

Penerapan Database Life Cycle Dalam Merancang Basis Data Relasional Pengolahan Data Teks

Nofiyani^{1*}, Yesi Puspita Dewi²

^{1,2}Fakultas Teknologi Informasi, Teknik Informatika, Universitas Budi Luhur, Jakarta, Indonesia
Jl. Ciledug Raya, Petukangan Utara, Jakarta Selatan, 12260, DKI Jakarta
E-mail: ^{1*}nofiyani@budiluhur.ac.id, ²yesi.puspitadewi@budiluhur.ac.id
(*: corresponding author)

Abstrak— Web adalah sumber utama informasi berbasis teks yang tersedia untuk kita akses, dengan beragam kejadian maupun kegiatan yang dijadikan berita atau opini sebagai sumber informasi. Pengelompokan berita mampu memberikan dukungan untuk memudahkan pengambilan informasi dari berita tersebut. Basis data relasional menjadi salah satu solusi utama untuk mengorganisasi dan mengelola data secara efektif. Salah satu metode yang dapat digunakan untuk membangun basis data relasional adalah pendekatan Database Life Cycle (DLC). Perancangan basis data relasional terdiri dari Conceptual Database Design, Logical Database Design dan Physical Database Design. Dalam penelitian ini menghasilkan 8 (delapan) tipe entitas konseptual, menghasilkan diagram hubungan entitas dari kedelapan entitas tersebut dan menghasilkan rancangan pisikal antara lain tdata, tentry, tkategori, tpreprocessing, tslangword, tprobabilitas_kata, tprobabilitas_kategori dan tclassify.

Kata Kunci— Basis Data, Data Relasional, DBLC dan Berita

Abstract— The web is the main source of text-based information available for us to access, with various events and activities that are used as news or opinions as a source of information. News grouping can provide support to make it easier to retrieve information from the news. Relational databases are one of the main solutions for organizing and managing data effectively. One method that can be used to build a relational database is the Database Life Cycle (DLC) approach. Relational database design consists of Conceptual Database Design, Logical Database Design and Physical Database Design. This research produces 8 (eight) types of conceptual entities, produces entity relationship diagrams from these eight entities and produces physical designs including tdata, tentry, tcategory, tpreprocessing, tslangword, tprobabilities_words, tprobabilities_categories and tclassify.

Keyword—Databases, Relational Data, DBLC and News

I. PENDAHULUAN

Perkembangan teknologi informasi yang pesat telah mendorong kebutuhan untuk pengelolaan data yang lebih efisien dan terstruktur. Dalam konteks ini, basis data relasional menjadi salah satu solusi utama untuk mengorganisasi dan mengelola data secara efektif. Basis data relasional menyimpan informasi dalam bentuk tabel, di mana antara tabel satu dengan yang lainnya memiliki relasi. Dengan adanya basis data pengelolaan informasi jadi lebih mudah dan efisien serta

keakuratan dan keandalan data lebih terjamin [1]. Namun, untuk mencapai kinerja optimal dalam pengelolaan data, perancangan basis data relasional yang tepat sangatlah penting.

Model basis data relasional, jika diatur dan beroperasi dengan baik maka akan mampu memberikan informasi kepada semua afiliasi dan departemen perusahaan [2]. Salah satu metode yang dapat digunakan untuk membangun basis data relasional adalah pendekatan Database Life Cycle (DLC), dimana terdiri dari serangkaian tahapan yang mencakup perencanaan, analisis, desain, implementasi, pengujian, hingga pemeliharaan, secara sistematis dan terstruktur [3]. Sedangkan menurut Putri Swastika Sukmanasari System Development Life Cycle (SDLC) terdiri dari tahapan perencanaan, analisis, desain, pengembangan, pengujian, implementasi, dan pemeliharaan [4]. Setiap tahapan memiliki peran penting dalam memastikan bahwa basis data yang dibangun dapat memenuhi kebutuhan pengguna dan berjalan dengan efisien.

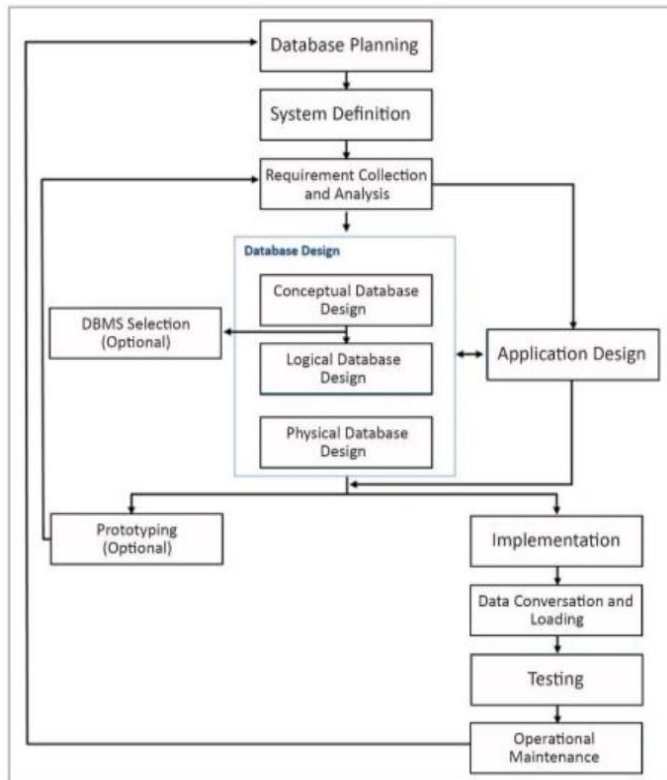
Web adalah sumber utama informasi berbasis teks yang tersedia untuk kita akses [5]. Begitu banyak kejadian maupun kegiatan yang dijadikan berita atau opini yang dipaparkan di website atau biasa di sebut berita online [6].

Berdasarkan uraian tersebut dibutuhkan database dalam pengelompokan berita untuk memudahkan pengambilan informasi dari berita tersebut. Melalui analisis dan implementasi model basis data relasional yang tepat, diharapkan dapat tercipta solusi pengelolaan data yang tidak hanya efisien, tetapi juga mempercepat dan mempermudah pengguna dalam mencari informasi yang dibutuhkan.

Untuk itu fokus penelitian ini adalah membangun basis data dengan menggunakan pendekatan *database life cycle*, dalam pengelompokan data teks berita.

II. METODE PENELITIAN

Metode pengembangan database yang digunakan oleh penulis adalah DBLC (siklus hidup basis data), yang menggambarkan tentang fase-fase dalam kehidupan sebuah database. Gambar 1 berikut menunjukkan Tahapan metode DBLC [7].



Gambar 1. Tahapan Metode DBLC

Berdasarkan gambar diatas terdapat tiga tahapan dalam desain basis data yang terdiri dari:

a. Conceptual Database Design

Pada tahapan ini adalah menentukan user yang akan terlibat, data apa saja yang perlu dimasukkan dalam sistem dan informasi apa yang ingin dihasilkan sistem.

b. Logical Database Design

Tahapan ini adalah mengidentifikasi atribut, domain, dan kunci kandidat untuk setiap entitas yang terbentuk berdasarkan identifikasi tipe entitas [8].

c. Physical Database Design

Tahapan ini adalah transformasi perancangan logis dalam bentuk fisik ke dalam DBMS terpilih [8].

III. HASIL DAN PEMBAHASAN

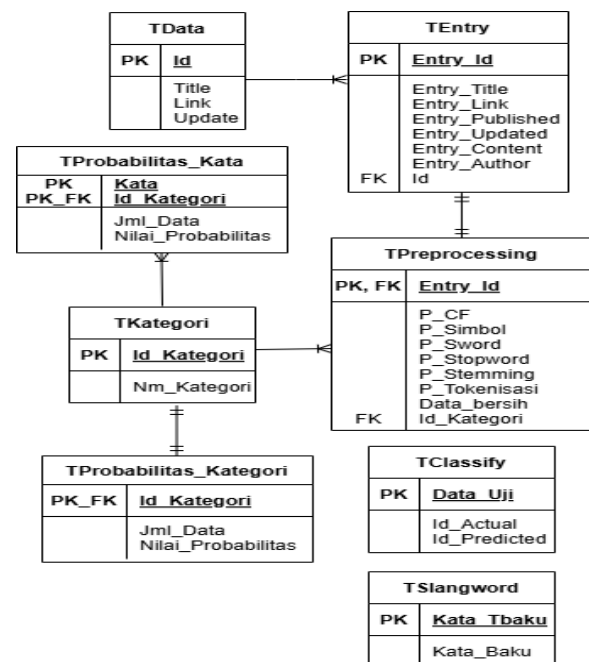
A. Conceptual Database Design

Conceptual database design adalah membuat perencanaan dalam bentuk konseptual yang digambarkan dalam identifikasi entitas dan diagram konseptual [9]. Tabel 1 dan Gambar 2 menunjukkan rancangan Conceptual database design.

TABEL I
IDENTIFIKASI TIPE ENTITAS

Nama Entity	Keterangan Entity	Kegiatan
TData	Berisi informasi mengenai kata kunci pencarian berita, lokasi dan kapan berita dibuat.	Pengelompokan data berita berdasarkan kata kunci pencarian dan satu kode pencarian bisa terdiri

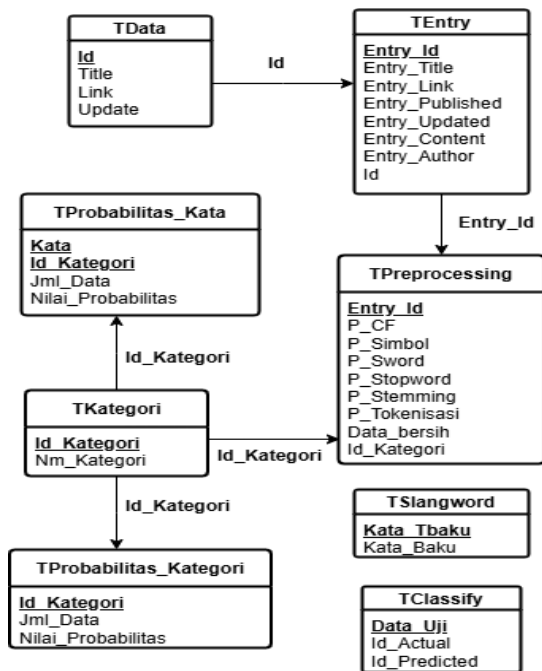
		dari beberapa berita dan terdiri dari satu kategori pencarian.
TEntry	berisi informasi mengenai detail berita.	Pengelompokan data berita berdasarkan kode kata kunci pencarian dan satu data berita mempunyai satu data bersih
TKategori	Berisi informasi kategori pengelompokan berita	Pengelompokan kategori berita berdasarkan kode kata kunci pencarian
TPreprocessing	Berisi informasi mengenai data berita yang sudah dilakukan proses pembersihan	Pengelompokan data bersih berdasarkan data entry dan satu data entry bisa terbentuk satu data bersih
TSlangword	Berisi kamus kata tidak baku dan kata baku	Dapat berisi daftar kata tidak baku dan kata baku yang digunakan dalam proses pembersihan data
TProbabilitas_Kata	Berisi informasi mengenai nilai probabilitas setiap kata data bersih	Pengelompokan kata berdasarkan data bersih dan satu kata terdiri dari beberapa kategori berita
TProbabilitas_Kategori	Berisi informasi mengenai nilai probabilitas setiap kategori berita	Pengelompokan nilai probabilitas berdasarkan kategori berita
TClassify	Berisi informasi mengenai data uji	Satu data uji terdiri dari satu kategori berita



Gambar 2. E-R Diagram Konseptual

B. Logical Database Design

Pada tahap logical database design dilakukan identifikasi atribut beserta penentuan *Primary key* dan *foreign key* pada setiap entitas. Penentuan key bertujuan untuk relasi antar entitas dan sebagai acuan untuk mendefinisikan sebuah baris data [10]. Logical database design dapat ditunjukkan pada gambar 3.



Gambar 3. Logical Record Structure (LRS)

C. Physical Database Design

Physical database design adalah transformasi dari logical database design terhadap DBMS yang dipilih sehingga dapat melakukan penyimpanan secara fisik. Xampp MySql merupakan DBMS yang dapat digunakan dalam pengembangan basis data pengelompokan berita dalam penelitian ini.

TABEL II
 SPESIFIKASI TABEL TDATA

Field	Type	Length
Id	Varchar	300
Title	Tinytext	255
Link	Tinytext	255
Update	Tinytext	255

TABEL III
 SPESIFIKASI TABEL TENTRY

Field	Type	Length
Entry_Id	Varchar	300
Entry_Title	Tinytext	255
Entry_Link	Tinytext	255
Entry_Published	Tinytext	255
Entry_Updated	Tinytext	255
Entry_Content	Tinytext	255

Entry_Author	Tinytext	255
Id	Tinytext	300
	Varchar	

TABEL IV
 SPESIFIKASI TABEL TKATEGORI

Field	Type	Length
Id_Kategori	Int	11
Nm_Kategori	Varchar	50

TABEL V
 SPESIFIKASI TABEL TPREPROCESSING

Field	Type	Length
Entry_Id	Varchar	300
P_CF	Tinytext	255
P_Symbol	Tinytext	255
P_Sword	Tinytext	255
P_Stopword	Tinytext	255
P_Stemming	Tinytext	255
P_Tokenisasi	Tinytext	255
Data_bersih	Tinytext	255
Id_Kategori	Int	11

TABEL VI
 SPESIFIKASI TABEL TSLANGWORD

Field	Type	Length
Kata_Tbaku	Varchar	50
Kata_Baku	Varchar	50

TABEL VII
 SPESIFIKASI TABEL TPROBABILITAS_KATA

Field	Type	Length
Kata	Varchar	50
Id_Kategori	Int	11
Jml_Data	Int	11
Nilai_Probabilitas	Float	10,10

TABEL VIII
 SPESIFIKASI TABEL TPROBABILITAS_KATEGORI

Field	Type	Length
Id_Kategori	Int	11
Jml_Data	Int	11
Nilai_Probailitas	Float	10,10

TABEL IX
 SPESIFIKASI TABEL TCLASSIFY

Field	Type	Length
Data_Uji	Tinytext	255
Id_Actual	Int	11
Id_Predicted	Int	11

Hasil perancangan basis data dalam penelitian ini terbentuk 8 tabel yang dibuat dengan aplikasi MySQL-Front. Berikut ini merupakan perintah-perintah SQL dalam mendefinisikan tabel-tabel dalam database yang dipilih.

1. SQL Tabel TData

```
CREATE TABLE TDATA (  
    id VARCHAR (300) NOT NULL DEFAULT "",  
    title TINYTEXT NULL DEFAULT NULL,  
    link TINYTEXT NULL DEFAULT NULL,  
    update TINYTEXT NULL DEFAULT NULL,  
    PRIMARY KEY (id)  
)
```

2. SQL Tabel TEntry

```
CREATE TABLE TENTRY (  
    entry_id VARCHAR (300) NOT NULL,  
    entry_title TINYTEXT NOT NULL,  
    entry_link TINYTEXT NOT NULL,  
    entry_published TINYTEXT NOT NULL,  
    entry_updated TINYTEXT NOT NULL,  
    entry_content TINYTEXT NOT NULL,  
    entry_author TINYTEXT NOT NULL,  
    id VARCHAR (300) NOT NULL DEFAULT "",  
    PRIMARY KEY (entry_id)  
)
```

3. SQL Tabel TKategori

```
CREATE TABLE TKATEGORI (  
    id_kategori int(11) NOT NULL auto_increment,  
    nm_kategori varchar (50) NOT NULL default "",  
    PRIMARY KEY (id_kategori)  
)
```

4. SQL Tabel TPreprocessing

```
CREATE TABLE TPREPROCESSING (  
    entry_id VARCHAR (300) NOT NULL,  
    p_cf TINYTEXT NOT NULL,  
    p_simbol TINYTEXT NOT NULL,  
    p_sword TINYTEXT NOT NULL,  
    p_stopword TINYTEXT NOT NULL,  
    p_stemming TINYTEXT NOT NULL,  
    p_tokenisasi TINYTEXT NOT NULL,  
    data_bersih TINYTEXT NOT NULL,  
    id_kategori int(11),  
    PRIMARY KEY (entry_id)  
)
```

5. SQL Tabel TSlangword

```
CREATE TABLE TSLANGWORD (  
    kata_tbaku VARCHAR (50) NOT NULL,  
    kata_baku VARCHAR (100) NOT NULL  
)
```

6. SQL Tabel TProbabilitas_Kata

```
CREATE TABLE TPROBABILITAS_KATA (  
    kata VARCHAR (50) NOT NULL,  
    id_kategori Int(11) NOT NULL,  
    jml_data Int(11) NULL DEFAULT 0,  
    nilai_probabilitas float (10,10) NULL DEFAULT 0,  
    PRIMARY KEY (kata, id_kategori)  
)
```

7. SQL Tabel TProbabilitas_Kategori

```
CREATE TABLE TPROBABILITAS_KATEGORI (  
    id_kategori Int(11) NOT NULL,  
    jml_data Int(11) NULL DEFAULT 0,  
    nilai_probabilitas float (10,10) NULL DEFAULT 0,  
    PRIMARY KEY (id_kategori)  
)
```

8. SQL Tabel TClassify

```
CREATE TABLE TCLASSIFY (  
    data_uji TINYTEXT NOT NULL,  
    id_actual Int(11) NULL,  
    id_predicted Int(11) NULL  
)
```

IV. KESIMPULAN

Metode Database Life Cycle (DLC) digunakan dalam merancang basis data dan telah menghasilkan database relasional. Entitas yang dibutuhkan berdasarkan identifikasi tipe entitas terdiri dari 8 entitas yang saling berhubungan dalam diagram entitas. Menghasilkan hubungan model data logis yang mewakili entitas, hubungan, dan atribut yang telah diidentifikasi sebelumnya. Xampp MySql digunakan dalam perancangan Physical Database Design dengan jumlah tabel yang terbentuk 8 tabel.

REFERENSI

- [1] M. I. Ramadan, "Pengantar Basis Data," 2023. [Online]. Available: <https://www.researchgate.net/publication/368080171>
- [2] S. G. Rezeki, M. Irwan, and P. Nasution, "Peranan Penggunaan Basis Data dalam Sistem Informasi Manajemen," *IJM: Indonesian Journal of Multidisciplinary*, vol. 1, no. 4, pp. 1243-1251, 2023.
- [3] O. A. Arrohman and R. P. Wibawa, "Analisis Efisiensi Aplikasi Administrasi 'Jembul Smart Village' Dalam Mempercepat Pembuatan Surat Digital," *Jurnal IT-Edu*, vol. 09, no. 03, pp. 31-39, 2024.
- [4] P. S. Sukmanasari, Marsono, F. Alqodri, and A. P. P. Triono, "Desain dan Implementasi Aplikasi SISPEDA (Sistem Informasi Pengajaran Dana) sebagai Dasar Ajuan Uang Muka Kerja (UMK) Rutin dan Kegiatan pada Fakultas Teknik Universitas Negeri Malang Berorientasi Industri 5.0 untuk Meningkatkan Akuntabilitas Keuangan," *3rd SEMINASTIKA 2024*, pp. 1-7.
- [5] A. Firdaus and W. I. Firdaus, "Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi : (Sebuah Ulasan)," *JUPITER: Jurnal Penelitian Ilmu dan Teknologi Komputer*, vol. 13, no. 1, pp. 66-78, 2021.
- [6] C. N. Daiman, A. Y. Rahman, and F. Nudiyansyah, "Klasifikasi Teks Berita Breaking News Di Manggarai Menggunakan Long Short Term Memory (LSTM)," *Jurnal Mnemonic*, vol. 7, no. 2, pp. 170-174, 2024.
- [7] A. Widyanto, K. Khotop, H. Effendi, A. Pratama, EH., "Desain Model Database E-Commerce SLB Pembina Palembang dengan Menerapkan Metode Database Life Cycle," *Seminar Nasional Penelitian dan Pengabdian Kepada Masyarakat CORISINDO*, Institute Teknologi dan Bisnis STIKOM Bali, 2022, pp. 413-425.
- [8] S. Samidi and R. Hidayat, "Desain Model Database Mutasi Siswa Dengan Menerapkan Metode Database Life Cycle," *Technomedia J.*, vol. 8, no. 2, pp. 221-235, 2023.
- [9] Y. Bahtiar and D. Herwanto, "Perancangan Basis Data Penjualan Dengan Metode Database Lifecycle Pada Toko Lancar Elektrik," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, vol. 7, no. 2, pp. 169-176, 2022.
- [10] Y. B. Utomo, I. Kurniasari, and I. Yanuartanti, "Penerapan Knowledge Discovery In Database Untuk Analisa Tingkat Kecelakaan Lalu Lintas," *J. Tek. Inform. Kaputama*, vol. 7, no. 1, pp. 171-180, 2023.