

UDS and UDSDM Deep Schema Walk

We created an interactive index for most frequently used schemas called Deep Schema Walk:

[QPPAR datamarts overview | korq.github.io \(cms.gov\)](https://korq.github.io/cms.gov)

Features:

- schema view (index page): list of tables, with number of rows and columns
- table view: list of columns, with count of distinct values (cardinality), count of NULLs, and top 5 values
- the values for PII fields are not shown
- sort, filter-as-you-type, and pagination is enabled by jQuery-based DataTables

As a QPPAR developer, you frequently run into small questions/tasks such as:

- look up the exact name of a column in a table
- find the count of distinct values in a column (cardinality)
- find most frequent values in a column
- determine if a column has useful information or mostly NULLs
- verify the grain - number of distinct values in columns that make up the grain should equal the number of rows

These tasks are facilitated by the Deep Schema Walk.

Security audit was performed by Christopher Lamberson:

Brief Vulnerability Assessment and Audit for PII Leakage. I started with a manual review of the website's content, looking for PII leakage. In situations where a TIN may have been revealed, that information is listed accurately as encrypted. I had a long discussion about NPIs, which are revealed in this website, with more senior security engineers and project managers. It appears that NPIs are not considered PII. So, all good there. I could identify no other PII or possible PII being leaked. The functionality and features of this website don't allow for much in the way of an attack surface via the common webapp vulnerabilities like SQLi and XSS. Static site hosting like GitHub Pages are great in this regard.