Local database

One of the ways to get URL categories is to use the local database provided by SafeDNS. This database is getting updates on a daily basis and therefore the database always contains actual information about websites on the Internet. This database is currently supplied in sqlite3.

Updating the database

The database is provided in two sets: a binary file (sqlite3 base) and a patch (SQL constructions to bring the existing sqlite3 database to the actual state).

Here is the source to download the database in a binary file:
`https://url2cat.safedns.com/pubfilter/grandbase.db`

Here is the source to download the database in patches:
`https://url2cat.safedns.com/api/v1/update/<user_version>`
Where `<user_version>` is the PRAGMA parameter of the current database version. Inside the resulting patch, the first line indicates the new version of this parameter, and if it matches the requested version, no update is required.

You can check the current version of the database (PRAGMA parameter user_version) using the command:

```
xxd -l 4 -s 60 grandbase.db
```

The database can be updated by sending a GET request to the specified source with BASIC authorization parameters.

Database description

The database contains two tables:

1. Table result
2. Table cat

Table result

The table result contains hashed URL entries and their categories.

Table scheme:

Name	Data
domain_hash	Hash of a domain
path_hash	Hash of a path
cat_id	Categories list

With a primary key on the fields `domain_hash`, `path_hash`.

The data in the `cat_id` field is stored as a blob array to standardize the enumerated type, where each category is stored as an unsigned short.

Table cat

The table **cat** contains a list of entries with category names and identifiers.
Depending on customer requirements, the SafeDNS Octo database can be supplied with a different number of categories with more detailed categorization or unique category names.

Table scheme:

Name	Data
locale	localization identifier
cat_id	category identifier
name	category name

With a primary composite key `locale`, `cat_id`

Identifiers from the field `cat_id` of the "result" table are the foreign key to this table. The field `locale` contains the localization identifier of the language in which is written the name of the category in the name field.
The `cat_id` field contains the numerical category identifier, the `cat_id` data is not sequential. The name field contains localized category names.

Created 2022-08-22 11:53:38 UTC
Updated 2022-08-23 18:20:47 UTC