

Perancangan *Data Warehouse* Perguruan Tinggi untuk Kinerja Penelitian dan Pengabdian kepada Masyarakat

<http://dx.doi.org/10.28932/jutisi.v6i2.2557>

Agata Filiana^{#1}, Andhika Galuh Prabawati^{#2}, Maria Nila Anggia Rini^{#3}, Gloria Virginia^{#4}, Budi Susanto^{#5}

Informatika, Universitas Kristen Duta Wacana

Jl. Dr. Wahidin Sudirohusodo 5-25, Yogyakarta 55224

¹afiliana@ti.ukdw.ac.id

²andhika.galuh@staff.ukdw.ac.id

³nila@ti.ukdw.ac.id

⁴virginia@ti.ukdw.ac.id

⁵budsus@ti.ukdw.ac.id

Abstract — Accreditation is an evaluation to ensure the quality of a study program at higher education by completing two main documents for qualitative data and self-evaluation, namely Study Program Performance Report (LKPS) and Self Evaluation Report (LED) respectively. Data used for this process must be consistent and accurate, therefore a central repository is proposed. A data warehouse for the domain of Research and Community Service is built using the snowflake schema and a five-step approach methodology proposed for the educational data warehouse. The schema consisted of one fact table, ten dimension tables, and four bridge tables. Using a business intelligence tool, data from the data warehouse is queried and provided in a dashboard. The data warehouse successfully returned the qualitative data needed for LKPS. As for the evaluation data for LED, the data warehouse is able to fulfill seven out of ten requirements. The data for the remaining three are provided with a slightly different orientation to the requirement, this was due to the data source.

Keywords— data warehouse; accreditation; higher education;

I. PENDAHULUAN

A. Latar Belakang

Kualitas sebuah Perguruan Tinggi secara berkala dievaluasi oleh pemerintah. Badan Akreditasi Nasional Perguruan Tinggi (BAN-PT) merupakan sebuah badan evaluasi mandiri yang dibentuk oleh pemerintah yang memiliki tugas menjaga kelayakan Perguruan Tinggi. Pada Peraturan BAN-PT Nomor 1 Tahun 2020, disebutkan ada dua jenis akreditasi yaitu Akreditasi Perguruan Tinggi (APT) dan Akreditasi Program Studi (APS) yang dilakukan dengan menggunakan instrumen akreditasi yang disusun berdasarkan Standar Nasional Pendidikan Tinggi [1].

Mengikuti kebijakan yang dikeluarkan oleh BAN-PT, Program Studi (Prodi) Informatika Universitas Kristen Duta Wacana menyiapkan diri untuk melakukan APS dengan menggunakan Instrumen Akreditasi Program Studi 4.0 (IAPS 4.0) yang terdiri atas dua orientasi utama yaitu Laporan Kinerja Program Studi (LKPS) dan Laporan Evaluasi Diri (LED). LKPS diisi dengan data kuantitatif, sedangkan LED berisi penjelasan komprehensif. Pada IAPS 4.0 terdapat delapan indikator kinerja utama yaitu: Tata Pamong, Tata Kelola, dan Kerjasama, Mahasiswa, Sumber Daya Manusia, Keuangan, Sarana dan Prasarana, Pendidikan, Penelitian, Pengabdian Kepada Masyarakat, dan Luaran dan Capaian Tridharma. Salah satu esensi yang penting dalam APS adalah evaluasi pelaksanaan Tridharma Perguruan Tinggi yang meliputi pengajaran, penelitian dan pengabdian kepada masyarakat (PkM). Hal ini secara konsisten dievaluasi di setiap indikator kinerja utama.

Saat ini Prodi Informatika belum memiliki sebuah sistem terpusat yang dapat digunakan untuk mengisi borang akreditasi secara efektif dan efisien. Data yang digunakan untuk mengisi IAPS 4.0 didapatkan dari beberapa sistem internal dan rekapan manual dalam kurun waktu yang sudah ditentukan. Sifat IAPS 4.0 yang analitis mengharuskan prodi untuk mengolah data tersebut menjadi sebuah laporan yang dapat merangkum kondisi prodi dalam periode tertentu. Selain itu, data tersebut harus konsisten karena ada keterkaitan antar kriteria. Konsistensi data merupakan kunci dalam penulisan LKPS dan LED, sehingga penting bagi prodi untuk memiliki sebuah repositori terpusat untuk meminimalisasi *human error* dan inkonsistensi data.

Proses APS akan dilakukan secara berkala sehingga diperlukan sebuah solusi yang bersifat permanen dan dapat diperbaharui secara terus menerus. Penelitian ini

menawarkan *data warehouse* sebagai solusi yang tepat untuk permasalahan tersebut. Repositori ini berisi data historis dari berbagai sumber yang sudah dibersihkan dan ditransformasi melalui proses ETL (*Extraction, Transformation dan Loading*) [2]. Berbeda dengan *database* operasional, *data warehouse* dirancang khusus untuk analisis sehingga sifatnya esensial untuk proses pengambilan keputusan [3]. Data untuk kebutuhan APS dapat diambil dari *data warehouse* sesuai dengan granularitas dan dimensinya sehingga analisis dapat dilakukan dengan lebih efisien dan terpusat.

B. Batasan Masalah

Data warehouse yang dirancang untuk kebutuhan APS dapat dibagi menjadi beberapa domain sehingga memudahkan proses pembangunan. Fokus penelitian ini adalah pada pengembangan *data mart* atau *departmental data warehouse* untuk Penelitian dan Pengabdian kepada Masyarakat (PkM). Kedua komponen ini merupakan bagian dari Tridharma Perguruan Tinggi dan juga dibahas secara detail pada IAPS 4.0 di berbagai kriteria. Penelitian dan PkM memiliki karakteristik yang mirip, termasuk dosen dan mahasiswa yang terlibat serta luaran yang dihasilkan, sehingga keduanya dapat disatukan ke dalam domain yang sama. Penelitian ini merupakan bagian dari sistem *data warehouse* akreditasi prodi.

Secara khusus yang menjadi bahasan pada penelitian ini adalah tabel-tabel berikut ini pada IAPS 4.0: Tabel 6.a Penelitian DTSP yang melibatkan mahasiswa, Tabel 7 PkM DTSP yang melibatkan mahasiswa, 3.b.2 Penelitian DTSP, 3.b.3 Pengabdian kepada Masyarakat (PkM) DTSP, 3.b.4 Publikasi Ilmiah DTSP, 3.b.7 Luaran Penelitian/PkM Lainnya oleh DTSP, 8.f.1 Publikasi Ilmiah mahasiswa, dan 8.f.4 Luaran penelitian/PkM lain yang dihasilkan mahasiswa.

C. Rumusan Masalah

Mempertimbangkan latar belakang dan batasan masalah yang ada, maka masalah untuk penelitian ini dapat dirumuskan sebagai berikut: Bagaimana merancang sebuah *data warehouse* yang dapat digunakan untuk mendukung proses pengisian LKPS dan LED yang berhubungan dengan Penelitian dan PkM?

D. Tujuan Penelitian

Tujuan dari penelitian ini adalah menyediakan sebuah repositori terpusat dengan data Penelitian dan PkM dosen yang terstruktur dan konsisten untuk menghasilkan analisis data bagi kebutuhan APS.

Moscoso-Zea, Paredes-Gualtor, dan Luján-Mora [4], penggunaan *data warehouse* di bidang pendidikan masih jauh di bawah industri lainnya. Penelitian tersebut mengungkapkan pendekatan *bottom-up* milik Kimball merupakan metodologi yang paling sering dipakai di bidang pendidikan dalam proses perancangan *data warehouse*. Cara kerja metodologi ini adalah dengan berfokus pada suatu bisnis area tertentu. Pendekatan ini cocok digunakan di bidang pendidikan karena data dalam institusi pendidikan terkadang tidak terintegrasi antar departemen dan unit. Selain itu, Moscoso-Zea, Paredes-Gualtor, dan Luján-Mora juga menilai ada kecenderungan institusi pendidikan untuk memakai *star schema* sebagai pendekatan implementasi *data warehouse*. *Star schema* merupakan skema dengan tabel fakta di tengah dan satu set tabel dimensi [2]. Skema ini populer digunakan karena sifatnya *denormalized* sehingga proses *query* menjadi lebih sederhana, namun kekurangannya adalah banyaknya redundansi data yang membutuhkan penyimpanan yang lebih besar [4]. Adanya kekurangan variasi dan implementasi dalam hal pengembangan *data warehouse* justru membuka peluang bagi industri pendidikan untuk melakukan eksplorasi yang lebih luas dan mendalam.

Di Indonesia, implementasi *data warehouse* dalam berbagai aspek lingkungan perguruan tinggi sudah mulai dilakukan. Seta, Wati, dan Isnainiyah [5] merancang *data warehouse* untuk perpustakaan kampus menggunakan metode Kimball dengan rancangan *snowflake schema* yang digunakan untuk membantu pihak perpustakaan kampus untuk mengambil keputusan. Berbeda dengan *star schema*, *snowflake schema* menghindari redundansi data dengan melakukan normalisasi terhadap tabel dimensi [2]. Sementara itu, Prasetyo, Soedijono, Amborowati [6] memanfaatkan *data warehouse* untuk perencanaan pemasaran perguruan tinggi dengan mengamati data calon mahasiswa baru. Laporan yang dihasilkan dari *data warehouse* digunakan untuk menentukan strategi promosi serta evaluasi pemasaran oleh pemimpin kampus.

Sehubungan dengan akreditasi, beberapa penelitian menunjukkan penggunaan *data warehouse* sebagai solusi yang dipilih untuk membantu perguruan tinggi dalam mengisi borang akreditasi. Windarto [7] menggunakan metodologi pengembangan *Business Life Cycle* dengan model *star schema* untuk membantu pengisian borang standar 3, yaitu mengenai data kemahasiswaan dan lulusan. Selain itu Windarto juga menggunakan *data warehouse* ini untuk mendukung pengambilan keputusan oleh manajemen fakultas, contohnya untuk mengetahui kualitas akademik baik dari mahasiswa maupun lulusannya. Irawati dan Kusumaningsih [8] juga menerapkan hal yang serupa yaitu memanfaatkan *data warehouse* sebagai pendukung dalam pengisian borang akreditasi standar 3. Irawati dan Kusumaningsih menggunakan *nine-step methodology* milik Kimball dalam perancangan *data warehouse* serta memakai skema bintang dengan alasan skema ini mudah untuk dipahami dibandingkan skema-skema lain yang ada.

II. TINJAUAN PUSTAKA

Data warehouse sudah menjadi bagian esensial dari proses pengambilan keputusan di banyak industri. Menurut

Dasawaty dan Waterkamp [9] merancang data warehouse menggunakan *star schema* untuk menjawab borang akreditasi standar 4, yaitu laporan yang berhubungan dengan dosen program studi termasuk jumlah dosen tetap dan dosen tidak tetap, aktivitas mengajar, serta prestasi dosen. Pada penelitian tersebut, Dasawaty dan Waterkamp berpendapat *data warehouse* juga meningkatkan keamanan data karena akses langsung ke data operasional menjadi berkurang, terutama dalam hal proses *query* data yang sering kali membebani sistem operasionalnya. Serupa dengan Irawati dan Kusumaningsih [8] dan Dasawaty dan Waterkamp [9], Sinaga dan Girsang [10] juga menggunakan *nine-step methodology* dan skema bintang untuk merancang *data warehouse* yang digunakan untuk menjawab borang akreditasi standar 3 dan 4 tentang mahasiswa dan alumni, serta Sumber Daya Manusia (SDM). Hasil analisis penelitian ini dilihat dalam bentuk *bar chart* dan *pie chart* untuk membandingkan hasil dari tahun ke tahun.

Pengembangan *data warehouse* berkaitan dengan kepentingan akreditasi memiliki kemiripan yang sejalan dengan pendapat Moscoso-Zea, Paredes-Gualtor, dan Luján-Mora [4] tentang popularitas penggunaan skema bintang dalam perancangan *data warehouse* di lingkup pendidikan. Pendekatan *bottom-up* juga terlihat menjadi opsi yang sesuai untuk pengembangan *data warehouse* yang berkaitan dengan akreditasi karena berfokus pada satu area bisnis tertentu, dalam hal ini standar akreditasi tertentu. Kriteria akreditasi yang menjadi fokus pada beberapa penelitian yang disebutkan adalah standar tentang mahasiswa dan lulusan, serta SDM prodi. Pada IAPS 4.0 standar 3 dan 4 diubah menjadi kriteria 2 dan 3 untuk mahasiswa dan SDM. Pada penelitian yang disebutkan belum ada yang membahas tentang pembuatan *data warehouse* untuk domain Penelitian dan PkM. Pada IAPS 4.0, domain tersebut berkaitan secara khusus dengan kriteria 6 dan 7, yaitu Penelitian dan PkM, namun memiliki implikasi juga terhadap kriteria yang lain.

III. LANDASAN TEORI

A. Akreditasi

Menurut Peraturan Menteri Nomor 32 Tahun 2016 [11], akreditasi disebut sebagai Sistem Penjaminan Mutu Eksternal yang merupakan bagian dari Sistem Penjaminan Mutu Pendidikan Tinggi. Pada Pasal 2 dijelaskan dua tujuan akreditasi yang utama yaitu menentukan kelayakan prodi/perguruan tinggi, serta menjamin mutu prodi/perguruan tinggi di bidang akademik maupun non akademik. Ada lima prinsip dasar dari akreditasi yaitu: independen, akurat, obyektif, transparan, dan akuntabel. Hasil akreditasi berupa status Terakreditasi atau Tidak Terakreditasi. Jika Terakreditasi, maka akan diberi peringkat sesuai dengan skor nilai akreditasi yang didapat yaitu Unggul, Baik Sekali, dan Baik.

IAPS 4.0 merupakan instrumen terbaru yang dikeluarkan oleh BAN-PT yang berlaku sejak 1 April 2019 sesuai

dengan Peraturan BAN-PT Nomor 2 Tahun 2019. Didalamnya ada beberapa pembaharuan, termasuk berbasis evaluasi diri, berorientasi pada *outputs* dan *outcomes*, spesifik terhadap jenis institusi, dan merupakan bagian integral dari CQI (*Continuous Quality Improvement*) [12]. Seperti yang sudah dijelaskan di bagian Latar Belakang, terdapat sembilan kriteria yang harus dipenuhi oleh prodi yang ditulis dalam bentuk LKPS dan LED. Pada LKPS, tabel-tabel yang dibutuhkan sudah disiapkan sehingga hanya perlu diisi. Narasi tentang data yang disebutkan di LKPS secara lebih lengkap dijelaskan pada dokumen LED. Evaluasi diri merupakan bagian yang penting dalam proses APS karena merupakan upaya prodi dalam mengumpulkan dan menganalisis data yang dapat disimpulkan dan digunakan oleh pihak manajemen untuk mengelola institusinya [13]. Proses penulisan LED yang lebih detail ketimbang LKPS memungkinkan adanya penambahan data baru untuk mendukung analisis yang lebih mendalam. Data dan analisis yang ditulis pada LKPS dan LED harus memiliki benang merah, sehingga data harus memiliki konsistensi dan akurasi.

B. Definisi Data Warehouse

Data warehouse merupakan sebuah repositori berisi data historis yang dirancang khusus untuk mendukung proses pembuatan keputusan. Data pada repositori ini berasal dari sistem operasional dan sumber eksternal lainnya melalui proses ekstraksi, transformasi, integrasi dan pembersihan [2], [14]. Tujuan utama dari *data warehouse* adalah memberikan informasi tentang performa sebuah organisasi secara keseluruhan dan mendukung pengambilan keputusan yang lebih efisien dan akurat [2], [4]. Untuk itu, *data warehouse* memiliki struktur yang berbeda dengan *database* operasional yang digunakan untuk menunjang sistem operasional. *Database* operasional merupakan sistem OLTP (*Online Transaction Processing*) yang lebih tepat untuk penggunaan transaksional sehari-hari [15]. Sistem ini tidak dapat dipakai untuk proses analisis karena berbagai alasan, antara lain data di dalamnya terlalu detail, tidak ada sejarah data, dan memiliki performa yang buruk ketika harus melakukan *query* yang kompleks [2]. Selain itu, data yang dimiliki oleh sebuah organisasi mungkin saja tersebar di beberapa sistem operasional sehingga analisis tidak dapat dilakukan secara menyeluruh. *Data warehouse* memiliki empat karakteristik utama yaitu [14], [16]:

- berorientasi pada subjek (berfokus pada suatu area subjek tertentu)
- terintegrasi (data diambil dari berbagai sumber yang digabung menjadi satu kesatuan)
- non-volatile (data tidak dapat diperbaharui, hanya dapat diakses)
- time varying (setiap data bersifat akurat pada suatu waktu tertentu, ada sejarah data)

C. Tipe Skema Data Warehouse

Model multidimensional merupakan dasar dari *data warehouse* yang melihat data dari berbagai pandangan atau dimensi [2]. Terdapat dua komponen penting dalam model ini yaitu *dimensions* dan *facts*. Dimensi merupakan perspektif, sedangkan *fact* biasanya berhubungan dengan nilai numerik yang disebut *measure*. Pada dimensi, terdapat hirarki yang dapat menentukan granularitas atau tingkat detail [17].

Selain tabel fakta dan dimensi, ada kalanya ketika *bridge table* juga menjadi salah satu komponen pada skema *data warehouse*. *Bridge table* digunakan sebagai solusi relasi *many-to-many* [2]. Tabel ini dapat digunakan antara dua dimensi maupun antara tabel fakta dan dimensi.

Skema yang dapat digunakan dalam perancangan *data warehouse* yang sekaligus mendukung pandangan multidimensional yaitu *star schema*, *snowflake schema*, *starflake schema*, dan *fact constellation/galaxy schema* [2], [15].

- *Star schema* merupakan skema yang bersifat *denormalized* dengan satu tabel *fact* di tengah dan satu set tabel dimensi [2]. Kesederhanaan dari skema ini membuat *query* menjadi mudah, namun redundansi data menjadi kelemahannya yang menyebabkan integritas data menjadi susah untuk dipertahankan [15].
- *Snowflake schema* melakukan normalisasi data sehingga terlihat struktur hirarki pada dimensi dipisah menjadi beberapa tabel [18]. Proses yang disebut sebagai '*snowflaking*' ini menghindari persoalan redundansi data sekaligus meningkatkan efisiensi penyimpanan [4], [15]. Kelemahan yang jelas pada skema ini adalah *query* yang kompleks dengan adanya proses *join*.
- *Fact constellation/galaxy schema* merupakan skema yang memiliki beberapa tabel *fact* yang berbagi tabel dimensi [2].
- *Starflake schema* merupakan kombinasi dari *star* dan *snowflake schema* dimana ada beberapa dimensi yang dinormalisasi dan ada yang tidak dinormalisasi [2].

D. Proses ETL

ETL (*Extract, Transformation and Loading*) merupakan proses yang krusial dalam pembuatan *data warehouse*. ETL merupakan proses iteratif yang perlu dilakukan secara berkala [18]. Proses pertama, *extract*, merupakan ekstraksi data dari sumber data yang berbeda. Data dapat diambil dari berbagai sumber, dari *database* operasional yang berbeda hingga *flat files*. Selanjutnya, data dibersihkan, dan ditransformasi supaya sesuai dengan skema *data warehouse* melalui proses *transformation* [18]. Beberapa proses yang biasanya dilakukan pada langkah transformasi termasuk melakukan agregasi, kalkulasi, dan menghilangkan

duplikasi data [15], [18]. Proses terakhir dari ETL adalah *loading* dimana data yang sudah ditransformasi dimasukkan ke dalam tabel dimensi dan *fact* yang terkait.

E. Desain Data Warehouse

Terdapat dua metode populer dalam perancangan *data warehouse* yaitu pendekatan *top-down* oleh Inmon dan pendekatan *bottom-up* oleh Kimball [4]. Pendekatan *top-down* mengumpulkan semua kebutuhan organisasi di awal untuk membuat sebuah *data warehouse* sentral yang selanjutnya akan dipakai untuk membuat *data mart* dependen [2], [3]. *Data mart* merupakan bagian dari *data warehouse* yang hanya berfokus pada sebuah bisnis area tertentu, disebut juga sebagai *departmental data warehouse* [2]. Pendekatan *bottom-up* menekankan pembuatan *data mart* untuk setiap area bisnis terlebih dahulu, sehingga *data warehouse* dibangun secara bertahap [4]. Mengingat kompleksitas data yang ada pada instansi pendidikan, Moscoso-Zea, Paredes-Gualtor, dan Luján-Mora berpendapat pendekatan ini yang lebih cocok diterapkan dalam pembuatan *data warehouse* pendidikan. Penggunaan pendekatan ini dapat menghasilkan *data mart* lebih cepat sehingga dapat menekan biaya dalam menghasilkan *report* yang dibutuhkan [2].

Terdapat beberapa pendapat tentang pendekatan yang dapat dipakai dalam pengembangan *data warehouse*. Vaisman dan Zimányi [2] berpendapat tidak ada persetujuan umum tentang tahapan yang harus diikuti dalam pengembangan *data warehouse*. Namun pada penelitian yang dilakukan oleh Moscoso-Zea, Paredes-Gualtor, dan Luján-Mora, disimpulkan lima tahapan yang disarankan untuk pengembangan *data warehouse* di ranah pendidikan, yaitu:

1. *Information needs analysis* dan *requirement analysis*
2. Analisis sumber data
3. Desain *data warehouse* dan *multidimensional modeling*
4. Proses ETL
5. Pengembangan sistem, aplikasi, *reporting*, *dashboard* dan OLAP

IV. METODE PENELITIAN

Pengembangan *data warehouse* untuk kebutuhan APS dilakukan dengan pendekatan *bottom-up*. *Data warehouse* akreditasi prodi dibagi menjadi beberapa domain utama yaitu: Penerimaan Mahasiswa Baru (PMB), Sumber Daya Manusia (SDM), Mahasiswa, Alumni, Penelitian dan PkM, Penunjang, dan Kegiatan. Pendekatan *bottom-up* membantu dalam pembuatan *data warehouse* akreditasi dengan cara berfokus pada setiap domain sebelum disatukan menjadi sebuah *data warehouse* global. Sesuai dengan Batasan Masalah, fokus penelitian ini adalah pada pengembangan *data mart* untuk domain Penelitian dan PkM Dosen. Mengingat istilah *data mart* juga dapat disebut sebagai *departmental data warehouse* maka istilah *data warehouse*

akan dipakai untuk menjelaskan *data mart* untuk domain tersebut.

Pengembangan *data warehouse* untuk Penelitian dan PkM Dosen akan dilakukan melalui lima tahapan yang disarankan oleh Moscoso-Zea, Paredes-Gualtor, dan Luján-Mora [4]:

1. **Analisis kebutuhan.** Hal ini diambil dari tabel pada LKPS yang berhubungan dengan Penelitian dan PkM Dosen. Selain itu, pihak prodi juga menambahkan beberapa analisis kebutuhan yang diperlukan untuk proses evaluasi diri pada LED.
2. **Analisis Sumber Data.** Sumber data diambil dari sistem internal kampus yang mencatat publikasi dari penelitian dan pengabdian dosen. Selain itu, ada juga data yang direkap secara manual oleh prodi.
3. **Desain data warehouse dan multidimensional modeling.** Pada tahapan ini, dilakukan desain *data warehouse* dengan menentukan tabel *fact* dan dimensi yang dibutuhkan.
4. **Proses ETL.** Proses ini akan dilakukan menggunakan Pentaho Data Integration Community Edition. *Data warehouse* akan disimpan pada *database* MariaDB.
5. **Reporting.** Pengambilan data untuk laporan akan dilakukan melalui Metabase. Metabase merupakan sebuah *Business Intelligence tool* yang dapat

digunakan untuk berbagai kebutuhan analisis, salah satunya adalah *reporting*. Laporan dapat disajikan dalam bentuk *dashboard* menggunakan berbagai macam *chart* yang tersedia, atau dapat diunduh dengan berbagai format seperti Microsoft Excel. Metabase dapat dikoneksikan ke *data warehouse* yang tersedia untuk menghasilkan laporan yang cepat untuk mengisi LKPS dan LED.

V. PEMBAHASAN

A. Analisis Kebutuhan

Domain Penelitian dan PkM Dosen secara khusus berkaitan dengan beberapa tabel pada LKPS. Tabel I merupakan daftar tabel pada LKPS yang diidentifikasi memiliki hubungan dengan domain Penelitian dan PkM berikut dengan kebutuhan analisis dan rentang waktu yang diharapkan. TS merupakan singkatan dari Tahun Semester yang meliputi tahun ajaran ganjil, genap, dan pendek.

Selain kebutuhan dari LKPS, pihak prodi juga menyimpulkan beberapa kebutuhan data yang akan digunakan sebagai evaluasi diri pada LED. Tabel II menunjukkan kebutuhan yang terkait dengan LED.

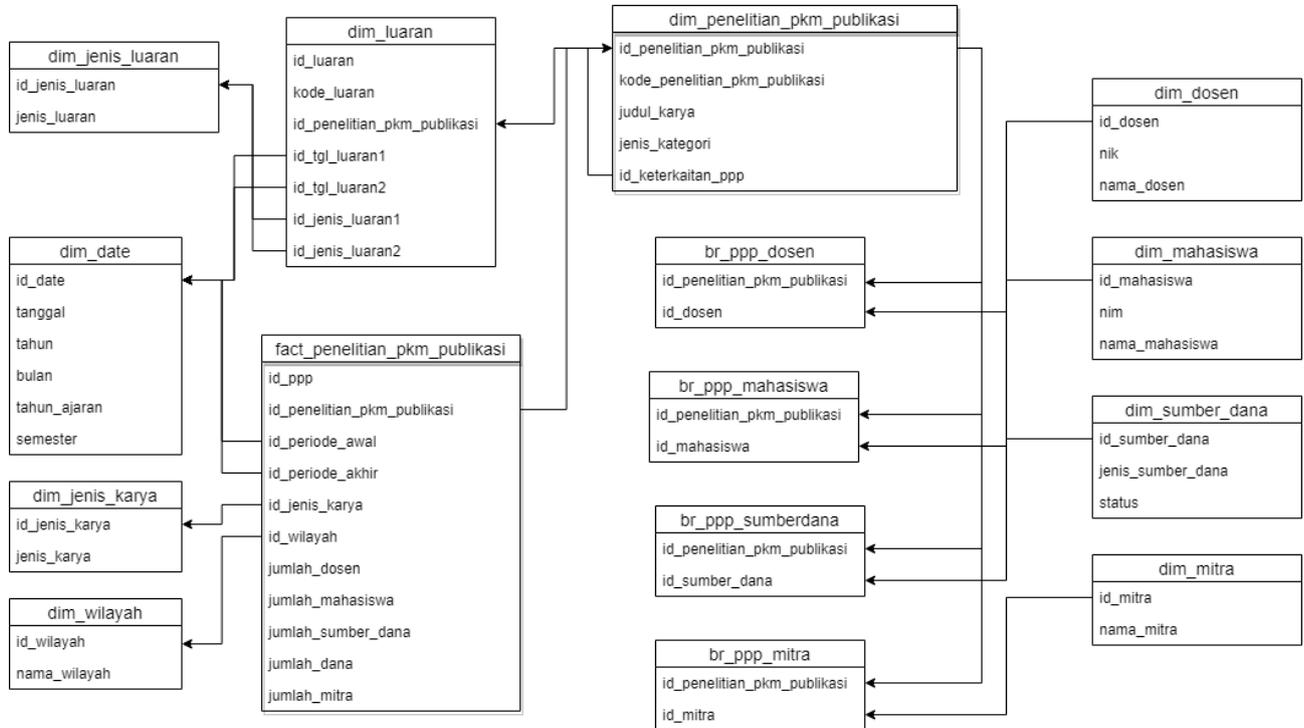
TABEL I
TABEL LKPS YANG BERKAITAN DENGAN PENELITIAN DAN PKM DOSEN

| Tabel pada LKPS | Analisis | Waktu |
|-----------------|---|------------------|
| 3.b.2 | Jumlah judul Penelitian DTPS berdasarkan Sumber Pembiayaan | TS hingga TS-2 |
| 3.b.3 | Jumlah judul PkM DTPS berdasarkan Sumber Pembiayaan | TS hingga TS-2 |
| 3.b.4 | Jumlah judul Publikasi Ilmiah DTPS berdasarkan Media Publikasi | TS hingga TS-2 |
| 3.b.7 | Luaran Penelitian/PkM Lainnya oleh DTPS dengan keterangan: Judul Luaran, Tahun, Keterangan | 3 tahun terakhir |
| 6.a | Penelitian DTPS yang melibatkan mahasiswa dengan keterangan Nama Dosen, Tema Penelitian sesuai Roadmap, Nama Mahasiswa, Judul Kegiatan, dan Tahun | TS hingga TS-2 |
| 7 | PkM DTPS yang melibatkan mahasiswa dengan keterangan Nama Dosen, Tema PkM sesuai Roadmap, Nama Mahasiswa, Judul Kegiatan, dan Tahun | TS hingga TS-2 |
| 8.f.1 | Jumlah judul Publikasi Ilmiah Mahasiswa berdasarkan Media Publikasi | TS hingga TS-2 |
| 8.f.4 | Luaran Penelitian/PkM Lainnya yang dihasilkan mahasiswa dengan keterangan: Judul Luaran Penelitian/PkM, Tahun, Keterangan | 3 tahun terakhir |

TABEL II
KEBUTUHAN DATA UNTUK PENGISIAN LED

| No. | Analisis | Waktu |
|-----|---|----------------|
| 1 | Jumlah dana penelitian setiap DTPS per TS | TS hingga TS-2 |
| 2 | Jumlah dana penelitian dari sumber internal setiap DTPS per TS | TS hingga TS-2 |
| 3 | Jumlah dana penelitian dari sumber eksternal setiap DTPS per TS | TS hingga TS-2 |
| 4 | Rata-rata dana penelitian per TS | TS hingga TS-2 |

| No. | Analisis | Waktu |
|-----|--|----------------|
| 5 | Rata-rata dana penelitian dari sumber internal per TS | TS hingga TS-2 |
| 6 | Rata-rata dana penelitian dari sumber eksternal per TS | TS hingga TS-2 |
| 7 | Jumlah DTSP yang terlibat dalam penelitian per TS | TS hingga TS-2 |
| 8 | Jumlah DTSP yang terlibat dalam PkM per TS | TS hingga TS-2 |
| 9 | Jumlah Mitra yang terlibat dalam penelitian | TS hingga TS-2 |
| 10 | Jumlah Mitra yang terlibat dalam PkM | TS hingga TS-2 |



Gambar 1. Snowflake Schema Penelitian dan PkM

B. Analisis Sumber Data

Saat ini, universitas menyediakan sebuah sistem informasi yang digunakan untuk mencatat publikasi yang dihasilkan oleh dosen. Setiap dosen dapat mencatat publikasi yang dihasilkan termasuk judul publikasi, dana, penelitian atau PkM yang terkait, serta anggota yang terlibat dalam publikasi tersebut. Namun hal ini tidak cukup untuk memenuhi kebutuhan analisis LKPS dan LED. Hal utama yang menjadi masalah adalah setiap dosen dapat memasukkan judul publikasi secara mandiri sehingga dapat menyebabkan inkonsistensi data apabila terdapat beberapa anggota dalam publikasi tersebut. Sistem ini menekankan pada publikasi saja dan tidak mencatat data terkait penelitian dan PkM secara lengkap. Untuk hal tersebut, pihak prodi melakukan rekap secara manual yang harus disesuaikan juga dengan data dari sistem internal. Untuk keperluan pembangunan *data warehouse*, data dari kedua sumber tersebut disatukan.

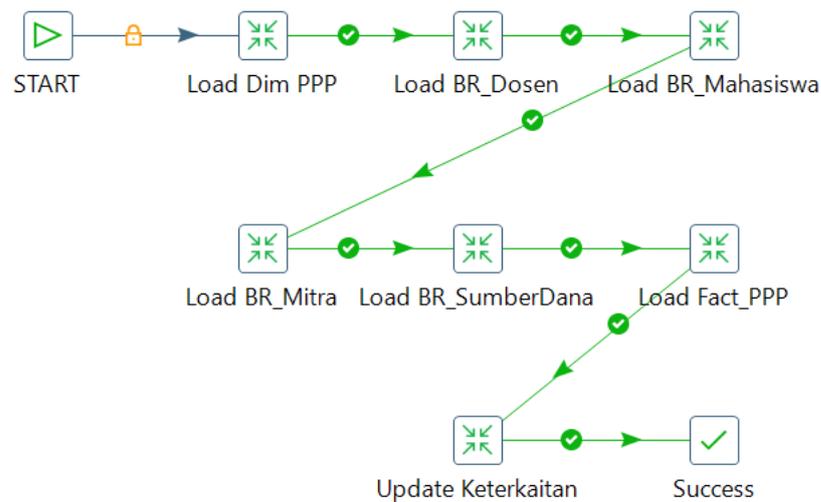
Data berhubungan dengan luaran harus direkap secara manual karena saat ini belum ada sistem internal yang mencatat luaran lainnya. Dalam pencatatannya, semua data

luaran dimasukkan, termasuk yang tidak berhubungan dengan penelitian maupun PkM.

C. Desain Data Warehouse dan Multidimensional Modeling

Berdasarkan kebutuhan analisis serta sumber data yang dimiliki maka ditentukan beberapa tabel dimensi yaitu:

1. **dim_date**
Dimensi waktu yang digunakan untuk menyimpan detail tanggal berikut dengan hirarkinya yaitu bulan dan tahun. Selain itu di sini terdapat informasi tentang tahun_ajaran dan semester.
2. **dim_mahasiswa**
Dimensi tentang keterangan mahasiswa.
3. **dim_dosen**
Dimensi tentang keterangan dosen.
4. **dim_mitra**
Dimensi tentang keterangan mitra.
5. **dim_wilayah**
Dimensi tentang wilayah, yang terdiri dari lokal, regional, nasional, dan internasional.
6. **dim_sumber_dana**



Gambar 2. Job untuk domain Penelitian dan PkM

Dimensi tentang keterangan sumber dana termasuk jenis sumber dana dari pihak internal maupun eksternal.

7. dim_jenis_karya

Dimensi ini menyimpan jenis karya publikasi untuk melakukan kategorisasi publikasi seperti jurnal internasional, jurnal nasional tidak terakreditasi, prosiding seminar nasional, dan sebagainya.

8. dim_jenis_luaran

Dimensi ini menyimpan jenis luaran seperti hak cipta, *book chapter*, dan sebagainya.

9. dim_penelitian_pkm_publicasi

Dimensi berisi data penelitian, PkM, dan publikasi yaitu kode, judul karya, dana, dan jenis kategori (penelitian/PkM/publikasi). Pada dimensi ini terdapat juga relasi *self-reference* khusus untuk jenis kategori publikasi dengan penelitian atau PkM yang terkait.

10. dim_luaran

Dimensi yang berisi semua luaran lainnya yang dihasilkan oleh prodi. Untuk luaran yang memiliki dimensi, terdapat relasi ke tabel *dim_penelitian_pkm_publicasi*. Satu baris pada dimensi ini menunjuk ke satu kelompok luaran yang memiliki beberapa jenis luaran. Contohnya, satu kelompok luaran bisa saja memiliki luaran jenis hak cipta dan buku saku.

Melihat data yang ada, maka dapat disimpulkan bahwa untuk setiap penelitian dan PkM melibatkan beberapa mahasiswa, dosen, mitra, dan sumber dana. Hal ini juga berlaku untuk publikasi. Untuk itu dibutuhkan beberapa *bridge table* untuk mengatasi relasi *many-to-many* ini.

1. *br_ppp_dosen*
Menjembatani antara tabel *dim_penelitian_pkm_publicasi* dengan tabel *dim_dosen*.
2. *br_ppp_mahasiswa*
Menjembatani antara tabel *dim_penelitian_pkm_publicasi* dengan tabel *dim_mahasiswa*.
3. *br_ppp_sumberdana*
Menjembatani antara tabel *dim_penelitian_pkm_publicasi* dengan tabel *dim_sumber_dana*.
4. *br_ppp_mitra*
Menjembatani antara tabel *dim_penelitian_pkm_publicasi* dengan tabel *dim_mitra*.

Untuk menjawab analisis yang dibutuhkan untuk LKPS dan LED, dibutuhkan sebuah skema *snowflake* dimana terdapat satu tabel *fact* dengan beberapa dimensi yang ternormalisasi, yaitu *fact_penelitian_pkm_publicasi*. Tabel fakta ini mencatat aktivitas penelitian, publikasi dan PkM yang dilihat dari *dim_time*, *dim_wilayah*. Pada tabel fakta ini juga terdapat beberapa *measure*: *jumlah_dosen*, *jumlah_mahasiswa*, *jumlah_sumber_dana*, dan *jumlah_dana*.

Gambar 1 menunjukkan skema *snowflake* untuk domain Penelitian dan PkM. Skema ini dapat menjawab kebutuhan yang disebutkan pada Tabel I dan sebagian Tabel II. Khusus untuk kebutuhan nomor 1-3 pada Tabel II terkait dana, jumlah dan rata-rata dana yang didapatkan dari data *warehouse* berorientasi pada Penelitian atau PkM yang terlibat, dan bukan setiap DTSP. Hal ini disebabkan oleh

sumber data yang merangkum dana berdasarkan Penelitian dan PkM.

D. Proses ETL

Sesuai dengan skema yang telah dirancang, proses ETL dilakukan menggunakan Pentaho Data Integration (*Community Edition*). *Database* yang digunakan untuk *data warehouse* adalah MariaDB ColumnStore. Proses ETL dilakukan secara iteratif untuk semua dimensi dan fakta yang berkaitan dengan domain Penelitian dan PkM.

Proses ETL pertama kali dilakukan untuk *dim_date* yang menyimpan keterangan tanggal secara lengkap termasuk bulan, tahun, nama hari, nama bulan, tahun ajaran dan semester. Pada Pentaho Data Integration (PDI) digunakan *Generate Rows* untuk mendapatkan tanggal secara otomatis, selanjutnya pemetaan tahun ajaran dan semester dilakukan menggunakan sumber data prodi yang telah memetakan tanggal mulai dan tanggal selesai untuk setiap kombinasi tahun ajaran dan semester. Selanjutnya proses ETL dilakukan untuk dimensi berikut ini: *dim_wilayah*, *dim_jenis_karya*, *dim_sumber_dana*, *dim_mitra*, *dim_dosen*, dan *dim_mahasiswa*. Sumber data untuk dimensi tersebut diambil dari sistem internal UKDW dan dimasukkan ke *data warehouse* menggunakan langkah *Combination Lookup/Update* dan *Update*. Langkah ini memastikan tidak ada duplikat data serta apabila di kemudian hari ada perubahan pada data, maka data akan diperbaharui.

Berikutnya proses ETL dilanjutkan untuk *dim_penelitian_pkm_publicasi*, *br_ppp_dosen*, *br_ppp_mahasiswa*, *br_ppp_sumberdana*, *br_ppp_mitra*, dan *fact_penelitian_pkm_publicasi*. Sumber data yang digunakan untuk tabel dimensi, fakta, dan *bridge* tersebut diambil dari data sejarah transaksi sistem publikasi internal UKDW. Sistem ini digunakan oleh dosen untuk mencatat publikasi serta keterangan yang terkait seperti judul publikasi, dana, penelitian atau PkM yang terkait, serta dosen atau mahasiswa yang terlibat. Persoalan yang muncul adalah publikasi yang sama dapat diisi oleh beberapa dosen yang terkadang mengakibatkan ketidakkonsistenan judul publikasi sehingga sukar untuk melakukan pengelompokan. Untuk itu sebelum proses ETL pada PDI, dilakukan pengecekan secara manual berdasarkan data sejarah transaksi dari sistem internal dengan surat perjanjian penelitian atau PkM yang bersangkutan serta rekapan daftar penelitian dan PkM milik prodi. Sumber data keenam tabel tersebut diambil dari hasil rekapan akhir ini.

Transformasi untuk *dim_penelitian_pkm_publicasi* dilakukan terlebih dahulu dengan menggunakan langkah *Combination Lookup/Update* dan *Update*. Selanjutnya dilakukan transformasi untuk semua tabel *bridge* yang meliputi langkah *Database Lookup* untuk mengambil *id_penelitian_pkm_publicasi* pada *dim_penelitian_pkm_publicasi* serta *id_dosen*, *id_mahasiswa*, *id_sumber_dana*, dan *id_mitra* dari tabel *dim_dosen*, *dim_mahasiswa*, *dim_sumber_dana*, dan *dim_mitra* secara

berurutan. Langkah ini krusial agar data konsisten dan memastikan bahwa data tersebut sudah tercatat di *data warehouse*. Berikutnya tabel *fact_penelitian_pkm_publicasi* dijalankan dengan sumber data yang sama, ditambah dengan *SQL query* dari tabel *dim_penelitian_pkm_publicasi* yang dilakukan JOIN ke tabel-tabel *bridge* untuk menghasilkan *measures* *jumlah_dosen*, *jumlah_mahasiswa*, *jumlah_sumber_dana*, *jumlah_dana*, dan *jumlah_mitra* yang digabungkan melalui langkah *Stream Lookup*. Yang terakhir, dilakukan transformasi untuk *update self-reference* pada *dim_penelitian_pkm_publicasi* untuk menentukan relasi publikasi dengan penelitian atau PkM yang terkait. Transformasi untuk keenam tabel tersebut dirangkum pada sebuah job yang menjalankan transformasi secara berurutan seperti yang terlihat pada Gambar 2.

Khusus untuk *dim_luaran*, sumber data yang digunakan pada proses ETL diambil dari rekapan manual karena saat ini belum ada sistem yang mencatat luaran penelitian dan PkM selain publikasi. Rekapan ini digunakan sebagai sumber data transformasi untuk *dim_luaran*. Proses transformasi data meliputi langkah *Database Lookup* untuk mengambil *id_penelitian_pkm_publicasi* dari *dim_publicasi_pkm_penelitian* yang berhubungan dengan luaran tersebut. Selain itu langkah *Database Lookup* juga dilakukan ke *dim_date* dan *dim_jenis_luaran*.

E. Reporting

Metabase digunakan untuk mengambil data dari *data warehouse*. Untuk pengisian LKPS dan LED, dibuatkan sebuah *dashboard* sederhana yang menampilkan data yang diperlukan menggunakan *query SQL*. Khusus untuk LKPS, data diperlihatkan dalam bentuk tabel menyesuaikan dengan tabel yang sudah disiapkan oleh BAN-PT. *Dashboard* tersebut dapat dilihat pada Gambar 3 untuk LKPS dan Gambar 4 untuk LED.

Penggunaan skema *snowflake* jelas menambahkan JOIN pada *query*, namun dengan adanya *bridge table*, semakin banyak JOIN yang harus dilakukan. Secara khusus pada tabel 6.a dan 7 pada LKPS tentang penelitian dan PkM yang melibatkan mahasiswa, digunakan CROSS JOIN untuk menghasilkan data mahasiswa dan dosen. Pemakaian *bridge table* mengakibatkan bertambahnya kompleksitas *query*.

VI. KESIMPULAN DAN SARAN

Sebuah *data warehouse* sebagai repositori terpusat yang berisi data tentang Penelitian dan PkM dosen berhasil dirancang. Keberadaan *data warehouse* membantu prodi dalam menghasilkan data yang konsisten dan terstruktur guna mengisi LKPS dan LED untuk IAPS 4.0. Perancangan *data warehouse* menggunakan skema *snowflake* dan metodologi perancangan untuk *data warehouse* pendidikan. Berdasarkan kebutuhan yang ditetapkan untuk LKPS dan LED, rancangan *data warehouse* ini dapat menjawab semua kebutuhan yang disebutkan di LKPS. Skema ini juga dapat menjawab tujuh dari sepuluh kebutuhan pada LED.

LKPS: Penelitian dan PkM

Domain Penelitian dan PkM



| 3.b.7 Luaran Penelitian/PkM Lainnya oleh DTSPS | | | | |
|--|------------------------|----------------------|---------------|---|
| Tahun Ajaran | Nama Dosen | Jenis Luaran | No. HKI | Judul Ciptaan |
| 2016/2017 | Lukas Chrisantyo A. A. | Teknologi Tepat Guna | - | Pembuatan Aplikasi Registrasi dan Event Management System dengan Metode Uji |
| 2016/2017 | Prihadi Benny Waluyo | Teknologi Tepat Guna | - | Pembuatan Aplikasi Registrasi dan Event Management System dengan Metode Uji |
| 2018/2019 | Antonius Rachmat C. | Hak Cipta | EC00201852527 | Algoritma dan Pemrograman dengan Bahasa C: konsep, teori, dan implementasi |

Rows 1-3 of 61

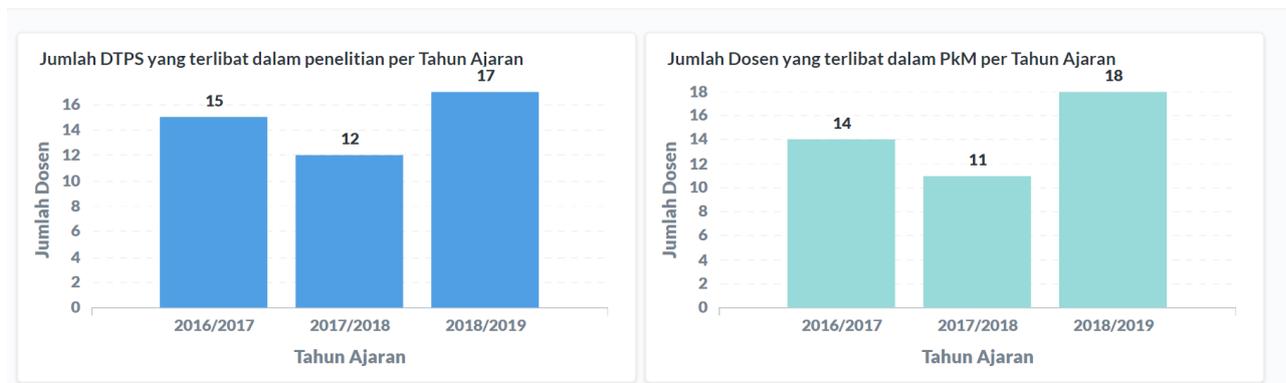
| 6.a Penelitian DTSPS yang melibatkan mahasiswa | | | |
|--|-----------------------------|------------------|---|
| Tahun Ajaran | Nama Dosen | Nama Mahasiswa | Judul Karya |
| 2016/2017 | Antonius Rachmat C. | JAYADI KURNIAWAN | KLASIFIKASI SENTIMEN MENGGUNAKAN ALGORITMA NAIVE BAYES PADA DATASET KOMEN |
| 2016/2017 | Yuan Lukito | JAYADI KURNIAWAN | KLASIFIKASI SENTIMEN MENGGUNAKAN ALGORITMA NAIVE BAYES PADA DATASET KOMEN |
| 2016/2017 | Aloysius Airlangga Bajuadji | INGGAR SAPUTRA | PERANCANGAN SISTEM INFORMASI KESISWAAN MENGGUNAKAN METODE SYSTEM DEVEL |

Rows 1-3 of 104

Gambar 3. Dashboard untuk Keperluan LKPS

LED: Penelitian dan PkM

Domain Penelitian dan PkM



Gambar 4. Dashboard untuk Keperluan LED

Meskipun demikian, rancangan ini masih dapat diperbaiki untuk mengoptimalkan analisis, seperti misalnya penambahan *measure* jumlah DTSPS yang aktif pada tahun ajaran yang bersangkutan sehingga dapat memberikan gambaran yang lebih jelas tentang prosentase keaktifan dosen dalam menjalankan Penelitian dan PkM.

UCAPAN TERIMA KASIH

Penulis berterimakasih kepada PUSPINDIKA UKDW, Kevin Kent, Bayu Manunggal Kristiawan, dan Thalia Maria Camilo yang telah membantu dalam pengumpulan data penelitian ini.

DAFTAR PUSTAKA

- [1] Badan Akreditasi Nasional Perguruan Tinggi, *Peraturan Badan Akreditasi Nasional Perguruan Tinggi Nomor 1 Tahun 2020 tentang Mekanisme Akreditasi untuk Akreditasi Yang Dilakukan oleh Badan Akreditasi Nasional Perguruan Tinggi*. 2020.
- [2] A. Vaisman dan E. Zimányi, *Data Warehouse Systems: Design and Implementation*. Berlin Heidelberg: Springer-Verlag, 2014.
- [3] P. Ponniah, *Data Warehousing Fundamentals for IT Professionals*, 2 edition. Wiley, 2011.
- [4] O. Moscoso-Zea, J. Paredes-Gualtor, dan S. Luján-Mora, "A Holistic View of Data Warehousing in Education," *IEEE Access*, vol. 6, p. 64659-64673, Okt 2018, doi: 10.1109/ACCESS.2018.2876753.
- [5] H. B. Seta, T. Wati, dan I. N. Isnainiyah, "Data Warehouse Development For UPN 'Veteran' Jakarta Library (Perancangan Data Warehouse pada Perpustakaan UPN 'Veteran' Jakarta)," *Pekommas*, vol. 2, no. 2, p. 161-166, Okt 2017, doi: 10.30818/jpkm.2017.2020206.
- [6] A. Prasetyo, W.B. Soedijono, dan A. Amborowati, "Perancangan Data Warehouse untuk Mendukung Perencanaan Pemasaran Perguruan Tinggi," *Telematika*, vol. 10, no. 1, p. 1-22, Feb 2017, doi: 10.35671/telematika.v10i1.500.
- [7] W. Windarto, "Pemanfaatan Data Warehouse sebagai Sarana Penunjang Penyusunan Borang Akreditasi Standar 3 pada Fakultas Teknologi Informasi Universitas Budi Luhur," *Telematika MKOM*, vol. 3, no. 2, p. 1-11, Agu 2016.
- [8] Y. Irawati dan D. Kusumaningsih, "Analisa dan Perancangan Data Warehouse untuk Menunjang Borang Akreditasi Standar 3 pada Fakultas Ilmu Komunikasi Universitas Budi Luhur," *Telematika MKOM*, vol. 4, no. 2, p. 173-182, Jun 2016.
- [9] E. S. Dasawaty dan J. Waterkamp, "Data Warehouse Sebagai Sarana Penunjang Penyusunan Borang Akreditasi Standar 4 Pada Program Studi Sistem Informasi Institut Bisnis dan Informatika Kwik Kian Gie," *J. Inform. Dan Bisnis*, vol. 2, no. 1, p. 36-49, 2013.
- [10] A. S. Sinaga dan A. S. Girsang, "University Accreditation using Data Warehouse," *J. Phys. Conf. Ser.*, vol. 801, p. 012030, Jan 2017, doi: 10.1088/1742-6596/801/1/012030.
- [11] Menteri Riset, Teknologi, Dan Pendidikan Tinggi, *Peraturan Menteri Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia Nomor 32 Tahun 2016 tentang Akreditasi Program Studi dan Perguruan Tinggi*. 2016.
- [12] Tim Penyusun Instrumen IAPT 3.0 dan IAPS 4.0, "Instrumen Akreditasi Perguruan Tinggi: Perkembangan Terkini Akreditasi Perguruan Tinggi." Badan Akreditasi Nasional Perguruan Tinggi (BAN-PT), Mei 27, 2019, [Online]. Tersedia: <https://ldikti8.ristekdikti.go.id/wp-content/uploads/2019/05/1.-Sosialisasi-APS-4.0-BAN-PT-INSTRUMEN.pdf>.
- [13] Tim Penyusun Instrumen IAPT 3.0 dan IAPS 4.0, "Instrumen Akreditasi Program Studi 4.0: Penyusunan Laporan Evaluasi Diri." Badan Akreditasi Nasional Perguruan Tinggi (BAN-PT), Mei 27, 2019, [Online]. Tersedia: https://ldikti8.ristekdikti.go.id/wp-content/uploads/2019/05/4.-Sosialisasi-APS-4.0-BAN-PT-PENYUSUNAN_LED.pdf.
- [14] G. S. Reddy, R. Srinivasu, M. P. C. Rao, dan S. Rikkula, "Data Warehousing, Data Mining, OLAP and OLTP Technologies are Essential Elements to Support Decision-Making Process in Industries," *International Journal on Computer Science and Engineering*, vol. 2, no. 9, p. 2865-2873, Des 2010.
- [15] P. Bhatia, *Data Mining and Data Warehousing: Principles and Practical Techniques*, 1 edition. Cambridge, United Kingdom ; New York, NY: Cambridge University Press, 2019.
- [16] W. H. Inmon, *Building the Data Warehouse Fourth Edition*, 4 ed. Indianapolis, Ind: Wiley, 2005.
- [17] O. B. Romero, "A Survey of Multidimensional Modeling Methodologies," *Int. J. Data Warehous. Min.*, Diakses: Jun 01, 2020. [Online]. Tersedia: https://www.academia.edu/10459695/A_Survey_of_Multidimensional_Modeling_Methodologies.
- [18] W. Lemahieu, S. vanden Broucke, dan B. Baesens, *Principles of Database Management: The Practical Guide to Storing, Managing and Analyzing Big and Small Data*, 1 edition. New York, NY: Cambridge University Press, 2018.