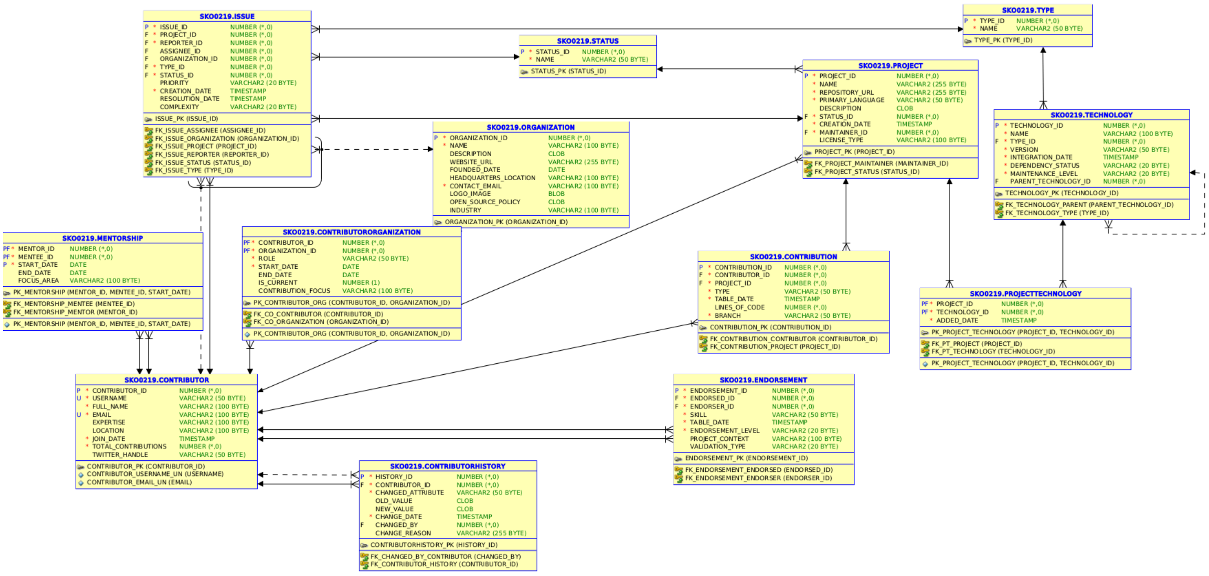
Technologie Databázových Systémů I

Semestrální projekt 2025
Boris Skok
SKO0219

Diagram datového relačního modelu


SKO0219.ISSUE
P * ISSUE_ID NUMBER (*,0)
F * PROJECT_ID NUMBER (*,0)
F * REPORTER_ID NUMBER (*,0)
F ASSIGNEE_ID NUMBER (*,0)
F ORGANIZATION_ID NUMBER (*,0)
F * TYPE_ID NUMBER (*,0)
F * STATUS_ID NUMBER (*,0)
PRIORITY VARCHAR2 (20 BYTE)
* CREATION_DATE TIMESTAMP
RESOLUTION_DATE TIMESTAMP
COMPLEXITY VARCHAR2 (20 BYTE)
ISSUE_PK (ISSUE_ID)
FK_ISSUE_ASSIGNEE (ASSIGNEE_ID)
FK_ISSUE_ORGANIZATION (ORGANIZATION_ID)
FK_ISSUE_PROJECT (PROJECT_ID)
FK_ISSUE_REPORTER (REPORTER_ID)
FK_ISSUE_STATUS (STATUS_ID)
FK_ISSUE_TYPE (TYPE_ID)
SKO0219.STATUS
P * STATUS_ID NUMBER (*,0)
* NAME VARCHAR2 (50 BYTE)
STATUS_PK (STATUS_ID)
SKO0219.TYPE
P * TYPE_ID NUMBER (*,0)
* NAME VARCHAR2 (50 BYTE)
TYPE_PK (TYPE_ID)
SKO0219.PROJECT
P * PROJECT_ID NUMBER (*,0)
* NAME VARCHAR2 (255 BYTE)
* REPOSITORY_URL VARCHAR2 (255 BYTE)
* PRIMARY_LANGUAGE VARCHAR2 (50 BYTE)
DESCRIPTION CLOB
F * STATUS_ID NUMBER (*,0)
* CREATION_DATE TIMESTAMP
F * MAINTAINER_ID NUMBER (*,0)
LICENSE_TYPE VARCHAR2 (100 BYTE)
PROJECT_PK (PROJECT_ID)
FK_PROJECT_MAINTAINER (MAINTAINER_ID)
FK_PROJECT_STATUS (STATUS_ID)
SKO0219.TECHNOLOGY
P * TECHNOLOGY_ID NUMBER (*,0)
* NAME VARCHAR2 (100 BYTE)
F * TYPE_ID NUMBER (*,0)
* VERSION VARCHAR2 (50 BYTE)
* INTEGRATION_DATE TIMESTAMP
* DEPENDENCY_STATUS VARCHAR2 (20 BYTE)
* MAINTENANCE_LEVEL VARCHAR2 (20 BYTE)
F PARENT_TECHNOLOGY_ID NUMBER (*,0)
TECHNOLOGY_PK (TECHNOLOGY_ID)
FK_TECHNOLOGY_PARENT (PARENT_TECHNOLOGY_ID)
FK_TECHNOLOGY_TYPE (TYPE_ID)
SKO0219.ORGANIZATION
P * ORGANIZATION_ID NUMBER (*,0)
* NAME VARCHAR2 (100 BYTE)
DESCRIPTION CLOB
WEBSITE_URL VARCHAR2 (255 BYTE)
FOUNDED_DATE DATE
HEADQUARTERS_LOCATION VARCHAR2 (100 BYTE)
* CONTACT_EMAIL VARCHAR2 (100 BYTE)
LOGO_IMAGE BLOB
OPEN_SOURCE_POLICY CLOB
INDUSTRY VARCHAR2 (100 BYTE)
ORGANIZATION_PK (ORGANIZATION_ID)
SKO0219.MENTORSHIP
PF * MENTOR_ID NUMBER (*,0)
PF * MENTEE_ID NUMBER (*,0)
P * START_DATE DATE
END_DATE DATE
FOCUS_AREA VARCHAR2 (100 BYTE)
PK_MENTORSHIP (MENTOR_ID, MENTEE_ID, START_DATE)
FK_MENTORSHIP_MENTEE (MENTEE_ID)
FK_MENTORSHIP_MENTOR (MENTOR_ID)
PK_MENTORSHIP (MENTOR_ID, MENTEE_ID, START_DATE)
SKO0219.CONTRIBUTORORGANIZATION
PF * CONTRIBUTOR_ID NUMBER (*,0)
PF * ORGANIZATION_ID NUMBER (*,0)
* ROLE VARCHAR2 (50 BYTE)
* START_DATE DATE
END_DATE DATE
IS_CURRENT NUMBER (1)
CONTRIBUTION_FOCUS VARCHAR2 (100 BYTE)
PK_CONTRIBUTOR_ORG (CONTRIBUTOR_ID, ORGANIZATION_ID)
FK_CO_CONTRIBUTOR (CONTRIBUTOR_ID)
FK_CO_ORGANIZATION (ORGANIZATION_ID)
PK_CONTRIBUTOR_ORG (CONTRIBUTOR_ID, ORGANIZATION_ID)
SKO0219.CONTRIBUTION
P * CONTRIBUTION_ID NUMBER (*,0)
F * CONTRIBUTOR_ID NUMBER (*,0)
F * PROJECT_ID NUMBER (*,0)
* TYPE VARCHAR2 (50 BYTE)
* TABLE_DATE TIMESTAMP
LINES_OF_CODE NUMBER (*,0)
* BRANCH VARCHAR2 (50 BYTE)
CONTRIBUTION_PK (CONTRIBUTION_ID)
FK_CONTRIBUTION_CONTRIBUTOR (CONTRIBUTOR_ID)
FK_CONTRIBUTION_PROJECT (PROJECT_ID)
SKO0219.PROJECTTECHNOLOGY
PF * PROJECT_ID NUMBER (*,0)
PF * TECHNOLOGY_ID NUMBER (*,0)
* ADDED_DATE TIMESTAMP
PK_PROJECT_TECHNOLOGY (PROJECT_ID, TECHNOLOGY_ID)
FK_PT_PROJECT (PROJECT_ID)
FK_PT_TECHNOLOGY (TECHNOLOGY_ID)
PK_PROJECT_TECHNOLOGY (PROJECT_ID, TECHNOLOGY_ID)
SKO0219.CONTRIBUTOR
P * CONTRIBUTOR_ID NUMBER (*,0)
U * USERNAME VARCHAR2 (50 BYTE)
* FULL_NAME VARCHAR2 (100 BYTE)
U * EMAIL VARCHAR2 (100 BYTE)
EXPERTISE VARCHAR2 (100 BYTE)
LOCATION VARCHAR2 (100 BYTE)
* JOIN_DATE TIMESTAMP
* TOTAL_CONTRIBUTIONS NUMBER (*,0)
TWITTER_HANDLE VARCHAR2 (50 BYTE)
CONTRIBUTOR_PK (CONTRIBUTOR_ID)
CONTRIBUTOR_USERNAME_UN (USERNAME)
CONTRIBUTOR_EMAIL_UN (EMAIL)
SKO0219.CONTRIBUTORHISTORY
P * HISTORY_ID NUMBER (*,0)
F * CONTRIBUTOR_ID NUMBER (*,0)
* CHANGED_ATTRIBUTE VARCHAR2 (50 BYTE)
OLD_VALUE CLOB
NEW_VALUE CLOB
* CHANGE_DATE TIMESTAMP
F CHANGED_BY NUMBER (*,0)
CHANGE_REASON VARCHAR2 (255 BYTE)
CONTRIBUTORHISTORY_PK (HISTORY_ID)
FK_CHANGED_BY_CONTRIBUTOR (CHANGED_BY)
FK_CONTRIBUTOR_HISTORY (CONTRIBUTOR_ID)
SKO0219.ENDORSEMENT
P * ENDORSEMENT_ID NUMBER (*,0)
F * ENDORSED_ID NUMBER (*,0)
F * ENDORSER_ID NUMBER (*,0)
* SKILL VARCHAR2 (50 BYTE)
* TABLE_DATE TIMESTAMP
* ENDORSEMENT_LEVEL VARCHAR2 (20 BYTE)
PROJECT_CONTEXT VARCHAR2 (100 BYTE)
VALIDATION_TYPE VARCHAR2 (20 BYTE)
ENDORSEMENT_PK (ENDORSEMENT_ID)
FK_ENDORSEMENT_ENDORSED (ENDORSED_ID)
FK_ENDORSEMENT_ENDORSER (ENDORSER_ID)

DD S01 L02: Data vs Information

- Data: Raw values stored in database columns

Example: Contributor.join_date = ‘2023-05-12 09:30:00’ Example: Project.primary_language = ‘Java’

- Information: Meaningful interpretation of data

Example: “5 contributors joined in March 2024”
Example: “JavaScript projects have 30% more contributions than Python projects”

DD S02 L02: Entities/Attributes

- Entity: Project (table representing a distinct concept)
- Instance: Row with project_id=101, name=‘Typst Docs’, repository_url=‘…’
- Attribute: Project.primary_language (property describing an entity)
- Identifier: project_id (PRIMARY KEY uniquely identifying each project)

DD S03 L01: Database Relations

- Project ↔ Contribution (1:N)

Each project receives many contributions
Every contribution must belong to exactly one project (mandatory)

- Contributor ↔ Mentorship (M:N)

A contributor can mentor multiple people
A mentee can have multiple mentors

- Technology (Self-Referencing)

Parent-child relationship for technology dependencies
A technology may have zero or one parent technology

DD S30 L04 Matrix diagram

TODO

DD S04 L01: Supertypes/Subtypes

- Supertype: Contributor (common attributes: username, email, join_date)
- Subtype: Maintainer (identified by Project.maintainer_id reference, additional business rules)

DD S04 L02: Business Rules

- “Contributors need ≥10 contributions to become maintainers”
- “Issues must be assigned to either a contributor OR organization (XOR)”
- “Mentorship cannot be self-referential (mentor ≠ mentee)”
- “Projects in ‘Archived’ status cannot receive new contributions”

DD S05 L01: Binding Types

- Portable: ProjectTechnology (junction table with only FKs)
- Non-portable: Contribution.project_id (direct FK in child table)

DD S05 L03: M:N Relationships

- Without info: ProjectTechnology (just project_id + technology_id)
- With info: ContributorOrganization (additional attributes: role, start_date)

DD S06 L01: Identifying Relationship

- Contribution depends on Project
- Transferred key: project_id becomes part of Contribution's composite PK

DD S06 L02-04: Normalization

- 1NF: No repeating groups (Contributor.expertise stores one value)
- 2NF: All attributes depend on full PK (Contribution.lines_of_code depends on both contributor_id+project_id)
- 3NF: No transitive dependencies (Issue.priority depends only on issue_id, not other non-key attributes)

DD S07 L01: ARC (Exclusive Relationship)

- Implemented in Issue table:

```
CHECK (
  (assignee_id IS NULL AND organization_id IS NOT NULL) OR
  (assignee_id IS NOT NULL AND organization_id IS NULL) OR
  (assignee_id IS NULL AND organization_id IS NULL)
)
```

DD S07 L02: Hierarchical/Recursive

- Technology.parent_technology_id references same table

```
SELECT LEVEL, name FROM Technology
CONNECT BY PRIOR technology_id = parent_technology_id
START WITH parent_technology_id IS NULL
```

DD S07 L03: Historical Data

ContributorHistory table tracks:

- changed_attribute (e.g., 'email')
- old_value/new_value
- change_timestamp
- changed_by (audit trail)

DD S09 L01/02: Change Tracking

- Temporal: ContributorOrganization.end_date marks role termination
- Journaling: Project.version_history stores schema changes

DD S10 L01: Readability

- Consistent naming (snake_case)
- Named constraints (fk_contributor_org)
- Logical table grouping in diagrams

DD S10 L02: Generic Modeling

Type table reused for:

- `Technology.type_id`
- `Issue.type_id`

DD S11 L01: Integrity Constraints

- Entity: PRIMARY KEY on all identifier columns
- Attribute: Project.name NOT NULL
- Binding: FOREIGN KEY with ON DELETE CASCADE
- User-defined: CHECK (end_date > start_date)