

Database: linkedin_data

=====

Table: profiles

Description: LinkedIn individual profile information

Primary Key: profile_id

Columns:

- profile_id (BIGINT): Unique profile identifier
- id (VARCHAR): Internal LinkedIn identifier
- first_name (VARCHAR): First name of the profile
- last_name (VARCHAR): Last name of the profile
- full_name (VARCHAR): Full name of the profile
- headline (TEXT): LinkedIn headline
- summary (TEXT): Profile summary or about section
- location (VARCHAR): Profile location (city/region)
- profile_country (VARCHAR): Country code of profile
- profile_language (VARCHAR): Language of profile
- job_title (VARCHAR): Job title
- current_job_title (VARCHAR): Current job title
- current_company_name (VARCHAR): Name of current company
- current_company_url (TEXT): URL of current company
- number_connections (INT): Number of LinkedIn connections
- number_followers (INT): Number of followers
- languages (JSON): List of languages spoken
- skills (JSON): List of skills
- education (JSON): Education history records
- experiences (JSON): Work experience records
- volunteer_experiences (JSON): Volunteer experience records
- open_to_work (BOOLEAN): Indicates if profile is open to work
- birth_date (DATE): Birth date

- linkedin_profile_url (TEXT): Full LinkedIn profile URL
 - linkedin_profile_handle (VARCHAR): Public LinkedIn handle
 - linkedin_profile_id (VARCHAR): LinkedIn profile ID
 - connection_degree (VARCHAR): Connection degree (e.g., 1st, 2nd)
 - extracted_at (DATETIME): Data extraction timestamp
 - FileName (VARCHAR): Source file name
-

Table: companies

Description: Company information from LinkedIn and Crunchbase

Primary Key: company_id

Columns:

- company_id (BIGINT): Unique company identifier
- company_name (VARCHAR): Name of the company
- company_industry (VARCHAR): Industry category
- company_category (VARCHAR): Business category
- company_type (VARCHAR): Company type (e.g., Public, Private)
- company_description (TEXT): Company description
- company_tagline (VARCHAR): Company tagline
- company_website (VARCHAR): Official website
- company_domain (VARCHAR): Company domain name
- company_phone (VARCHAR): Company phone number
- company_city (VARCHAR): City where company is located
- company_country (VARCHAR): Country of company
- company_geographic_area (VARCHAR): Geographic region
- company_headquarters (TEXT): Headquarters address
- company_locations (JSON): List of company locations
- company_company_size (VARCHAR): Company size range
- company_employees_count (INT): Exact employee count
- company_employees_range (VARCHAR): Employee range category

- company_followers_count (INT): LinkedIn follower count
 - company_year_founded (INT): Year founded
 - company_founded_on (INT): Founded year (alternative field)
 - company_industries (JSON): List of industries
 - company_specialities (JSON): List of specialties
 - company_last_funding_raised (DECIMAL): Last funding amount raised
 - company_last_funding_type (VARCHAR): Type of last funding
 - company_last_funding_date (DATE): Date of last funding
 - company_last_funding_currency (VARCHAR): Currency of last funding
 - company_crunchbase_url (TEXT): Crunchbase profile URL
 - company_linkedin_company_url (TEXT): LinkedIn company URL
 - company_linkedin_company_id (VARCHAR): LinkedIn company ID
 - company_logo_url (TEXT): Company logo URL
 - company_updated_at (DATETIME): Last updated timestamp
-

Table: profile_company_relation

Description: Relationship table linking profiles to companies

Columns:

- profile_id (BIGINT): Reference to profiles.profile_id
- company_id (BIGINT): Reference to companies.company_id
- relation_type (VARCHAR): Type of relation (current, past, etc.)
- job_title (VARCHAR): Job title in the company
- start_date (DATE): Employment start date
- end_date (DATE): Employment end date
- is_current (BOOLEAN): Indicates if this is the current position

Relationships:

- profile_company_relation.profile_id → profiles.profile_id (many-to-one)
 - profile_company_relation.company_id → companies.company_id (many-to-one)
-