

Visualizing PostgreSQL Databases

Customized entity-relationship diagrams
using an intermediate JSON-table-schema representation

ibu radempa

Python user group Vienna

2015-09-27

Create some tables using psql and visualize them

Existing tools

Most useful tools are (semi-)proprietary or for proprietary OSs

- ▶ PostgreSQL GUI design tools
- ▶ PostgreSQL GUI tools

Examples of free tools

- ▶ PostgreSQL autodoc: stable perl tool since 2002 ([commits](#))
- ▶ [schemaspy](#): displays a single table with its relations ([sample](#))
- ▶ [Autodoc-Generator](#): generate HTML (rather inactive)

Diagramming software

- ▶ dia
- ▶ yEd

example: input format (excerpt):

```
[Book] {color:orange}
*Book ID
  Subtitle
  Author
  Publisher
  Language
  Summary
  Item ID*
# Relationships
[Item] *--* [Source]
[Item] 1--1 [Location]
[Item] 1--* [Lending]
[Item] *--* [Attachment]
[Event] *--1 [Item] <contains->
[Book] ---- [Item] <is a->
```

Existing tools - missing features

Combination of features I was missing

- ▶ visualize information in table/column comments, in particular 'soft' foreign key relations
- ▶ let (foreign key relation) links end at the respective columns
- ▶ display foreign key relations for multiple columns
- ▶ support PostgreSQL-specific features (e.g., data types); no ORM
- ▶ visual grouping of tables (color, position)
- ▶ individual styling of relations
- ▶ free software license
- ▶ easily hackable, ideally written in python

Tooling modularity

PostgreSQL $\xrightarrow{\text{pg_jts}}$ JSON $\xrightarrow{\text{jts_erd}}$ SVG

- ▶ **pg_jts** produces a *generalized JSON table schema* data structure from a live PostgreSQL database
- ▶ **jts_erd** visualizes a database described by a *generalized JSON table schema* using **PyGraphviz**

Advantages of using *JSON table schema* as intermediate format

- ▶ **jts_render** can be used with other adapters (e.g., MySQL)
- ▶ JSON description of database can be used for other purposes

JSON table schema (JTS)

- ▶ simple [DDL](#) for describing a single table using JSON notation
- ▶ not just for SQL tables, but tabular data in general
- ▶ readable and well-structured
- ▶ can be extended in many ways
- ▶ current version: [1.0-pre9](#)
- ▶ maybe an upcoming standard for table description
- ▶ part of [dataprotocols suite](#), [okfn.org](#)
- ▶ has python package [json-table-schema](#)
- ▶ needs to be extended in case of many coherent tables

- ▶ version 0.0.1
- ▶ uses PostgreSQL system catalog (pg_% tables)
- ▶ needs a live database connection (read only)
- ▶ extracts various information from a single database
- ▶ →source

Generalized JTS, “jts-pg”

JSON object for a single PostgreSQL database

- ▶ namespaces
 - ▶ tables (extension of JTS)
 - ▶ columns
 - ▶ foreign keys

Produces a jts-pg object as output

Differences from current draft

- ▶ use 'column' instead of 'field' (cf. [issue 134](#))

- ▶ version 0.0.1
- ▶ takes a jts-pg object as input
- ▶ uses PyGraphviz to render tables as nodes (of shape record) and (foreign key) relations as edges
- ▶ multicolumn relations are rendered with one or two intermediate nodes (of shape point)
- ▶ takes various options (f.i., omitting column information)
- ▶ `→source`

Missing features

- ▶ **no manual (table) node positioning**, no node pinning
- ▶ **no memory of node positions**
- ▶ views not displayed
- ▶ table inheritance
- ▶ sequences, triggers, functions not displayed
- ▶ (polishing)

Roadmap

- ▶ put source on github
- ▶ add unit tests for pg_jts; support other versions of PostgreSQL, Python3
- ▶ add diagram title, “created with ...” footer, table status ('deprecated')
- ▶ add Javascript functionality for highlighting links/columns/tables: click on a column highlights it and shows details, also for edges (foreign keys)
- ▶ add options for showing tables statistics (number of rows, max/min value for selected fields (e.g. update_time), ...)
- ▶ add input check for jts_erd using [voluptuous](#)
- ▶ improve pg_jts (e.g., table inheritance/[EERD](#))
- ▶ add Javascript functionality for positioning (table) nodes
- ▶ add SQL live editing functionality (with JSON RPC server)

Outlook

Somewhen there might be

- ▶ converters from other RDBMSs to JTS
- ▶ a diff tool for JSON table schemas; schema versioning

Thanks.

Questions?