# RDSv3 FAQs

May 13, 2025

***Q: Do older versions of RDSv3 databases need to be saved for use with future releases?***

A: Yes, a full version of the RDSv3 SQLite database files will need to be saved, and used when delta releases are published, since the delta release will build upon a full database set. The delta publication will not be a full SQLite database file, but will instead be a SQL file containing INSERT, UPDATE, and DELETE statements, which when run in a full RDSv3 SQLite database file, will update the database to the latest set of hashes and metadata released by the NSRL.

***Q: If I alter the SQLite database file, will I need to download the full database again in order to make updates using the delta publications?***

A: If you make changes within the SQLite database publication itself, then yes, you will need to acquire the unaltered SQLite database file again. If you plan to make changes to the database file, then it is recommended that you make a local copy of the database to remain unaltered, that can later be updated to the latest version via the delta publications.

***Q: Why do I find additional hash entries in the RDSv3 version when compared to the RDSv2 version of the same set?***

A: One of the new data additions to the RDSv3 set, the 'path' column in the METADATA table, is the main reason users will see a difference of over 300 million more hashes in the RDSv3 modern set, compared to the RDS 2.XX version. In RDS 2.XX, any file within the same software package with the same name as another within the same software package is removed. This essentially removes any duplicate rows from NSRLFile.txt. In RDSv3, the METADATA table includes the 'path' column, which differentiates two identical files with the same name. In RDS 2.XX, files with the same name, but different paths would be duplicate entries, and would be removed, while in RDSv3, the entries would not be duplicates, as they would have different paths, therefore the RDSv3 publication has a significantly higher hash count. Comparing RDS 2.76 and RDS 2022.03.1, which were published at the same time from the same source data, we see that the RDS 2.76 Modern set has 209,064,705 hashes, while the RDS 2022.03.1 Modern set has 552,038,839 million hashes.

***Q: Why do I find additional distinct hash entries in the RDSv3 version when compared to the RDSv2 version of the same set?***

A: Additional hash entries can be found in the RDSv3 sets when compared to the same set of the RDSv2 version, due to the ability of the SQLite database being able to handle more characters that the text files of the RDSv2 can support. In RDSv2, some lines are removed from the publication file, as the file names contain characters that would break the formatting of the text files, causing errors while try to read the files. There is no error with the extra data in the RDSv3 publication. It is a result of the SQLite database being able to handle more characters.

***Q: I only care about the hash values being published. Do I need to download the full database, and the full database delta files each time?***

A: No. There has been much interest by our users for including a minimal hash database version of the RDSv3 publication, which will reduce the size of the database and delta file downloads, by only including data that is equivalent to the old RDS 2.XX text files previously published by the NSRL. The minimal database publication mirrors the FILE, MFG, OS, and PKG views of the full publication database, but will only contain the set of FILE data that includes just distinct SHA256 hashes (reference the minimal RDSv3 database schema provided in the published RDSv3.pdf document).

Currently, the NSRL is only planning to publish the minimal database for the Modern hash set, as this set has received the most interest for the inclusion of a minimal database. The NSRL may publish minimal databases for other hash sets, if there is sufficient demand.

***Q: I like the minimal hash set, but I would only like distinct hash values from this set. How can I get just the distinct hash values?***

A: The minimal database schema includes a view (pre-coded database query) which provides just the distinct SHA256, SHA1, and MD5 hash values for every file in the RDSv3 publication. A simple select query of all the data in this view will provide just the distinct hash values (no duplicated hash values) for the hash set in which the minimal database was made.

***Q: I need the data published in the RDSv3 publications in the same text file format as the RDS 2.XX publications. Will the NSRL publish in a text file format? Alternatively, can a text file format be created from an RDSv3 database?***

A: The NSRL will no longer be publishing the RDS data set in a text format. However, the NSRL has published detailed instructions on the RDS downloads page in a file named RDSv3_to_RDSv2_text_files.pdf. By following these instructions, an RDS 2.XX format set of text files, and UDF image, can be created using the RDSv3 minimal database publications.

***Q: How do I generate dbhash values for an RDS database?***

A: DBhash is a SQLite program that is designed to generate a unique database hash specifically for SQLite databases. The dbhash program must be installed on your computer which houses your RDS databases (installation procedures vary dependent on your operating system. More information on the dbhash program can be found here: https://www.sqlite.org/dbhash.html . Running the dbhash program, once installed, is as simple as navigating to the directory, via command line, in which your RDS database is stored, and running the dbhash program against your database.

For example, running the following:

```
dbhash RDS_2025.03.1_modern.db
```

Produces the following output:

```
25eddfb966c362cd9870d3d081db6ddabf64e3d0 RDS_2025.03.1_modern.db
```

The direct output of the dbhash program, as seen here, is what gets published in the dbhashes.txt file with each publication.