

**RANCANG BANGUN *DATA MART* IDENTITAS NASABAH
DARI DATA TIDAK TERSTRUKTUR PT XYZ**



SKRIPSI

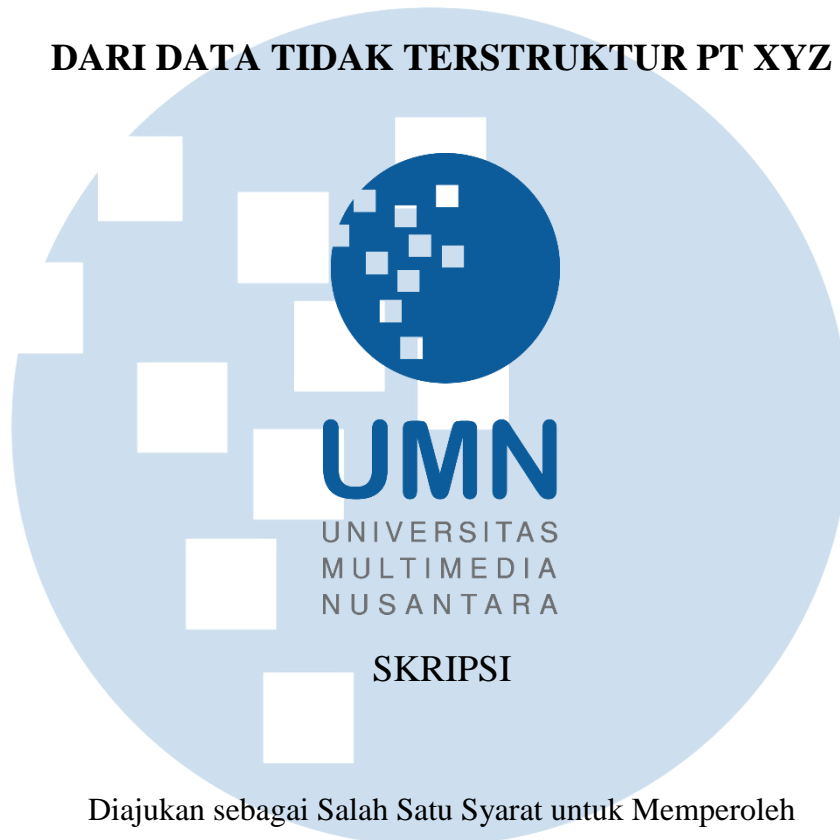
Angelica Olivia Anthony

000000050653

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2024

**RANCANG BANGUN DATA MART IDENTITAS NASABAH
DARI DATA TIDAK TERSTRUKTUR PT XYZ**



SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Komputer (S.Kom)

Angelica Olivia Anthony

0000050653

PROGRAM STUDI SISTEM INFORMASI

TEKNIK DAN INFORMATIKA

UNIVERSITAS MULTIMEDIA NUSANTARA

TANGERANG

2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Angelica Olivia Anthony

Nomor Induk Mahasiswa : 00000050653

Program studi : Sistem Informasi

Skripsi dengan judul:

RANCANG BANGUN *DATA MART* IDENTITAS NASABAH DARI DATA TIDAK TERSTRUKTUR PT XYZ

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan skripsi, saya bersedia menerima konsekuensi dinyatakan **TIDAK LULUS** untuk Tugas Akhir yang telah saya tempuh.

Tangerang, 15 Mei 2024



A handwritten signature in black ink, appearing to read 'Angelica', is written over the stamp.

Angelica Olivia Anthony

UMM
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PERSETUJUAN

Skripsi dengan judul

RANCANG BANGUN *DATA MART* IDENTITAS NASABAH DARI DATA TIDAK TERSTRUKTUR PT XYZ

Oleh

Nama : Angelica Olivia Anthony

NIM : 00000050653

Program Studi : Sistem Informasi


Fakultas : Teknik dan Informatika

Telah disetujui untuk diajukan pada

Sidang Ujian Skripsi Universitas Multimedia Nusantara

Tangerang, 15 Mei 2024

Pembimbing 1


15/05/2024
Suryasari, S.Kom., M.T.
0323088301

Pembimbing 2


Iwan Prasetyawan, S.Kom., M.M.

Ketua Prodi Sistem Informasi


15.05.2024
Ririn Ikana Desanti, S.Kom., M.Kom.

UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

RANCANG BANGUN DATA MART IDENTITAS NASABAH DARI DATA TIDAK TERSTRUKTUR PT XYZ

Oleh

Nama : Angelica Olivia Anthony
NIM : 00000050653
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

Telah diujikan pada hari Senin, 20 Mei 2024
Pukul 14.00 s.d 16.00 dan dinyatakan
LULUS
Dengan susunan pengujian sebagai berikut.

Ketua Sidang



3/6/24

Ririn Ikana Desanti, S.Kom., M.Kom.
313058001

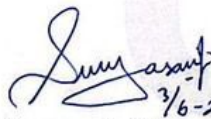
Penguji



03/05/24
ak

Dinar Ajeng Kristiyanti, S.Kom, M.Kom.
330128801

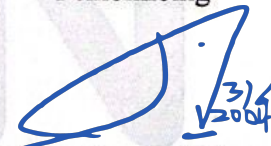
Pembimbing



3/6-2024.

Suryasari, S.Kom., M.T.
323088301

Pembimbing



3/6/2024

Iwan Prasetiawan, S.Kom., M.M.

Ketua Program Studi Sistem Informasi



3/6/24

Ririn Ikana Desanti, S.Kom., M.Kom

iv

LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHASISWA

Yang bertanda tangan dibawah ini:

Nama : Angelica Olivia Anthony

Nomor Induk Mahasiswa : 00000050653

Program Studi : Sistem Informasi

Jenjang : S1

Judul Karya Ilmiah :

RANCANG BANGUN DATA MART IDENTITAS NASABAH DARI DATA TIDAK TERSTRUKTUR PT XYZ

Menyatakan dengan sesungguhnya bahwa saya bersedia:

Memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya di repositori Knowledge Center, sehingga dapat diakses oleh Civitas Akademika/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial dan saya juga tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.

Saya tidak bersedia, dikarenakan:

Dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)*.

Tangerang, 20 Mei 2024



Angelica Olivia Anthony

* Jika tidak bisa membuktikan LoA jurnal/HKI selama 6 bulan kedepan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

KATA PENGANTAR

Puji Syukur atas selesainya penulisan skripsi dengan judul: “Rancang Bangun *Data Mart* Identitas Nasabah dari Data Tidak Terstruktur PT XYZ” ini yang dilakukan untuk memenuhi syarat untuk mencapai gelar Sarjana Jurusan Sistem Informasi Pada Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan tugas akhir ini, sangatlah sulit bagi saya untuk menyelesaikan tugas akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada

1. Bapak Dr. Ninok Leksono, M.A., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ibu Suryasari, S.Kom., M.T., selaku Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Bapak Iwan Prasetiawan, S.Kom., M.M, selaku Pembimbing kedua yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
6. PT XYZ, sebagai Perusahaan yang telah memberikan kesempatan saya untuk melakukan penelitian di perusahaan ini.
7. Bapak *Head of IT* Data Arsitek PT XYZ, yang telah membantu memberikan informasi dan arahan dalam menyusun skripsi ini.
8. Bapak *IT Technology Architecture* PT XYZ, yang telah membantu memberikan informasi dan arahan dalam menyusun skripsi ini
9. Steven Nathaniel, Vincentius Adrian dan Monica Auwdya, yang telah memberikan dukungan sehingga penulis dapat menyelesaikan skripsi ini.

10. Keluarga, yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.

Semoga skripsi ini dapat bermanfaat bagi para pembaca baik sebagai sumber informasi maupun sumber inspirasi.

Tangerang, 15 Mei 2024



Angelica Olivia Anthony

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

RANCANG BANGUN DATA MART IDENTITAS NASABAH DARI DATA TIDAK TERSTRUKTUR PT XYZ

Angelica Olivia Anthony

ABSTRAK

Data warehouse merupakan tempat penyimpanan data yang lengkap dan konsisten dengan karakteristik berorientasi subjek, terintegrasi, tidak volatil, dan bervariasi waktu yang berguna untuk mendukung pengambilan keputusan. Sebagai perbandingan, *Data mart* merupakan tipe *data warehouse* yang mengakomodasi kebutuhan data spesifik dari departemen atau fungsi bisnis tertentu. Perusahaan *Multifinance*, PT XYZ saat ini memiliki tantangan dalam mengolah dan mengintegrasikan data tidak terstruktur berupa foto identitas, sehingga pembangunan sebuah data mart identitas nasabah dapat menjadi solusi untuk menstrukturisasi data tersebut. *Data Mart* Identitas Nasabah dibangun dengan menggunakan metode *bottom-up* dan proses ETL dengan memanfaatkan sistem *cloud* AWS. Dengan menggunakan sejumlah *software* AWS seperti AWS S3, Textract, Amazon RDS, dan AWS Lambda, data tidak terstruktur dari foto-foto identitas dapat diolah dan disimpan dalam struktur yang terorganisir di dalam database Amazon RDS. Melalui integrasi teknologi *cloud*, penelitian ini bertujuan untuk meningkatkan efisiensi dan skalabilitas dalam pengelolaan data identitas nasabah. Hasil dari penelitian ini, sebanyak 317 data dari 350 data berhasil dimasukkan ke dalam sistem *data mart* identitas. Hal ini dikarenakan tidak terbacanya teks pada gambar. Dihasilkan juga *view* yang dapat dimodifikasi untuk mempermudah perusahaan dalam memahami nasabahnya seperti daerah mana yang memiliki nasabah terbanyak sehingga mendukung proses pengambilan keputusan bagi pihak *Top Management*.

Kata kunci: Amazon RDS, AWS Lambda, AWS S3, *Data Mart* Identitas Nasabah, Textract



RANCANG BANGUN DATA MART IDENTITAS NASABAH DARI DATA TIDAK TERSTRUKTUR PT XYZ

Angelica Olivia Anthony

ABSTRACT (English)

A data warehouse is a complete and consistent data storage system characterized by being subject-oriented, integrated, non-volatile, and time-variant, which is useful for supporting decision-making. In comparison, a data mart is a type of data warehouse that accommodates the specific data needs of a particular department or business function. Multifinance company PT XYZ currently faces challenges in processing and integrating unstructured data in the form of identity photos, so the development of a customer identity data mart can be a solution to structure this data. The Customer Identity Data Mart is built using a bottom-up method and the ETL process, leveraging the AWS cloud system. By using several AWS software tools such as AWS S3, Textract, Amazon RDS, and AWS Lambda, unstructured data from identity photos can be processed and stored in an organized structure within the Amazon RDS database. Through the integration of cloud technology, this research aims to improve efficiency and scalability in managing customer identity data. As a result of this research, 317 out of 350 data entries were successfully inserted into the identity data mart system. This was due to the unreadability of text in some images. Additionally, modifiable views were generated to help the company better understand its customers, such as identifying regions with the highest number of customers, thereby supporting the decision-making process for Top Management.

Keywords: Amazon RDS, AWS Lambda, AWS S3, Data Mart Customer Identity, Textract



DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHASISWA v	
KATA PENGANTAR	vi
ABSTRAK	viii
<i>ABSTRACT (English)</i>	ix
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan dan Manfaat Penelitian	5
1.4.1 Tujuan Penelitian	5
1.4.2 Manfaat Penelitian	5
1.5 Sistematika Penulisan	6
BAB II LANDASAN TEORI	8
2.1 Penelitian Terdahulu	8
2.2 Teori tentang Topik Skripsi	12
2.2.1 Industri <i>Multifinance</i>	12
2.2.2 Data Tidak Terstruktur	13
2.2.3 <i>Data Mart</i>	13
2.2.4 <i>Extract, Transform, Loading (ETL)</i>	15
2.2.5 <i>Entity Relationship Diagram (ERD)</i>	16
2.2.6 Model Konseptual <i>Data Warehouse</i>	18
2.2.7 <i>Star Schema</i>	19
2.2.8 <i>Optical character recognition (OCR)</i>	20
2.3 Teori tentang Framework / Algoritma yang digunakan	21

2.3.1	Texttract.....	21
2.3.2	Cloud Computing	21
2.4	Teori tentang Tools / Software yang digunakan.....	22
2.4.1	Amazon Web Services (AWS)	22
2.4.2	Amazon S3.....	23
2.4.3	AWS Lambda.....	23
2.4.4	Amazon RDS	24
2.4.5	PostgreSQL.....	24
2.4.6	Amazon EC2.....	25
BAB III METODOLOGI PENELITIAN		27
3.1	Gambaran Umum Objek Penelitian.....	27
3.1.1	PT XYZ.....	27
3.1.2	KTP	28
3.1.3	NPWP.....	29
3.1.4	SIM.....	29
3.1.5	JSON File.....	29
3.2	Metode Penelitian	30
3.2.1	Alur Penelitian.....	30
3.2.2	Metode Pengembangan Sistem	31
3.2.2.1	Metode Kimball (<i>Bottom-up</i>).....	31
3.3	Desain Arsitektur Data Mart.....	33
3.2.2.2	Desain Arsitektur Sistem	45
3.3	Teknik Pengumpulan Data.....	48
3.3.1	Studi Literatur.....	48
3.3.2	Wawancara	49
BAB IV ANALISIS DAN HASIL PENELITIAN		50
4.1	Analisa Masalah dan Kebutuhan Penelitian.....	50
4.1.2	Proses Bisnis dan Sistem pada PT XYZ.....	50
4.2	Konfigurasi Sistem	52
4.2.1	Konfigurasi S3.....	52
4.2.2	Konfigurasi Lambda.....	52
4.2.2.1	Lambda 'TransformasiData'	52

4.2.2.2	Lambda ‘testskripsi’	53
4.2.3	Konfigurasi VPC	54
4.2.4	Konfigurasi RDS	54
4.2.5	Rancangan ERD Data Mart	55
4.2.6	Hasil Transformasi Data	59
4.2.6.1	Tabel Dimensi ‘KTP’	59
4.2.6.2	Tabel Dimensi ‘SIM’	60
4.2.6.3	Tabel Dimensi ‘NPWP’	60
4.2.6.4	Tabel Dimensi ‘Invoice’	61
4.2.6.5	Tabel Fakta ‘Data’	62
4.2.7	View Data Mart	62
4.3	Hasil dan diskusi	65
BAB V SIMPULAN DAN SARAN		67
5.1	Simpulan	67
5.2	Saran	68
DAFTAR PUSTAKA		69
LAMPIRAN		74

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR TABEL

Tabel 2 1 Penelitian terdahulu.....	8
Tabel 2 2 tabel komponen ERD	17
Tabel 2 3 tabel relasi ERD	18
Tabel 3.1 Perbedaan metode	32
Tabel 3.2 Logika KTP.....	37
Tabel 3.3 algoritma tanggal_lahir dan tempat_lahir	37
Tabel 3.4 algoritma berlaku_hingga	38
Tabel 3.5 Logika SIM	40
Tabel 3.6 algoritma mengambil no_sim.....	41
Tabel 3.7 Logika NPWP	42
Tabel 3.8 Ekstrasi JSON	43
Tabel 3.9 Perhitungan angsuran_perbulan	44
Tabel 3.10 Proses pemanggilan lambda testskripsi untuk load.....	44
Tabel 4.1 Tabel dim_ktp	56
Tabel 4.2 Tabel dim_npwp	56
Tabel 4.3 Deskripsi tabel no_sim.....	57
Tabel 4.4 Deskripsi tabel invoice.....	58
Tabel 4.5 Deskripsi tabel data.....	58
Tabel 4.6 Hasil dan diskusi	65

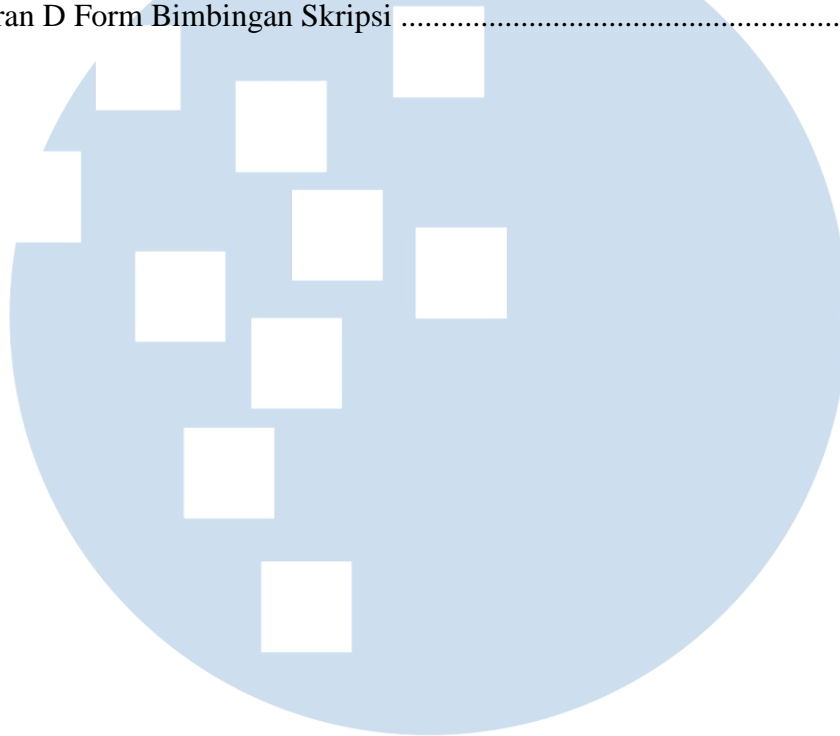


DAFTAR GAMBAR

Gambar 2. 1 Data Mart.....	14
Gambar 2. 2 ETL.....	15
Gambar 2. 3 Star Schema.....	20
Gambar 3.1 Alur penelitian.....	30
Gambar 3.2 Arsitektur Data Mart	33
Gambar 3.3 ETL Proses	34
Gambar 3.4 unzip file.....	35
Gambar 3.5 library	35
Gambar 3.6 Foto KTP.....	36
Gambar 3.7 Atribut KTP.....	36
Gambar 3.8 Output KTP	38
Gambar 3.9 Foto SIM	39
Gambar 3.10 Atribut SIM	39
Gambar 3.11 Output SIM.....	41
Gambar 3.12 Foto NPWP	41
Gambar 3.13 Atribut NPWP	42
Gambar 3.14 Output NPWP.....	43
Gambar 3.15 File JSON	43
Gambar 3.16 Arsitektur Sistem.....	45
Gambar 3.17 Virtual Private Cloud	47
Gambar 4 1 Proses Bisnis dan Sistem PT XYZ.....	50
Gambar 4 2 Konfigurasi S3	52
Gambar 4 3 Permission policies TransformasiData.....	53
Gambar 4 4 Permission testskripsi.....	53
Gambar 4 5 RDS Koneksi.....	54
Gambar 4 6 Konfigurasi VPC.....	54
Gambar 4 7 Konfigurasi RDS.....	55
Gambar 4 8 ERD Star Schema.....	55
Gambar 4 9 Tabel dimensi KTP	59
Gambar 4 10 Tabel dimensi SIM	60
Gambar 4 11 Tabel dimensi NPWP.....	60
Gambar 4 12 Tabel dimensi Invoice	61
Gambar 4 13 Tabel dimensi Data.....	62
Gambar 4 14 View Data Mart.....	62
Gambar 4 15 View tipe SIM A	63
Gambar 4 16 View tipe SIM B	63
Gambar 4 17 View tipe SIM C	63
Gambar 4 18 View KTP Jabodetabek.....	64
Gambar 4 19 View KTP Non Jabodetabek	64
Gambar 4 20 View nasabah wanita.....	64
Gambar 4 21 View nasabah pria	65

DAFTAR LAMPIRAN

Lampiran A Wawancara Head of IT Data Arsitektur	74
Lampiran B Wawancara Tim IT Arsitek	75
Lampiran C Turnitin	76
Lampiran D Form Bimbingan Skripsi	86



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA