

BAB II

LANDASAN TEORI

2.1 Penelitian Terdahulu

Terdapat beberapa penelitian terdahulu yang mendasari atau dapat dijadikan sebagai acuan dalam pengerjaan tugas akhir ini. Penelitian tersebut terdiri dari 10 judul yang dipublikasikan diantara tahun 2020 dan 2025 yang dianggap relevan dan dapat menjadi acuan pembelajaran penyusunan penelitian ini. Tabel 2.1 di bawah ini akan menjelaskan hasil yang diperoleh pada setiap penelitian-penelitian tersebut.

Tabel 2. 1 Penelitian Terdahulu

No	Sumber	Metode	Hasil Penelitian
1.	[11] Susiana S, Agata F, Lussy E, Gloria V. (2024). <i>Pengembangan Data Mart Keuangan untuk Organisasi Nirlaba</i> . Jurnal Teknik Informatika dan Sistem Informasi, 10(2), Agustus 2024.	Kimball – <i>Bottom Up Approach Four Step Methodology</i>	Penelitian ini menghasilkan independent data mart finansial dengan skema starflake untuk analisis dan penyimpanan data historis. Perancangan data mart mengikuti metode Kimball <i>Four step</i> dan mampu mendukung visualisasi data melalui dashboard dan tabel pivot.
2.	[12] Irfan Zaelani (2021). <i>Implementasi Data Mart Terhadap Sistem Penjualan pada Perusahaan Bidang Distributor di PT. Eugen Trimathema</i> . Jurnal Penelitian Mahasiswa Teknik dan Ilmu Komputer, 1(2), November 2021.	Kimball – <i>Bottom Up Approach Four step Methodology</i>	Penelitian ini menghasilkan data mart menggunakan metode Kimball <i>Bottom-Up Approach Four Step Methodology</i> guna mendukung analisis data dan pengambilan keputusan bisnis di PT. Eugen Trimathema. Hasil implementasi menunjukkan bahwa data mart yang dibangun memudahkan pengolahan data dan membantu perusahaan dalam mengambil keputusan yang lebih efektif.

No	Sumber	Metode	Hasil Penelitian
3.	[13] Hendra M, Bambang P. D. P, Rikie K, Muhammad G, Robertus S (2023). <i>Analysis and Design of Data Warehouse and Data Mart Budget</i> . Journal of Intelligent Software Systems, 2(1), 2023.	<i>Kimball – Bottom Up approach Nine Step Methodology</i>	Penelitian ini menghasilkan desain data warehouse dan data mart anggaran menggunakan metode Kimball dengan <i>nine-step methodology</i> . Hasil penelitian menunjukkan bahwa data mart yang dirancang dapat menyimpan dan mengelola data anggaran secara lebih terstruktur untuk mendukung analisis keuangan universitas.
4.	[14] Tasya M. S, Angeline C, Andrew H. M, Dedi T, Jap Tji B (2024). <i>Perancangan Data Mart untuk Manajemen Data Penjualan pada Kedai Kopi X di Jakarta</i> . Journal of Information Technology and Computer Science, 7(6), 2024.	<i>Kimball – Bottom Up Approach Nine Step Methodology</i>	Penelitian ini menghasilkan data mart penjualan dengan star schema menggunakan metode <i>Nine-step Kimball</i> untuk mendukung analisis data dan pengambilan keputusan di kedai kopi X. Hasil implementasi menunjukkan bahwa data mart yang dibangun dapat mengoptimalkan pemrosesan data penjualan serta meningkatkan efisiensi dalam manajemen data.
5.	[15] □ Putu L. B. S. J. Amertha, Rukmi S. H, Made S (2020). <i>Data Warehouse Design for the Bank X with Inmon Approach</i> . International Journal of Engineering and Emerging Technology, 5(2), 2020.	<i>Bill Inmon Methodology</i>	Penelitian ini membangun desain database Data Warehouse untuk Bank X dengan menggunakan pendekatan Inmon. Proses perancangan dilakukan melalui tahapan seperti penentuan subjek utama, definisi subjek, pembuatan relasi antar objek, penentuan dimensi, dan pembuatan data mart independen. Hasil dari penelitian ini berupa tiga tabel fakta (pinjaman, tabungan, dan portofolio) serta

No	Sumber	Metode	Hasil Penelitian
			dua data mart (pinjaman dan tabungan) yang dapat digunakan oleh masing-masing departemen untuk mendukung analisis dan pengambilan keputusan.
6.	[16] Alaa H, Marwah K. H, Zahraa A, Rabab H. S (2021). <i>Implementing Data-Driven Decision Support System Based on Independent Educational Data Mart</i> . International Journal of Electrical and Computer Engineering, 11(6), 2021.	Bottom Up Approach	Penelitian ini mengembangkan <i>data-driven decision support system (DSS)</i> berbasis independent data mart yang mengelola data nilai siswa di Al-Iskandaria Primary School. Menggunakan implementasi <i>OLAP</i> dan <i>Key Performance Indicator (KPI)</i> , hasil evaluasi menunjukkan bahwa DSS ini memenuhi aspek privasi, keamanan, serta kinerja, sehingga efektif dalam mendukung keputusan akademik jangka pendek maupun panjang.
7.	[17] Hasdi P, Benardo A (2023). <i>Penerapan Data Warehouse dan Dashboard Berbasis Kimball Nine-Step untuk Meningkatkan Kualitas Informasi dan Pengambilan Keputusan</i> . Jurnal Sistem Informasi, 15(1), 2023.	Kimball – Bottom Up Approach Nine Step Methodology	Penelitian ini menerapkan metode Kimball <i>Nine-Step</i> untuk membangun data warehouse dan dashboard untuk PT Debe Tour guna mendukung pengelolaan dan analisis data. Hasil penelitian menunjukkan bahwa sistem yang dikembangkan mampu meningkatkan kualitas informasi serta mendukung pengambilan keputusan di perusahaan.
8.	[18] Jerrico A. Mok, D. T (2024). <i>Sales Data Mart Design at PT.XYZ</i> . Journal of Information Technology and	Kimball – Bottom Up Approach Nine Step Methodology	Penelitian ini merancang data mart penjualan di PT. XYZ menggunakan metode Kimball <i>Nine-Step Methodology</i> untuk mendukung pengelolaan dan analisis data. Hasil implementasi menunjukkan bahwa data mart ini

No	Sumber	Metode	Hasil Penelitian
	Computer Science, 7(6), 2024.		meningkatkan efisiensi pengelolaan data penjualan serta mendukung pengambilan keputusan yang lebih cepat dan akurat.
9.	[19] Ihab N, Jasim D, Alaa H, Ali A, Wid A, Mohammed A, Aqeel H (2022). <i>OLAP Mining with Educational Data Mart to Predict Students' Performance</i> . International Journal of Computing and Informatics, 46(5), 2022.	Bottom Up Approach	Sebagai dasar untuk analisis akademik, telah dibangun <i>educational data mart</i> berbasis data dari Alexandria Private Elementary School menggunakan OLAP. Hasil penelitian menunjukkan bahwa metode <i>clustering with expectation-maximization</i> mencapai akurasi tertinggi, yaitu 96.76% dalam memprediksi kinerja siswa dan 96.12% dalam memprediksi nilai siswa dibandingkan dengan algoritma lainnya.
10.	[20] Alessandro C, Angel Q, Mercedes V, Juan J, Lidia S (2024). <i>Data Mart in Business Intelligence with Ralph Kimball for Commercial Sales</i> . In <i>Proceedings of the Computational Methods in Systems and Software</i> , 2024.	Kimball – Bottom Up Approach Four Step Methodology	Dengan menerapkan metode Ralph Kimball dan model <i>star schema</i> , telah dibangun data mart penjualan untuk mendukung analisis bisnis di organisasi. Hasil penelitian menunjukkan bahwa implementasi data mart ini, bersama dengan proses ETL dan visualisasi menggunakan Tableau, berkontribusi pada peningkatan penjualan lebih dari 30%, dengan beberapa kategori produk mengalami peningkatan hingga 45%.

Secara keseluruhan, penelitian-penelitian sebelumnya menunjukkan kesamaan dalam pengembangan *independent data mart*, data mart, maupun data warehouse, penggunaan proses ETL, serta penerapan konsep BI dalam mendukung pengelolaan

data akademik. Dalam konteks skripsi ini yang berfokus pada dukungan teknologi terhadap manajemen data pendidikan tinggi, artikel [16], [19], [48] juga mengangkat studi kasus di bidang pendidikan. Ketiganya menghadapi tantangan dalam mengelola data akademik yang tersebar di berbagai sumber dan belum terstruktur, sehingga solusi berbasis *data mart* dipilih untuk mengintegrasikan data, menyederhanakan proses analisis, serta meningkatkan efisiensi pengambilan keputusan akademik.

Kesamaan ini juga tercermin dalam penelitian ini, di mana perancangan *independent data mart* untuk perhitungan dan pelaporan *Expected Learning Outcomes* (ELO) bertujuan menyajikan data akademik yang lebih terorganisir dan siap digunakan untuk analisis lebih lanjut. Meskipun artikel [16] dan [19] menekankan pada pengembangan *Decision Support System* (DSS) dan penerapan *data mining*, serta artikel [48] lebih berfokus pada efisiensi pelaporan nilai siswa di tingkat sekolah, pendekatan teknis yang mereka gunakan seperti proses ETL dan pemusatan data melalui *data mart* menjadi referensi relevan dalam perancangan sistem pada skripsi ini.

Selain itu, perancangan data mart tidak terlepas dari pemilihan pendekatan yang tepat guna memastikan proses pembangunan berjalan secara efisien dan menghasilkan sistem yang akurat. Secara umum, terdapat empat pendekatan yang dapat digunakan dalam membangun arsitektur data warehouse maupun data mart, yaitu: (1) top-down, (2) bottom-up, (3) hybrid, dan (4) federated [23]. Sebagai contoh, beberapa penelitian yang tercantum pada Tabel 2.1 membahas implementasi pendekatan top-down melalui metodologi Bill Inmon serta pendekatan bottom-up melalui metodologi Kimball. Dalam konteks skripsi ini, pendekatan Kimball dinilai lebih relevan karena memungkinkan pengembangan data mart independen yang terfokus dan cepat dalam mendukung analisis *Expected Learning Outcomes* (ELO), tanpa memerlukan pembangunan sistem terpusat atau data warehouse terlebih dahulu sebagaimana pada pendekatan top-down Inmon. Pemilihan pendekatan ini diharapkan dapat meningkatkan efisiensi perancangan serta adaptabilitas sistem terhadap kebutuhan analisis akademik yang dinamis.

Pendekatan lainnya yang bisa diperhatikan adalah jenis *dimensional modeling* yang digunakan. Delapan dari sepuluh penelitian di atas menggunakan *star schema*, yang menunjukkan bahwa pendekatan ini memberikan hasil dan manfaat nyata dalam implementasi data mart. *Star schema* sendiri banyak digunakan karena kemudahan dalam implementasi, kecepatan dalam eksekusi kueri OLAP, serta struktur yang lebih sederhana dibandingkan dengan *snowflake schema* [24]. Di sisi lain, dalam perancangan ETL, mayoritas penelitian terdahulu menggunakan SSIS atau Pentaho sebagai *ETL tools*. Misalnya, jurnal [11] menggunakan Pentaho karena alat ini bersifat *open-source*, fleksibel, serta mendukung berbagai format sumber data, sehingga lebih ekonomis dan dapat dikustomisasi sesuai kebutuhan tanpa ketergantungan pada ekosistem Microsoft. Sementara itu, jurnal [16] menggunakan SSIS karena terintegrasi secara langsung dengan *Microsoft SQL Server*, memiliki performa yang optimal dalam lingkungan berbasis SQL Server, serta menyediakan fitur *drag-and-drop* yang memudahkan pengembangan proses ETL tanpa perlu banyak *scripting*. Jika dikaitkan dengan penelitian ini, SSIS lebih relevan karena data mart yang akan dikembangkan menggunakan SQL Server sebagai basis penyimpanan data. Selain itu, SSIS memiliki kemampuan *native integration* dengan ekosistem Microsoft, termasuk SSMS dan SSAS, yang memungkinkan proses ETL berjalan lebih optimal dan efisien dalam mengelola data akademik untuk perhitungan dan pelaporan ELO.

Penelitian ini secara signifikan mengisi celah dalam literatur dengan berfokus pada perancangan *independent data mart* untuk pelaporan *Expected Learning Outcomes* (ELO) mahasiswa, sebuah area yang belum banyak dieksplorasi secara mendalam. Pada artikel [19], pengembangan *data mart* berhasil mengintegrasikan data siswa sekolah dasar ke dalam *star schema* yang kemudian digunakan dalam sistem pendukung keputusan (*Decision Support System*) untuk memprediksi performa dan kelulusan siswa. Dengan menerapkan sembilan algoritma *data mining*, penelitian tersebut mencapai tingkat akurasi sangat tinggi, yaitu 96,76% untuk prediksi performa dan 96,12% untuk prediksi nilai akhir menggunakan algoritma *Expectation-Maximization Clustering*. Keunggulan utama penelitian [19] terletak pada kekuatan analitiknya melalui integrasi *OLAP cube* dan

data mining berbasis algoritma pembelajaran mesin, serta keberhasilannya dalam menyajikan model prediktif yang akurat.

Namun, penelitian [19] memiliki keterbatasan dalam hal ruang lingkup dimensi analisis, karena hanya mengeksplorasi dimensi terbatas yaitu kelas, mata pelajaran, dan status kelulusan, tanpa mengaitkannya dengan kompetensi pembelajaran yang lebih spesifik. Selain itu, pendekatan yang digunakan lebih menekankan pada prediksi performa, bukan pelaporan berbasis hasil pembelajaran (*outcome-based reporting*) seperti ELO. Di sisi lain, penelitian [16] juga memiliki fokus utama pada prediksi performa siswa melalui pemantauan nilai akademik secara umum. Keunggulan penelitian [16] adalah kemampuannya dalam menangkap tren performa siswa menggunakan algoritma klasifikasi, namun belum menyentuh kedalaman analisis terhadap pencapaian kompetensi spesifik atau pemetaan nilai terhadap ELO tertentu.

Sebagai pembeda utama, penelitian ini menasar dimensi yang lebih kaya dan kontekstual dengan sistem akademik perguruan tinggi, seperti semester, mata kuliah, jenis pembelajaran (teori/praktikum), dan capaian pembelajaran (ELO) A sampai J. Penelitian ini juga tidak berfokus pada prediksi performa, melainkan pada fungsi utama *data mart* sebagai repositori data historis (*snapshot*) yang terstruktur dan siap dianalisis untuk pelaporan outcome. Selain itu, penelitian ini mengintegrasikan nilai UTS, UAS, tugas, dan nilai ELO spesifik dari berbagai jenis, mata kuliah, kelas, dan semester, sehingga memungkinkan analisis mendalam terhadap kompetensi yang telah dikuasai oleh mahasiswa, termasuk perbandingan antar mata kuliah yang memiliki ELO serupa—kemampuan analisis yang belum tersedia dalam dua penelitian sebelumnya. Penggunaan metodologi Kimball *Bottom-Up Nine-Step*, integrasi data dari Excel yang tidak terstruktur, serta pengujian kualitas data dan dashboard melalui *heuristic evaluation* dan *Think-Aloud Protocol* juga menambah kekuatan kontribusi penelitian ini dari aspek teknis dan usability.

Meski demikian, penelitian ini belum mengeksplorasi pendekatan prediktif seperti yang dilakukan pada penelitian [16] dan [19]. Kedua penelitian tersebut

telah menunjukkan potensi kuat dalam penggunaan *data mining* untuk klasifikasi dan prediksi performa siswa. Oleh karena itu, sebagai saran pengembangan, penelitian ini dapat dikembangkan lebih lanjut dengan mengintegrasikan modul analitik prediktif menggunakan algoritma pembelajaran mesin untuk mengantisipasi pencapaian ELO berdasarkan data historis mahasiswa. Selain itu, integrasi dengan OLAP cube juga dapat menjadi alternatif pengayaan fitur analisis multidimensi secara interaktif.

Sebagai kesimpulan, penelitian ini memberikan kontribusi baru dalam pengembangan *independent data mart* untuk pelaporan ELO yang fleksibel, mandiri, dan sesuai dengan kebutuhan institusi pendidikan tinggi di Indonesia. Dengan cakupan dimensi yang lebih luas, pendekatan integrasi data yang efisien, dan fokus pada outcome daripada prediksi, penelitian ini memperluas arah pemanfaatan data mart di dunia pendidikan tidak hanya sebagai alat pemantauan, melainkan sebagai fondasi utama dalam mendukung kebijakan berbasis capaian pembelajaran.

2.2 Teori Penelitian

2.2.1 Peran Business Intelligence pada Data Mart

Sistem Business Intelligence mengintegrasikan data operasional dan historis dengan memanfaatkan alat analitik guna menghasilkan informasi yang bernilai, yang selanjutnya dapat digunakan oleh pengambil keputusan untuk mendukung proses pengambilan keputusan secara lebih efektif dan berbasis data [25]. Kepemilikan data analitis saja belum memberikan nilai tambah bagi pemiliknya. Agar data dapat digunakan secara efektif diperlukan visualisasi data yang tepat, mudah dibaca, serta dapat dipahami. Visualisasi ini dapat dilakukan melalui pemanfaatan alat Business Intelligence. Salah satu alat yang paling mudah digunakan dan populer adalah Microsoft Power BI atau Tableau Desktop. Selain fungsi visualisasi, alat ini juga menyediakan fitur lanjutan yang memungkinkan transformasi data serta penciptaan metrik baru untuk mendeskripsikan secara kuantitatif area yang dianalisis [26].

Hal serupa juga terjadi di bidang pendidikan. Perguruan tinggi atau seperti dalam studi kasus penelitian ini menghasilkan data berjumlah besar yang heterogen, yang apabila dianalisis dengan tepat dapat menjadi dasar dalam pengambilan keputusan dengan berbasis data, bukan sekadar intuisi. Namun, analisis terhadap data dalam jumlah besar ini tidak mungkin dilakukan tanpa dukungan teknologi dan perangkat modern yang mana di dalam ruang lingkup Business Intelligence [20]. Dalam konteks ini, integrasi konsep Business Intelligence ke dalam proses akademik memiliki potensi besar untuk meningkatkan efisiensi dan kualitas pengambilan keputusan. BI memungkinkan pemanfaatan berbagai alat analitik yang mendukung kebutuhan informasi dari berbagai pemangku kepentingan, seperti mahasiswa, dosen, administrator, hingga pengambil kebijakan [27].

Pendekatan Business Intelligence (BI) dirancang sebagai suatu kerangka interoperabilitas yang memanfaatkan teknologi untuk mengintegrasikan data, mengelola volume data yang besar, serta mengembangkan query analitis dengan tetap memperhatikan aspek keamanan informasi dan tata kelola data yang baik [28]. Salah satu teknologi utama yang mendukung pendekatan ini adalah data mart. Dalam arsitektur Business Intelligence, data mart memainkan peran sebagai jembatan antara data warehouse dan proses analisis yang dilakukan oleh pengguna akhir. Data mart berfungsi sebagai repositori yang berisi subset data dari data warehouse yang telah melalui proses pembersihan, integrasi, dan pengorganisasian sesuai kebutuhan analisis spesifik, seperti untuk departemen atau unit bisnis tertentu. Jika menyediakan data yang lebih terfokus, relevan, dan mudah diakses, data mart mendukung pengambilan keputusan yang lebih cepat, akurat, dan berbasis data, selaras dengan tujuan utama implementasi Business Intelligence [29]. Oleh demikian, keberadaan data mart dalam kerangka Business Intelligence tidak hanya meningkatkan efisiensi proses analisis data, tetapi juga memperkuat kemampuan organisasi dalam merespons dinamika bisnis secara tepat dan strategis.

2.2.2 Independent Data Mart

Data mart merupakan bagian dari data warehouse yang berfungsi sebagai subkumpulan data terfokus pada suatu bidang atau departemen tertentu dalam organisasi. Data warehouse sendiri tersusun dari berbagai data mart kecil yang konsisten dan berorientasi subjek, dirancang untuk memenuhi kebutuhan informasi serta analisis pada tingkat departemen [26]. Sebuah data mart menyimpan data yang telah diringkas pada tingkat hierarki tertentu, dan berbeda dengan sumber data transaksional yang umumnya menggunakan skema relasional, data mart dirancang dalam bentuk skema *dimensional* guna mempermudah akses terhadap data [30]. Demikian, data mart dapat diartikan sebagai versi lebih ringkas dari data warehouse yang dirancang untuk memenuhi kebutuhan pengguna dalam skala yang lebih terbatas. Ia merupakan repositori data yang spesifik dan berorientasi subjek, yang dikembangkan untuk menjawab pertanyaan-pertanyaan tertentu dari pengguna tertentu [31]. Data mart hanya menyimpan data yang relevan untuk analisis lokal, dan berfokus pada satu topik atau area fungsional seperti penjualan, keuangan, pemasaran, sumber daya manusia, atau dalam konteks penelitian ini analisis capaian pembelajaran (ELO). Berdasarkan hal tersebut, data mart merepresentasikan data yang berasal dari satu proses bisnis tertentu. Berikut pada tabel 2.2 adalah 6 alasan data mart dibuat [31].

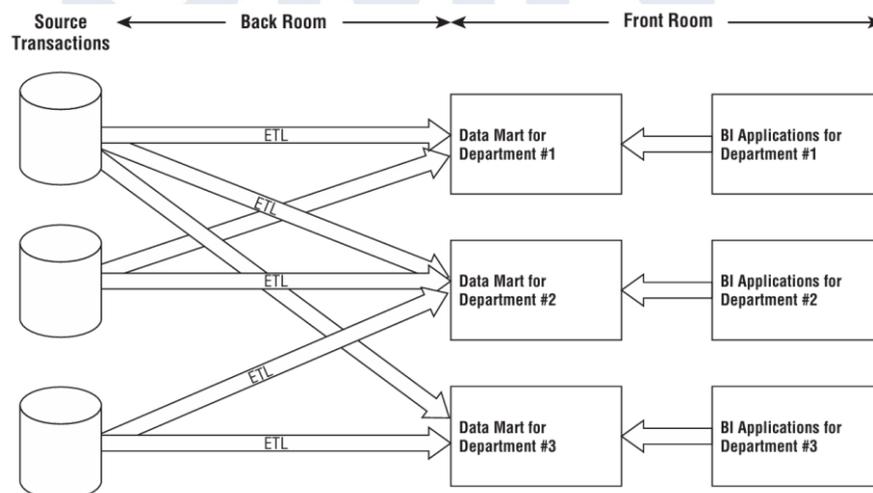
Tabel 2. 2 Alasan dibuatnya data mart

Alasan Dibuatnya Data Mart	
Kemudahan Pembuatan	Data mart relatif mudah untuk dikembangkan karena cakupannya terbatas pada satu subjek tertentu.
Waktu Respons	Meningkatkan waktu respons bagi pengguna akhir karena data yang disajikan lebih terfokus dan mudah diakses.
Biaya Lebih Rendah	Memiliki biaya pengembangan yang lebih rendah dibandingkan dengan penerapan data warehouse secara keseluruhan.

Kemudahan Pengembangan	Data mart lebih mudah untuk ditingkatkan atau dikembangkan seiring waktu sesuai kebutuhan.
Kesederhanaan	Karena hanya berfokus pada satu topik atau subjek, struktur dan penggunaannya menjadi lebih sederhana.
Ruang Lingkup Terbatas	Cakupan yang terbatas mempermudah pemahaman dan implementasi oleh pengguna.

Berdasarkan sumber datanya, data mart terbagi menjadi dua pendekatan yaitu data mart dependen dan data mart independen [31]. Data mart independen tidak bergantung pada gudang data pusat atau data warehouse maupun dengan data mart lainnya, karena pendekatan ini merancang data mart terlebih dahulu lalu baru membuka kesempatan untuk melakukan integrasi dengan data mart lainnya untuk merancang sebuah data warehouse [31]. Setiap data mart mengumpulkan informasi dari sumbernya dan bukan dari data warehouse. Data mart independen cocok untuk perusahaan kecil, tetapi hanya departemen tertentu yang perlu mengakses dan menganalisis informasi [32].

Arsitektur Data mart yang diilustrasikan oleh Kimball terdiri dari tiga komponen atau tiga layer. Berikut pada gambar 2.2 adalah ilustrasi dari arsitektur data mart [33].



Gambar 2. 1 Arsitektur data mart menurut Kimball

Pada gambar 2.2 Kimball mengilustrasikan arsitektur data mart yang terdiri dari tiga bagian yaitu *source transaction*, *back room*, dan *front room*. Apa yang terjadi diantara bagian arsitektur itu sendiri adalah source data dan data mart dijembatani oleh proses ETL, yang selanjutnya akan dimanfaatkan atau digunakan oleh user di bagian *front room* lewat tools BI. Ilustrasi arsitektur data mart menurut Kimball secara umum merepresentasikan fungsi dan peran data mart dalam ekosistem Business Intelligence. Namun, dalam implementasinya, arsitektur data mart dapat disesuaikan dan diinterpretasikan lebih rinci berdasarkan kebutuhan dan karakteristik masing-masing studi kasus.

2.2.3 Extract, Transform, Loading (ETL)

ETL (Extract, Transform, Load) merupakan tahapan dalam pemrosesan data dari berbagai sumber menuju data warehouse. Proses ini bertujuan untuk mengekstraksi data dari sumber yang berbeda, menyaring dan mengolahnya, kemudian mengintegrasikan data yang relevan ke dalam sistem penyimpanan terpusat seperti data warehouse. Selain berfungsi sebagai alat konsolidasi data, ETL juga dapat digunakan untuk menyatukan informasi dari sistem yang telah ada sebelumnya agar dapat dianalisis secara komprehensif. Melalui proses ETL, data yang tersimpan dalam data warehouse menjadi memenuhi karakteristik utama seperti bersifat historis, terintegrasi, ringkas, tidak berubah, dan memiliki struktur yang disesuaikan untuk kebutuhan analitik [34]. Fungsi dari masing-masing fase itu sendiri adalah sebagai berikut [35]:

1. Ekstraksi (*Extraction*): Proses ini melibatkan menghubungkan sistem sumber dan mengumpulkan data yang diperlukan untuk pemrosesan analitik.
2. Transformasi (*Transformation*): Data yang telah diekstrak kemudian diubah ke dalam format standar agar dapat digunakan secara konsisten dalam sistem target.
3. Pemuatan (*Loading*): Data yang telah ditransformasikan diimpor ke dalam data warehouse agar siap digunakan untuk analisis lebih lanjut.

ETL menganonimkan data sesuai dengan ketentuan regulasi, sehingga informasi yang bersifat rahasia dan sensitif dianonimkan sebelum dimuat ke dalam penyimpanan data target [36]. ETL menghilangkan data yang tidak diperlukan dari database operasional agar tidak dimuat ke dalam *data mart*. ETL juga melakukan penyesuaian data yang berasal dari berbagai sumber, serta menghitung ringkasan dan data turunan. ETL tools juga membantu dalam pemeliharaan metadata, memastikan bahwa informasi tentang struktur dan sumber data tetap terkelola dengan baik.

2.2.4 Entity Relationship Diagram (ERD)

Entity-Relationship Diagram (ERD) merupakan representasi visual yang digunakan dalam perancangan basis data untuk menunjukkan hubungan antar entitas atau objek dalam suatu sistem, termasuk atribut-atribut yang dimiliki oleh entitas tersebut [37]. ERD yang akan dirancang pada penelitian ini mengikuti metode *crow foot* yang mana bersifat intuitif, mudah dipahami, dan memberikan representasi yang jelas tentang hubungan antara entitas [38]. Diagram ini terdiri dari tiga komponen utama, yaitu entitas, atribut, dan hubungan (relasi) antar entitas yang menggambarkan struktur logis dari data yang akan digunakan.

1. Entitas

Entitas merujuk pada objek atau hal yang memiliki peran penting dan perlu direkam dalam sebuah basis data [39]. Objek ini bisa berupa individu, lokasi, benda, atau hal lain yang relevan dengan sistem yang dirancang. Setiap entitas umumnya memiliki sejumlah atribut yang menjelaskan ciri atau properti yang membedakannya dari entitas lainnya.

2. Atribut

Atribut merupakan detail atau informasi yang menjelaskan karakteristik dari suatu entitas [39]. Setiap entitas harus memiliki sebuah *primary key* yang berfungsi sebagai identitas unik, serta atribut lain yang bersifat deskriptif. Misalnya, jika entitas yang dimaksud adalah seseorang, maka atribut yang dimiliki bisa mencakup nama, alamat, dan informasi relevan lainnya.

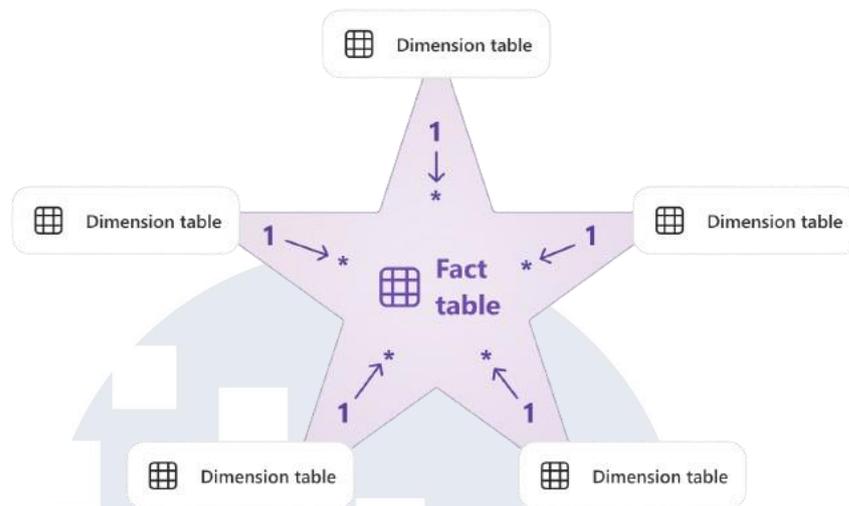
3. Relasi

Relasi menunjukkan bagaimana entitas saling terhubung dalam sebuah basis data, baik itu antara dua entitas maupun lebih [39]. Tipe hubungan ini bisa berupa *one-to-one* (satu entitas terhubung dengan satu entitas lainnya), *one-to-many* (satu entitas terhubung dengan banyak entitas lain), atau *many-to-many* (banyak entitas saling terhubung dengan banyak entitas lainnya). Sebagai contoh dalam penelitian ini, terdapat relasi antara entitas *data_id* dan *invoice*, di mana satu nasabah dapat memiliki beberapa pesanan (*one-to-many*). Hubungan semacam ini penting untuk memahami keterkaitan antar data serta interaksi yang terjadi di dalam sistem yang dibangun.

Dengan memanfaatkan ERD, perancang basis data dapat menggambarkan struktur data secara visual sehingga lebih mudah memahami hubungan antar entitas dalam sistem yang akan dibangun. Entity-Relationship Diagram juga berperan sebagai salah satu metode pemodelan yang digunakan dalam perancangan basis data, khususnya untuk menghasilkan skema konseptual yang merepresentasikan model data semantik, seperti pada perancangan model konseptual dalam Data Mart [37].

2.2.5 Skema Bintang

Star Schema atau skema bintang adalah salah satu model desain basis data yang umum digunakan dalam pembangunan *data warehouse*, terutama untuk kebutuhan analisis data. Skema ini menjadi pilihan yang cukup populer dibandingkan dengan skema lain seperti *snowflake* karena memiliki struktur yang lebih sederhana dan mendukung proses OLAP secara efisien [40]. Dalam skema bintang, terdapat dua komponen utama: data dimensi, yang memberikan konteks agregasi terhadap data, dan data fakta, yang merepresentasikan transaksi atau kejadian spesifik [40].



Gambar 2. 2 Skema bintang

Gambar 2.3, menggambarkan struktur dari *star schema*, di mana satu tabel fakta ditempatkan di bagian tengah dan dikelilingi oleh sejumlah tabel dimensi yang terhubung langsung melalui kunci relasi. Susunan ini membentuk pola visual menyerupai bintang, dengan tabel fakta sebagai inti, sementara tabel-tabel dimensi berperan sebagai "sinar" atau "lengan" dari bintang tersebut.

2.3 Teori tentang Framework/Algoritma yang Digunakan

2.3.1 Kimball *Bottom Up Approach Nine-Step Methodology*

Perancangan data mart dilakukan dengan mengikuti metodologi yang sistematis agar prosesnya berjalan secara efektif, efisien, dan akurat. Berbagai pendekatan tersedia dalam pembangunan arsitektur data warehouse yang efektif. Setiap data warehouse dirancang dengan karakteristik yang berbeda, bergantung pada kebutuhan pengguna bisnis di masing-masing area fungsional organisasi [23]. Salah satu framework yang umum digunakan dalam perancangan data mart yang mana juga diaplikasikan pada penelitian terdahulu yang menjadi acuan dari penelitian ini adalah framework Kimball *nine-step methodology* dan Bill Inmon *Methodology* [23]. Berikut pada tabel 2.3 dilakukan perbandingan dari kedua metodologi.

Tabel 2. 3 Perbandingan metodologi

Aspek	Bill Inmon	Ralph Kimball
-------	------------	---------------

Keseluruhan pendekatan	Metode top-down yang mana dimulai dengan merancang struktur data keseluruhan.	Metode bottom-up yang mana dimulai dengan level data paling rendah dan nantinya bisa dibangun ke struktur lebih kompleks.
Arsitektur	Enterprise-wide data warehouse yang menyediakan data ke data mart	Data mart dimodelkan berdasarkan kebutuhan bisnis dan konsistensi dicapai melalui bentuk data bus dan <i>dimensional modeling</i> .
Kompleksitas metode	Sangat kompleks	Cukup sederhana
Orientasi data	Berorientasi pada subjek atau data-driven	Berorientasi pada proses
Tujuan Utama	Fokus utamanya adalah membuat solusi teknis yang kuat dan rapi, menggunakan metode database yang sudah terbukti.	Fokus utamanya adalah membuat solusi yang mudah digunakan oleh pengguna, bisa di-query langsung.

Oleh karena itu, pada penelitian ini akan menggunakan Metode Kimball tepatnya Kimball *Nine-Step Methodology*. Pada penelitian ini, data mart dirancang untuk memenuhi kebutuhan spesifik, yaitu untuk pelaporan nilai ELO. Selain itu, data mart juga disusun dari file-file eksternal yaitu file excel hasil capaian mahasiswa dan bukan dari data warehouse. Dari kedua pertimbangan ini, metode kimball terbukti lebih relevan untuk diaplikasikan karena pendekatannya yang bersifat bottom-up atau merancang dari tingkat terendah. Data mart ini juga dirancang berdasarkan proses bisnis dan kebutuhan bisnisnya. Berdasarkan hal tersebut, penerapan metode Kimball dalam penelitian ini dinilai paling sesuai karena mampu menjawab kebutuhan pelaporan yang spesifik, memanfaatkan sumber data eksternal, serta menyesuaikan dengan alur proses bisnis yang ada. Pendekatan bottom-up yang dimiliki Kimball memungkinkan perancangan data mart yang lebih terfokus,

efisien, dan langsung dapat digunakan untuk analisis serta pengambilan keputusan.

2.4 Teori Tentang Tools/Software yang Digunakan

2.4.1 SQL Server Management Studio (SSMS)

Microsoft SQL Server merupakan sebuah sistem manajemen basis data relasional atau biasa dikenal dengan relational database management system (RDBMS) produk dari Microsoft. Bahasa query utamanya adalah Transact-SQL yang merupakan implementasi dari SQL standar ANS/ISO yang digunakan oleh Microsoft dan Sybase. Pada umumnya, SQL Server digunakan di dua bisnis yang memiliki basis data berskala kecil sampai dengan menengah, tapi kemudian berkembang dengan digunakannya SQL Server pada basis data yang lebih besar [37]. Bila diperhatikan dari literature review, terdapat 3 penelitian yang menggunakan SSMS sebagai database dan 1 penelitian yang menggunakan PostgreSQL. Berdasarkan hal tersebut, berikut pada tabel 2.4 disajikan perbandingan antara kedua database tersebut untuk menentukan database mana yang lebih relevan untuk digunakan pada penelitian ini.

Tabel 2. 4 Perbandingan database

Aspek Pemanding	SQL Server Management Studio (SSMS)	PostgreSQL
Definisi	SSMS adalah alat manajemen resmi dari Microsoft untuk mengelola SQL Server, termasuk konfigurasi, pemantauan, dan pengembangan basis data.	PostgreSQL adalah sistem manajemen basis data relasional open-source yang mendukung fitur canggih seperti MVCC, JSON, dan kepatuhan tinggi terhadap standar SQL.
Kelebihan	1. Antarmuka GUI yang user-friendly.	1. Gratis dan open-source. 2. Mendukung standar SQL dan fitur lanjutan (misalnya: JSON,

	<ol style="list-style-type: none"> 2. Terintegrasi langsung dengan Microsoft SQL Server. 3. Fitur debugging dan tuning yang lengkap. 4. Dukungan penuh dari Microsoft 	<ol style="list-style-type: none"> XML, indexing tipe data kompleks). 3. Cross-platform dan dapat berjalan di berbagai OS.
Kekurangan	<ol style="list-style-type: none"> 1. Hanya dapat digunakan dengan SQL Server. 2. Tidak open-source. 3. Hanya tersedia di Windows (secara resmi). 4. Memerlukan lisensi SQL Server untuk fitur-fitur enterprise 	<ol style="list-style-type: none"> 1. Tidak memiliki GUI resmi sebaik SSMS (meskipun ada pgAdmin dan lainnya). 2. Kurva pembelajaran lebih curam bagi pemula 3. Beberapa fitur enterprise memerlukan konfigurasi tambahan

Berdasarkan hasil studi pustaka dan perbandingan karakteristik antara SQL Server Management Studio (SSMS) dan PostgreSQL, maka diputuskan bahwa SSMS dipilih sebagai sistem manajemen basis data untuk penelitian ini. Pemilihan ini didasarkan pada sejumlah pertimbangan teknis dan relevansi terhadap kebutuhan penelitian. Pertama, SSMS merupakan alat manajemen resmi yang terintegrasi penuh dengan SQL Server, sehingga mendukung proses konfigurasi, pemantauan, serta pengembangan basis data secara terpusat dan efisien. Dukungan fitur seperti *query editor*, *index tuning*, serta *Integration Services Catalog* untuk pengelolaan proses ETL menjadikan SSMS sangat sesuai dengan kebutuhan pembangunan data mart yang stabil dan terotomatisasi. Selain itu, hasil tinjauan literatur menunjukkan bahwa SSMS lebih sering digunakan dalam studi-studi sebelumnya (3 dari 4 referensi), yang menunjukkan tingkat adopsi dan kepercayaan yang lebih tinggi dalam konteks

pengembangan data mart dan data warehouse. Bila dilakukan pertimbangan integrasi, kemudahan pengelolaan, serta dukungan komunitas dan dokumentasi dari Microsoft, maka SSMS dinilai sebagai pilihan yang paling relevan dan efektif untuk menyimpan dan mengelola data mart pada penelitian ini.

2.4.2 SQL Server Integration Services (SSIS)

SSIS (*SQL Server Integration Services*) adalah alat ETL (Extract, Transform, Load) milik Microsoft yang dirancang untuk memfasilitasi pemrosesan data, integrasi antar sumber data, serta otomatisasi alur kerja (*workflow*) dalam sistem basis data [39]. Pada penelitian ini, SSIS digunakan sebagai perangkat lunak utama dalam merancang paket ETL yang memuat data dari file Excel ke dalam *data mart* berbasis SQL Server. Dua tools yang umum digunakan untuk merancang paket ETL sebagaimana yang digunakan oleh penelitian yang dibahas pada tabel penelitian terdahulu adalah SSIS dan Pentaho. Berdasarkan hal tersebut, berikut pada tabel 2.5 adalah perbandingan dari SSIS dan Pentaho.

Tabel 2. 5 Perbandingan tools ETL

Aspek Pembeding	SSIS	Pentaho
Definisi	SQL Server Integration Services (SSIS) adalah komponen perangkat lunak basis data Microsoft SQL Server yang dapat digunakan untuk melakukan berbagai tugas migrasi data .	Pentaho adalah alat ETL open-source yang memungkinkan pengguna membuat transformasi dan pekerjaan ETL melalui antarmuka grafis.
Kelebihan	1. Antarmuka visual yang memudahkan pembuatan paket ETL tanpa perlu banyak penulisan kode.	1. Open-source dan bebas biaya lisensi, cocok untuk berbagai skala bisnis. 2. Mendukung berbagai sumber data, termasuk

	<p>2. Integrasi penuh dengan SQL Server Management Studio (SSMS) dan SQL Server Agent untuk eksekusi dan penjadwalan paket ETL.</p> <p>3. Mendukung berbagai sumber data dan memiliki kemampuan transformasi data yang kuat.</p>	<p>file flat, database relasional, dan NoSQL.</p> <p>3. Antarmuka drag-and-drop yang intuitif untuk membangun alur ETL.</p>
Kekurangan	<p>1. Keterbatasan integrasi sumber data non-microsoft.</p> <p>2. Kurva pembelajaran yang curam.</p>	<p>1. Kinerja yang kurang optimal pada dataset besar.</p> <p>2. Antarmuka Pengguna yang Tidak Konsisten dan Dapat Membingungkan</p>

Berikut adalah item yang digunakan pada SSIS di Visual Studio 2022 :

1. *Control Flow Item*

- a. *Data Flow Task* : merupakan sebuah *data flow* yang berisi proses-proses yang akan kita bangun didalamnya.
- b. *Execute SQL Task* : digunakan untuk menjalankan perintah SQL di luar pipeline data atau di Control Flow.
- c. *Script Task* : memberikan fleksibilitas dengan memungkinkan menulis kode kustom (VB.NET atau C#) untuk melakukan logika yang tidak bisa dilakukan oleh komponen SSIS biasa.

2. *Data Flow Sources*

- a. *OLE DB Source* : merupakan tempat menyimpan *source* yang akan digunakan. *Source* di *OLE DB Source* dapat berupa Syntak SQL atau table.
- b. *Excel Source* : Source dalam bentuk file Excel.

3. *Data Flow Transformations*

- a. *Agregate* : berfungsi untuk melakukan group by dan sum / count.
- b. *Conditional Split* : berfungsi untuk melakukan pemisahan nilai pada destinasi berbeda.
- c. *Data Conversion* : digunakan untuk mengubah tipe data kolom agar sesuai dengan kebutuhan tabel tujuan atau operasi selanjutnya.
- d. *Derived Column* : berfungsi untuk melakukan case when, perubahan nama kolom, dll.
- e. *Lookup* : digunakan untuk mencocokkan (match) data yang sedang diproses dengan data di sumber lain (biasanya database).
- f. *Merge Join* : berfungsi untuk melakukan penggabungan 2 buah table dengan fungsi join.
- g. *Multicast* : berfungsi untuk melakukan duplicate hasil.
- h. *OLE DB Command* : berfungsi untuk menjalankan perintah SQL untuk setiap baris data secara satu per satu.
- i. *Sort* : berfungsi untuk melakukan sorting
- j. *Union All* : berfungsi untuk menggabungkan 2 atau lebih dengan masing-masing harus memiliki nama kolom yang sama.

4. *Data Flow Destination*

- a. *OLE DB Destination* : merupakan tempat menyimpan destinasi setelah diproses dengan *data flow transformation*.
- b. *Flat File Destination* : *Destination* yang berupa file.txt .
- c. *Excel Destination* : *Destination* berupa file excel.

2.5 Teori Tentang Pengujian dan Evaluasi yang Digunakan

2.5.1 ETL Data Quality Testing

Dalam penelitian ini, dilakukan pengujian kualitas data (ETL Data Quality Testing) terhadap data yang dimuat ke dalam data mart guna memastikan bahwa seluruh proses ekstraksi, transformasi, dan pemuatan (ETL) berjalan secara akurat dan sesuai dengan standar kualitas data. Pengujian ini mengacu pada framework yang dikembangkan oleh Sara B., Tarek M., dan Abdelmgeid A. dalam artikel berjudul "*Automated ETL Testing on the Data Quality of a Data Warehouse*" [41]. Framework tersebut dirancang untuk melakukan pemeriksaan kualitas data dengan mengacu pada lima dimensi utama dan serangkaian *routine check* yang telah ditetapkan sebelumnya.

1. Completeness

Menunjukkan sejauh mana seluruh data yang relevan telah tersedia dan tidak terdapat nilai yang hilang (*missing*). Dimensi ini memastikan bahwa seluruh data yang dibutuhkan untuk analisis atau pelaporan benar-benar ada dan tercatat.

2. Consistency

Menilai konsistensi nilai data antar dataset atau antar tabel. Dimensi ini memastikan bahwa data tidak menunjukkan konflik logis atau perbedaan nilai untuk entitas yang sama di lokasi berbeda.

3. Uniqueness

Memastikan bahwa tidak ada duplikasi data di dalam sistem, khususnya pada atribut atau entitas yang seharusnya bersifat unik, seperti NIM mahasiswa atau kode mata kuliah.

4. Validity

Memastikan bahwa data sesuai dengan format, tipe data, atau rentang nilai yang telah ditentukan. Dimensi ini mencakup validasi terhadap struktur sintaks dan aturan bisnis terkait data.

5. Accuracy

Mengukur sejauh mana data menggambarkan kondisi sebenarnya di dunia nyata. Data dikatakan akurat apabila tidak terdapat kesalahan input dan nilai yang disimpan merepresentasikan objek atau kejadian dengan benar.

Untuk memvalidasi kelima dimensi tersebut, dilakukan serangkaian pengujian yang dikenal sebagai routine check. Routine check merupakan sekumpulan query atau prosedur otomatis yang dirancang untuk mencakup seluruh aspek pengujian kualitas data sesuai dimensi di atas. Dalam penelitian ini, terdapat sembilan jenis routine check yang diterapkan.

1. Record Count Validation dilakukan untuk memastikan jumlah data antara staging dan data mart konsisten.
2. Data Duplicate dilakukan untuk mengidentifikasi adanya duplikasi data yang tidak diinginkan.
3. Check Integrity Constraint dilakukan untuk memverifikasi integritas relasional antar tabel (misalnya foreign key).
4. Check Data Boundaries dilakukan untuk memastikan data berada dalam batas nilai yang diharapkan (range check).
5. Field Mapping dilakukan untuk memastikan bahwa kolom-kolom dari staging telah dimuat ke kolom yang sesuai di data mart
6. Hierarchy Level Integrity dilakukan untuk memastikan hubungan antar level hierarki (misalnya semester–mata kuliah–mahasiswa) tetap valid
7. Field Data Type dilakukan untuk memastikan data dimuat sesuai dengan tipe data yang ditentukan di skema data mart.
8. Field Length Check dilakukan untuk memastikan panjang data tidak melebihi kapasitas kolom yang disediakan.
9. Field-to-Field Comparison dilakukan untuk membandingkan nilai antar kolom untuk memastikan hubungan logis yang benar antar atribut.

2.5.2 Usability Testing

Usability evaluation merupakan metode yang digunakan untuk mengidentifikasi permasalahan spesifik terkait kegunaan suatu produk. Evaluasi ini bertujuan memastikan bahwa sistem yang dikembangkan dapat digunakan secara efisien, efektif, dan memuaskan oleh pengguna. Manfaat dari pelaksanaan evaluasi usability secara iteratif mencakup peningkatan prediktabilitas produk, peningkatan produktivitas dengan minimnya kesalahan

pengguna, kesesuaian yang lebih baik dengan kebutuhan pengguna, serta penghematan waktu dan biaya pengembangan [42].

Dalam penelitian ini, evaluasi usability terhadap dashboard data mart dilakukan menggunakan kombinasi dua metode, yaitu heuristic evaluation dan think-aloud protocol. Pendekatan ini diadaptasi dari studi Po-Yin Yen dan Suzanne Bakken dalam artikel berjudul “*A Comparison of Usability Evaluation Methods: Heuristic Evaluation versus End-User Think-Aloud Protocol – An Example from a Web-based Communication Tool for Nurse Scheduling*” [42]. Tujuan utama dari evaluasi ini adalah untuk menilai apakah dashboard yang dikembangkan telah layak digunakan dan diterima oleh pengguna (*user acceptance*).

Think-Aloud Protocol pertama kali dikembangkan oleh Lewis pada tahun 1982 untuk memahami proses kognitif dalam menyelesaikan tugas [43]. Metode ini mendorong partisipan untuk secara verbal mengungkapkan apa yang mereka lihat, pikirkan, lakukan, dan rasakan selama menjalankan suatu tugas [9]. Hal ini memungkinkan pengamat untuk menangkap proses kognitif yang berlangsung secara real-time dan mengidentifikasi permasalahan usability secara langsung. Penggunaan partisipan dari kalangan pengguna aktual (*intended users*) memberikan gambaran nyata tentang bagaimana sistem digunakan dan membantu mengungkap permasalahan praktis yang berkaitan dengan performa tugas [44]. Metode ini sesuai dengan konteks penelitian yang melibatkan jumlah pengguna yang terbatas, karena mampu memberikan insight mendalam secara kualitatif dari interaksi pengguna terhadap sistem.

Selain itu, evaluasi usability dalam penelitian ini juga menggunakan pendekatan heuristic evaluation, sebagaimana diterapkan dalam literatur oleh Nielsen [22]. Heuristic evaluation merupakan metode evaluasi berbasis prinsip-prinsip usability yang dilakukan dengan menilai antarmuka sistem terhadap sepuluh aturan atau *heuristik* usability. Sepuluh prinsip heuristik tersebut adalah:

1. *Visibility of System Status*

Sistem harus selalu memberi tahu pengguna tentang apa yang sedang terjadi, melalui umpan balik (feedback) yang tepat waktu dan dapat dimengerti.

2. *Match Between System and the Real World*

Sistem harus menggunakan bahasa, istilah, dan konvensi yang familiar bagi pengguna (bukan istilah teknis yang hanya dimengerti developer).

3. *User Control and Freedom*

Pengguna sebaiknya diberi opsi untuk membatalkan (undo) atau mengulang (redo) tindakan dengan mudah, terutama ketika mereka melakukan kesalahan.

4. *Consistency and Standards*

Sistem sebaiknya menggunakan elemen antarmuka dan istilah secara konsisten serta mengikuti standar desain umum.

5. *Error Prevention*

Lebih baik mencegah kesalahan daripada harus menampilkan pesan error setelahnya. Desain harus meminimalkan kemungkinan pengguna membuat kesalahan.

6. *Recognition Rather Than Recall*

Kurangi beban ingatan pengguna dengan menampilkan elemen, tindakan, atau pilihan yang terlihat. Pengguna tidak seharusnya mengingat informasi dari satu halaman ke halaman lain.

7. *Flexibility and Efficiency of Use*

Sistem harus efisien digunakan baik oleh pengguna pemula maupun mahir. Sediakan cara pintas (shortcuts) atau personalisasi untuk mempercepat interaksi.

8. *Aesthetic and Minimalist Design*

Tampilan tidak boleh memuat informasi yang tidak relevan atau tidak diperlukan. Desain harus sederhana dan fokus pada konten utama.

9. *Help Users Recognize, Diagnose, and Recover from Errors*

Pesan kesalahan harus jelas, menunjukkan masalahnya, dan memberikan saran bagaimana memperbaikinya.

10. Help and Documentation

Meskipun sistem idealnya mudah digunakan tanpa bantuan, tetap perlu menyediakan dokumentasi atau bantuan online yang mudah dicari dan relevan

Proses evaluasi dilakukan dengan mengelompokkan umpan balik atau pernyataan dari pengguna yang diperoleh selama sesi *think-aloud* ke dalam sepuluh kategori heuristik tersebut. Berdasarkan hal tersebut, setiap pernyataan, keluhan, atau kesan pengguna dapat dikaitkan dengan pelanggaran atau pemenuhan prinsip tertentu, sehingga memberikan pemetaan yang sistematis terhadap tingkat *usability* dari dashboard [42].

