesment

Dimension	Business Rule	Measure	Metric	Standard	Status
1. Completeness	1.1. Jumlah karyawan harus terisi	Jumlah karyawan yang terisi : 361 Jumlah karyawan yang tidak terisi : 56	84.48%	100%	Unacceptable
	1.2. Pekerjaan tech company harus terisi	Jumlah karyawan yang terisi : 361 Jumlah karyawan yang tidak terisi : 56	84.48%	100%	Acceptable
2. Uniqueness	2.1. ID tidak ada	Jumlah	100%	100%	Acceptable

	yg duplikat	karyawan yang duplikat : 0 Jumlah karyawan yang tidak duplikat : 417			
3. Timeliness	3.1. Kolom submit date terisi dengan tanggal dan jam	Jumlah kolom yang terisi : 417 Jumlah karyawan yang tidak terisi : 0	100%	100%	Acceptable
4. Validity	4.1. Kolom “improve mental health support for employees” terisi dengan kalimat.	Jumlah kolom yang terisi valid : 409 Jumlah karyawan yang tidak valid : 8	98.04%	100%	Unacceptable
5. Accuracy	5.1. Kolom “Family history mental illness” terisi yes atau no	Jumlah yang terisi “I don’t know” : 95 Jumlah yang terisi yes/no : 322	70.49%	90%	Unacceptable
6. Consistency	6.1. Kolom status pekerjaan teknologi (yes/no)	Jumlah terisi yes/no: 0 Jumlah terisi 1/0 : 417	0%	100%	Unacceptable

Dari hasil penilaian menggunakan data quality metric yang memiliki status unacceptable dimasukkan ke dalam tabel data quality requirement :

Table 3 Data Quality Requirement

No.	Bus. Rule No.	Business Rule	Requirement
1.	1.1.	Jumlah karyawan harus terisi lengkap	Field jumlah karyawan wajib isi
2.	4.1.	Kolom “improve mental health support for employees” terisi dengan kalimat	Field wajib isi, tidak boleh diisi dengan numeric.
3.	5.1.	Kolom “Family history mental illness” terisi yes atau no	Field wajib isi dan ada pilihan yes/no
4.	6.1.	Kolom status pekerjaan teknologi terisi yes atau no	Field wajib isi dan ada pilihan yes/no

Data quality requirement merupakan dokumen yang berisi business rule dan requirement rule dimana pada kolom requirement merupakan aturan atau kebutuhan yang akan diterapkan ke dalam pengembangan perangkat lunak agar pada proses berikutnya tidak terjadi kesalahan input dan output sehingga dapat meningkatkan kualitas data. Setiap metod pengembangan perangkat lunak mungkin berbeda, jika pada waterfall dimasukkan ke dalam requirement sedangkan pada agile dimasukkan sebagai product backlog.

V. KESIMPULAN

Untuk mendapatkan data yang berkualitas ada 2 tahap yang perlu dilakukan yaitu mengecek kualitas data dengan penilaian dan tahap berikutnya adalah dengan pengembangan aplikasi berbasis kualitas data. Untuk melakukan evaluasi terhadap data dapat menggunakan data quality metric untuk melakukan pengukuran dengan menggabungkan dimension dan business rule. Setelah evaluasi kualitas data, tahap berikutnya adalah peningkatan kualitas data dengan data quality development life cycle. Implementasi ini mesti dilakukan secara berkesinambungan yang dimulai dari tahap Plan Do Check dan Action agar kualitas data dapat terus ditingkatkan.

VI. REFERENSI

- Bernardi, F. A., Alves, D., Crepaldi, N., Yamada, D. B., Lima, V. C., & Rijo, R. (2023). Data Quality in Health Research: Integrative Literature Review. *Journal of Medical Internet Research*, 25, 1–20. <https://doi.org/10.2196/41446>
- DAMA. (2024). DAMA-DMBOK2 Framework. *The Data Mangement Association*, 1–27. <https://dama.org/sites/default/files/download/DAMA-DMBOK2-Framework-V2-20140317-FINAL.pdf>
- Declerck, J., Kalra, D., Stichele, R. Vander, & Coorevits, P. (2024). Frameworks, Dimensions, Definitions of Aspects, and Assessment Methods for the Appraisal of Quality of Health Data for Secondary Use: Comprehensive Overview of Reviews. *JMIR Medical Informatics*, 12. <https://doi.org/10.2196/51560>
- Hassenstein, M. J., & Vanella, P. (2022). Data Quality—Concepts and Problems. *Encyclopedia*, 2(1), 498–510. <https://doi.org/10.3390/encyclopedia2010032>
- Ismail, A., Suroso, A. I., & Hermadi, I. (2024). Data Governance Design with the DAMA-DMBOK Framework (Case Study: PT. XYZ). *International Journal of Research and Review*, 11(8), 210–221. <https://doi.org/10.52403/ijrr.20240823>
- Khong, I., Aprilia Yusuf, N., Nuriman, A., & Bayu Yadila, A. (2023). Exploring the Impact of Data Quality on Decision-Making Processes in Information Intensive Organizations. *APTISI Transactions on Management (ATM)*, 7(3), 253–260. <https://doi.org/10.33050/atm.v7i3.2138>
- Miller, R., Whelan, H., Chrubasik, M., Whittaker, D., Duncan, P., & Gregório, J. (2024). A Framework for Current and New Data Quality Dimensions: An Overview. *Data*, 9(12), 1–26. <https://doi.org/10.3390/data9120151>
- Morris, L. D., Nyongesa, M. W., & Sokiri, T. D. (2024). Factors Influencing Data Quality in Routine Health Information Systems in Maridi County , South Sudan. *South African Journal of Information Management*, 1–13.
- Pansara, R., Bhargav, S., & Adusumilli, K. (2023). Cultivating Data Quality to Strategies, Challenges, and Impact on Decision-Making. *International Journal of Management Education for Sustainable Development*, 6.
- Paramita, P. (2023). Public Health Information Standard Data Quality and Governance. *Journal of World Science*, 2(6), 817–824. <https://doi.org/10.58344/jws.v2i6.313>
- Prasad, A. (2024). Impact of Poor Data Quality on Business Performance: Challenges, Costs, and Solutions. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4843991>

-
- Rahmawati, R., Ruldeviyani, Y., Abdullah, P. P., & Ma'ruf Hudoarma, F. (2023). Strategies to Improve Data Quality Management Using Total Data Quality Management (TDQM) and Data Management Body of Knowledge (DMBOK): A Case Study of M-Passport Application. *CommIT Journal*, 17(1), 27–42. <https://doi.org/10.21512/COMMIT.V17I1.8330>
- Solomon, M., Addise, M., Tassew, B., Balcha, B., & Abebe, A. (2021). Data Quality Assessment and Associated Factors in the Health Management Information System Among Health Centers of Southern Ethiopia. *PLoS ONE*, 16(10 October), 1–13. <https://doi.org/10.1371/journal.pone.0255949>
- Wang, J., Liu, Y., Li, P., Lin, Z., Sindakis, S., & Aggarwal, S. (2024). Overview of Data Quality: Examining the Dimensions, Antecedents, and Impacts of Data Quality. *Journal of the Knowledge Economy*, 15(1), 1159–1178. <https://doi.org/10.1007/s13132-022-01096-6>
- Zhang, Y., Callaghan-Koru, J. A., & Koru, G. (2024). The challenges and opportunities of continuous data quality improvement for healthcare administration data. *JAMIA Open*, 7(3), 8–10. <https://doi.org/10.1093/jamiaopen/ooae058>