

Skema Database Sapa Tazkia Chatbot System

Database Engine: PostgreSQL 14+ atau MySQL 8.0+

A. Tabel dan Kolom

1. users

Tabel utama untuk menyimpan data pengguna (mahasiswa)

Kolom	Tipe Data	Key	Deskripsi
user_id	INT	PK	Primary Key (auto-increment)
nim	VARCHAR(20)	UK	Nomor Induk Mahasiswa (Unique)
full_name	VARCHAR(100)		Nama Lengkap Mahasiswa
email	VARCHAR(100)	UK	Email Mahasiswa (Unique)
password_hash	VARCHAR(255)		Password terenkripsi (bcrypt/argon2)
phone	VARCHAR(15)		Nomor telepon (opsional)
profile_picture_url	VARCHAR(255)		URL foto profil
program_studi_id	INT	FK	Foreign key ke PROGRAM_STUDI
angkatan	YEAR		Tahun angkatan

status	ENUM		Nilai: 'aktif', 'cuti', 'lulus', 'do'
created_at	TIMESTAMP		Waktu pembuatan akun (UTC)
updated_at	TIMESTAMP		Waktu update terakhir (UTC)
last_login	TIMESTAMP		Login terakhir (UTC)

2. program_studi

Master tabel untuk program studi

Kolom	Tipe Data	Key	Deskripsi
program_studi_id	INT	PK	Primary Key
program_studi_name	VARCHAR(100)	UK	Nama Program Studi
program_code	VARCHAR(10)	UK	Kode Program
faculty_id	INT	FK	Foreign key ke FACULTY
status	ENUM		'active', 'inactive'
created_at	TIMESTAMP		Waktu pembuatan (UTC)

Keuntungan: Memudahkan update data program studi, menghindari redundansi data

3. faculty

Master tabel untuk fakultas/jurusan

Kolom	Tipe Data	Key	Deskripsi
faculty_id	INT	PK	Primary Key
faculty_name	VARCHAR(100)	UK	Nama Fakultas
faculty_code	VARCHAR(10)	UK	Kode Fakultas
status	ENUM		'active', 'inactive'
created_at	TIMESTAMP		Waktu pembuatan (UTC)

4. session

Tabel untuk mengelola sesi login pengguna

Kolom	Tipe Data	Key	Deskripsi
session_id	VARCHAR(255)	PK	Session token (UUID v4)
user_id	INT	FK	Foreign key ke USERS (nullable untuk guest)

session_type	ENUM		'authenticated', 'guest'
ip_address	VARCHAR(45)		IP address pengguna
user_agent	TEXT		Browser/device info
is_active	BOOLEAN		Status sesi aktif (default: TRUE)
created_at	TIMESTAMP		Waktu session dibuat (UTC)
last_activity	TIMESTAMP		Aktivitas terakhir (UTC)
expires_at	TIMESTAMP		Waktu session expire (UTC)

5. chat_conversation

Tabel untuk menyimpan percakapan chat

Kolom	Tipe Data	Key	Deskripsi
conversation_id	INT	PK	Primary Key (auto-increment)
user_id	INT	FK	Foreign key ke USERS (NULL untuk guest)
session_id	VARCHAR(255)	FK	Foreign key ke SESSIONS
title	VARCHAR(200)		Judul Percakapan
model_version	VARCHAR(50)		Versi model AI yang digunakan
status	ENUM		'active', 'archived', 'deleted'

created_at	TIMESTAMP		Waktu conversation dibuat (UTC)
updated_at	TIMESTAMP		Update terakhir (UTC)

6. chat_messages

Tabel untuk menyimpan pesan individual dalam chat

Kolom	Tipe Data	Key	Deskripsi
message_id	INT	PK	Primary Key (auto-increment)
conversation_id	INT	FK	Foreign key ke CHAT_CONVERSATIONS
role	ENUM		'user', 'assistant', 'system'
content	TEXT		Isi pesan
content_type	ENUM		'text', 'file', 'link' (opsional)
tokens_used	INT		Jumlah token yang digunakan

response_time	FLOAT		Waktu respon (detik)
is_edited	BOOLEAN		Apakah pesan pernah diedit
created_at	TIMESTAMP		Waktu pesan dibuat (UTC)
updated_at	TIMESTAMP		Waktu edit terakhir (UTC)

7. chat_rate_limits

Tabel untuk membatasi jumlah chat (10 chat per 6 jam per user)

Kolom	Tipe Data	Key	Deskripsi
limit_id	INT	PK	Primary Key (auto-increment)
user_id	INT	FK	Foreign key ke USERS (nullable untuk guest)
session_token	VARCHAR(255)		Session token untuk tracking

chat_count	INT		Jumlah chat dalam window 6 jam
window_start	TIMESTAMP		Awal periode 6 jam (UTC)
window_end	TIMESTAMP		Akhir periode 6 jam (UTC)
last_chat_at	TIMESTAMP		Chat terakhir (UTC)
updated_at	TIMESTAMP		Update terakhir (UTC)

8. courses

Master tabel untuk mata kuliah

Kolom	Tipe Data	Key	Deskripsi
course_id	INT	PK	Primary Key

course_code	VARCHAR(20)	UK	Kode mata kuliah
course_name	VARCHAR(100)	UK	Nama mata kuliah
sks	INT		Jumlah SKS (default dari data, bisa override)
program_studi_id	INT	FK	Foreign key ke PROGRAM_STUDI
semester_offered	INT		Semester ditawarkan (1-8)
status	ENUM		'active', 'inactive', 'archived'
created_at	TIMESTAMP		Waktu pembuatan (UTC)

Keuntungan: Data mata kuliah terpusat, menghindari redundansi, mudah di-update

9. academic_grades

Tabel untuk menyimpan nilai akademik mahasiswa

Kolom	Tipe Data	Key	Deskripsi
grade_id	INT	PK	Primary Key (auto-increment)
user_id	INT	FK	Foreign key ke USERS
course_id	INT	FK	Foreign key ke COURSES
semester	VARCHAR(10)		Semester (format: 2024/1, 2024/2)
grade	VARCHAR(2)		Nilai huruf (A, A-, B+, B, B-, C+, C, D, E)
grade_point	DECIMAL(3,2)		Nilai angka (4.00, 3.50, dst)
created_at	TIMESTAMP		Waktu data dibuat (UTC)
updated_at	TIMESTAMP		Waktu update terakhir (UTC)

10. academic_summary

Tabel untuk menyimpan ringkasan akademik mahasiswa

Kolom	Tipe Data	Key	Deskripsi
summary_id	INT	PK	Primary Key (auto-increment)
user_id	INT	FK	Foreign key ke USERS
total_sks	INT		Total SKS yang diambil
ipk	DECIMAL(3,2)		IPK kumulatif
gpa_last_semester	DECIMAL(3,2)		GPA semester terakhir

semester_active	INT		Semester aktif saat ini
total_semesters_completed	INT		Total semester selesai
updated_at	TIMESTAMP		Update terakhir (UTC)

Notes: Tabel ini di-update setiap kali ada perubahan nilai atau akhir semester

11. knowledge_documents

Tabel metadata dokumen untuk RAG (Retrieval-Augmented Generation)

Kolom	Tipe Data	Key	Deskripsi
document_id	INT	PK	Primary Key (auto-increment)
title	VARCHAR(200)		Judul dokumen
description	TEXT		Deskripsi singkat dokumen
category_id	INT	FK	Foreign key ke DOCUMENT_CATEGORY
file_path	VARCHAR(255)		Path file di server
file_type	ENUM		'pdf', 'docx', 'txt', 'html'
file_size	INT		Ukuran file dalam bytes
vector_ids	JSON		Array ID di vector database (Qdrant)
chunks_count	INT		Jumlah chunks dari dokumen

status	ENUM		'active', 'archived', 'processing', 'failed'
last_indexed_at	TIMESTAMP		Waktu indexing terakhir (UTC)
uploaded_by	INT	FK	User yang upload (admin/staff)
uploaded_at	TIMESTAMP		Waktu upload (UTC)
updated_at	TIMESTAMP		Update terakhir (UTC)

12. document_category

Master tabel untuk kategori dokumen

Kolom	Tipe Data	Key	Deskripsi
category_id	INT	PK	Primary Key
category_name	VARCHAR(50)	UK	Nama kategori (pendaftaran, akademik, pembayaran, dll)
description	TEXT		Deskripsi kategori
display_order	INT		Urutan tampilan
status	ENUM		'active', 'inactive'
created_at	TIMESTAMP		Waktu pembuatan (UTC)

13. config

Tabel konfigurasi sistem yang fleksibel (key-value store)

Kolom	Tipe Data	Key	Deskripsi
id	VARCHAR(255)	PK	Primary key (UUID atau unique string)
key	VARCHAR(50)	Unique	Nama konfigurasi (unique)
value	VARCHAR(255)		Nilai konfigurasi

14. usage_statistics

Tabel untuk statistik penggunaan chatbot

Kolom	Tipe Data	Key	Deskripsi
stat_id	INT	PK	Primary Key (auto-increment)
date	DATE		Tanggal statistik
total_users	INT		Total user aktif
new_users	INT		User baru hari ini
total_conversations	INT		Total percakapan
total_messages	INT		Total pesan

total_tokens_used	INT		Total token digunakan
avg_response_time	FLOAT		Rata-rata waktu respon (detik)
max_response_time	FLOAT		Waktu respon maksimal (detik)
avg_tokens_per_message	FLOAT		Rata-rata token per pesan
created_at	TIMESTAMP		Waktu pembuatan (UTC)

Indexes:

CREATE INDEX idx_stats_date ON usage_statistics(date);

15. pdf_exports

Tabel untuk tracking export PDF transkrip

Kolom	Type Data	Key	Deskripsi
export_id	INT	PK	Primary Key (auto-increment)
user_id	INT	FK	Foreign key ke USERS
export_type	ENUM		'transcript', 'semester_grades', 'academic_summary'

file_path	VARCHAR(255)		Path file PDF di server
file_size	INT		Ukuran file dalam bytes
status	ENUM		'pending', 'completed', 'failed'
created_at	TIMESTAMP		Waktu export (UTC)

16. audit_logs

Tabel untuk mencatat akses data sensitif (keamanan)

Kolom	Tipe Data	Key	Deskripsi
log_id	INT	PK	Primary Key (auto-increment)
user_id	INT	FK	Foreign key ke USERS (nullable untuk system actions)
action	VARCHAR(50)		Action yang dilakukan (view_grades, download_transcript, login, dll)
resource	VARCHAR(100)		Resource yang diakses
old_value	TEXT		Nilai lama (untuk update)
new_value	TEXT		Nilai baru (untuk update)
ip_address	VARCHAR(45)		IP address user
user_agent	TEXT		Browser/device info

status	ENUM		'success', 'unauthorized', 'failed'
error_message	TEXT		Pesan error jika ada
created_at	TIMESTAMP		Waktu akses (UTC)

17. admin_users

Tabel untuk admin/staff yang mengelola sistem

Kolom	Tipe Data	Key	Deskripsi
admin_id	INT	PK	Primary Key
username	VARCHAR(50)	UK	Username admin
email	VARCHAR(100)	UK	Email admin
password_hash	VARCHAR(255)		Password terenkripsi
role	ENUM		'super_admin', 'admin', 'staff'
permissions	JSON		Array permissions yang dimiliki
status	ENUM		'active', 'inactive', 'suspended'
created_at	TIMESTAMP		Waktu pembuatan (UTC)

last_login	TIMESTAMP		Login terakhir (UTC)
------------	-----------	--	----------------------

B. Relasi Antar Tabel

Diagram Relasi (Simplified)

USERS (1)	(M) SESSIONS
USERS (1)	(M) CHAT_CONVERSATIONS
USERS (1)	(M) ACADEMIC_GRADES
USERS (1)	(1) ACADEMIC_SUMMARY
USERS (1)	(M) PDF_EXPORTS
USERS (1)	(M) AUDIT_LOGS
USERS (M)	(1) PROGRAM_STUDI
USERS (1)	(M) CHAT_RATE_LIMITS

PROGRAM_STUDI (M)	(1) FACULTY
PROGRAM_STUDI (1)	(M) COURSES

COURSES (1)	(M) ACADEMIC_GRADES
-------------	---------------------

SESSIONS (1)	(M) CHAT_CONVERSATIONS
--------------	------------------------

CHAT_CONVERSATIONS (1)	(M) CHAT_MESSAGES
------------------------	-------------------

DOCUMENT_CATEGORY (1)	(M) KNOWLEDGE_DOCUMENTS
-----------------------	-------------------------

Relasi Detail

Dari	Ke	Kolom	Tipe	Keterangan
users	program_studi	program_studi_id	N:1	Setiap user punya 1 program studi
users	sessions	user_id	1:M	1 user bisa punya banyak session
users	chat_conversations	user_id	1:M	1 user bisa punya banyak conversation
users	academic_grades	user_id	1:M	1 user bisa punya banyak nilai
users	academic_summary	user_id	1:1	1 user punya 1 summary
users	pdf_exports	user_id	1:M	1 user bisa export banyak PDF

users	audit_logs	user_id	1:M	1 user bisa punya banyak audit log
program_studi	faculty	faculty_id	N:1	Banyak program studi di 1 faculty
program_studi	courses	program_studi_id	1:M	1 program studi punya banyak course
courses	academic_grades	course_id	1:M	1 course punya banyak grade
sessions	chat_conversations	session_id	1:M	1 session punya banyak conversation
chat_conversations	chat_messages	conversation_id	1:M	1 conversation punya banyak message
document_category	knowledge_category	category_id	1:M	1 category punya banyak document

C. Normalisasi Database

Status Normalisasi: 3rd Normal Form (3NF)

1st Normal Form (1NF)

- **Status:** PASS
- Setiap kolom berisi nilai atomik (tidak ada repeating groups)
- Semua nilai adalah scalar

2nd Normal Form (2NF)

- **Status:** PASS
- Tidak ada partial dependency
- Setiap non-key attribute fully depend pada primary key
- Tabel ACADEMIC_GRADES: grade_point fully depend pada (user_id, course_id, semester)

3rd Normal Form (3NF)

- **Status:** PASS
 - Tidak ada transitive dependency
 - Tabel USERS: course_name dipindahkan ke tabel COURSES
 - Tabel USERS: program_studi dipindahkan ke tabel PROGRAM_STUDI
 - Tabel KNOWLEDGE_DOCUMENTS: category dipindahkan ke tabel DOCUMENT_CATEGORY
-