Analyzing Database Schema Evolution

Loup Meurice and Anthony Cleve

FNRS Research Project
« Data-Intensive Software System Evolution »
in collaboration with the University of Mons

UNIVERSITÉ
DE NAMUR

Context

- Objective: Understanding the evolution history of complex software systems

- Mining software repositories (MSR) techniques ➜ extraction of historical data

- Fewer studies have focussed on database systems and schemas

- Databases constitutes the heart of data-intensive systems

➜ **Understanding database schema evolution**

Approach

DAHLIA (DAtabase ScHema EvoLutIon Analysis) : a fully generic tool-supported approach

1. allowing one to extract database schema historical knowledge from a software project repository

2. proposing an interactive visualization for analyzing database schema history
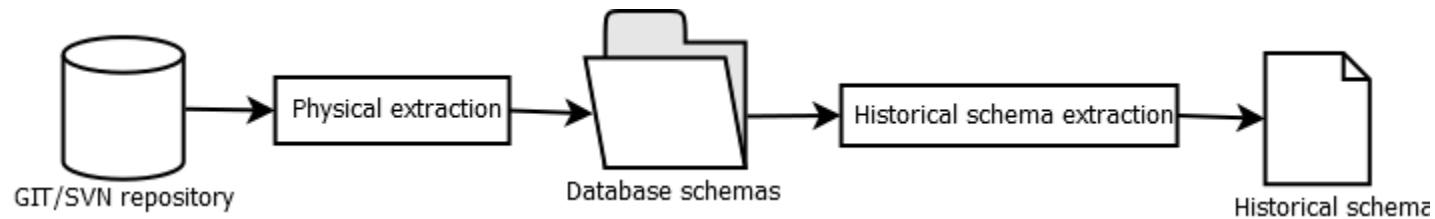
DAHLIA

- How does the database schema evolve over time?

- Which is the evolution trend?

- What are the most common evolution patterns that can be observed

- Which are the most stable tables?

- Who are the most qualified developers for achieving a given schema evolution phase?

DAHLIA applied to real-life case studies

1. **OSCAR** system: a full-featured Electronic Medical Record (EMR) software system for primary care clinics.

2. **MediaWiki**: a free and open source wiki software

3. **Tiki Wiki**: a free and open source wiki-based, content management system and Online office suite

4. **PrestaShop**: a free, open source e-commerce solution

Future directions

- Visualization of co-evolution between code and database (collaboration with the University of Mons)

- Applying DAHLIA to a large set of case studies (DISS)
  - Observation of the general evolution trend within DISS
  - Observation of the most frequent co-evolution patterns

Thank you

Questions?