

NPPEData.com Getting Started Guide

Last Updated: 6/12/2008

Introduction

This document is for those who have purchased or are thinking of purchasing an NPPEData.com data service. This includes the following services:

- Single Download – Flat Structure
- Single Download - Relational Structure
- 1 Year Subscription - Flat Structure - Update Service
- 1 Year Subscription - Relational Structure - Update Service
- Multi-Tenant - 1 Year Subscription
- Redistributable - 1 Year Subscription

This guide provides general information about our services and possibilities for implementing our data services for your application. It is intended to supplement the information found within our implementation and update guides that are available once you have activated a license for one or more of our services.

How our Services Work

Once you have selected a data service and your license is activated (this is automatic with credit card purchases and once we receive and post your check if you select to pay by check or money order), you can access the files available for download from “My Downloads” on the “Account Options” menu while logged in.

Each month, we will notify you by e-mail when the latest files are available to download on our web site. If you selected a shipping option, we begin processing shipped media immediately after we make the files available on our web site.

Initially, regardless of the service selected, you will have access to the latest full file set in your selected format. These downloads will be available to you for a minimum of 3 months after which point we archive the files (the complete set of files for processing and delivering files requires close to 50GB of disk space on our servers). If you need access beyond 90 days, we can restore the archive for you for the cost of a single support token (see below). Restored files will be available to you for an additional 90 days.

For our update services, once the first update occurs (when the following month’s data has been processed), you will also have access to any update file sets that you can access as well as that month’s full file set. These will also be available for a minimum of 3 months.

What's Available Online

Once you log into your account, you will be able to access a number of resources that will help you through your implementation:

- Relational Implementation Guide
- Flat File Implementation Guide
- Update Guide
- Technical FAQ
- Source Code Library

Technical Support

We can provide technical support for your specific implementation. Our technical support process uses tokens that are available for purchase on our site.

We never charge for resolving defects in the data that we caused nor for inquires about sales, services, licensing options, etc. We will make every effort to fix any defect as soon as possible after we are notified.

Technical support is available through "My Support" under the "Account Options" menu. All technical support questions must be entered through our ticketing system ("Create a New Support Ticket" within "My Support"). Technical support inquires submitted by other means will be referred back to the sender to be entered into our ticketing system

Please read our Support Policy carefully. It is available on our web site under "Support Policy".

Service Selection and Initial Implementation Strategy

How you implement the NPPES data is highly dependent on your application for it. Some of our clients may only need to update the full database monthly (essentially reloading it every month) while others may have their own data keyed to values in the NPPES data and require data that CMS removes from the NPPES data set to remain in their own databases in order to preserve long-term data integrity.

Your specific application for the data should also drive your choice of data structure (flat or relational) and how updates are applied (full refresh or incremental updates).

Other factors that you may want to consider in your choice of service and how you implement are:

- There are currently a large number of added and updated records every month (30,000+ adds and 50,000+ updates) with a smaller number of removed records (1000+).
- CMS continues to make changes to the column specifications for their CSV file. In 2008, they have so far added 37 columns. All of our full databases and updates are built using the most recent column specification. This means that you need to make the underlying table changes in order for the updates to be applied properly.
- Sometimes replacing the entire database becomes easier and faster than updating. With frequent changes and large updates, if you aren't required to maintain historical data, monthly replacement might be the simplest option. (Even if you are we offer strategies in our Update Guide to enable re-loading the full dataset without disturbing your existing data.)
- We have not enabled foreign key constraints on any of the relational tables. However we have performed checks on each value to ensure that each can be constrained. If you require constraints you may do so by creating them in your own database.
- Your database may not be 100% ANSI-SQL compliant. This could mean that the table creation scripts may need to be modified for most efficient use in your own database. For example, we provide no sizing information in our scripts or anything specific to a particular database vendor.

There are several initial implementation strategies that we recommend:

1. MySQL Tables

If you have a MySQL database running or it is not difficult for you to get one installed, by far the easiest option is to download the MySQL tables. All you need to do then is to shut down the MySQL service, and then copy the table files into a pre-existing database folder. When MySQL is restarted, it will automatically detect these new tables and you can instantly use the data.

You can also use this strategy if you need to populate data into another vendor's database (Oracle, SQL Server, Sybase, etc.) by using MySQL as a staging database and then using an ETL or other data transformation tool to populate the other database.

2. Bulk Loading using CSV or Pipe-Delimited Files

If you cannot or do not wish to use MySQL either as a primary or a staging database, the next fastest option is to use the bulk loading utility provided by your database vendor along with our CSV or Pipe-Delimited formats.

Since there are 300+ columns in the flat structure, we provide both DDL to generate the appropriate tables along with easy-to-use text files with lists of columns to make developing bulk load scripts and other utilities easier.

3. Loading via SQL Inserts

Using SQL Inserts is probably the slowest option, but it is also less error-prone and requires less coding than bulk loading.

All of our DML is ANSI-SQL compliant; however, the file sizes are also the largest of any of our formats. In all of our files, we include commits every 5,000 records, so that the datasets can be loaded on servers with limited resources. You may also have to consider date formats (ours are YYYY-MM-DD) prior to loading.

4. Writing your own Software

If you have decided to implement the data using a table structure that differs from ours, you probably will need to write your own programs to transform the data and populate your database. Obviously, we recommend using our own flat or relational table structures, however if you already have software written against an existing database, it may be easier to write software to perform the appropriate transformations as opposed to refactoring your software to query a new structure.

Update Strategy

Once you have the NPPES data implemented, you will need to form a strategy for updating the data on a monthly basis. CMS frequently makes changes to the underlying structure by adding columns. All of our updates are built against the most recent version of the structure.

Each month, if applicable, we will provide ANSI-SQL “alter” scripts to bring your data structure into sync with the current structure.

We provide ANSI-SQL insert and update scripts for all adds/updates/deletes. We do not actually delete rows from your database. Rather, we set our own deleted_flag column. This is so that if you have existing data keyed against NPPES data, you will not create orphaned dependencies in your data when you delete it. You can then use the deleted_flag to limit the ability to query and use deleted records in your application – or you can just simply delete it.

We also provide flat files in CSV and pipe-Delimited formats with adds/updates/deletes. Each file contains the full NPPES record for each change. CSV and pipe-delimited files are most compatible with our flat data structure. However, these files can also be used to update the relational structure with some additional programming.

We recommend the following strategies for updating:

1. Replace the Entire Database

This can be the easiest and quickest option, particularly if you don't have other transactional data that is dependent upon NPPES data.

2. Apply DDL and DML Scripts

If you are using our table structures, each monthly update becomes a matter of applying any table changes that are required via the scripts we provide and then running the ANSI-SQL compliant insert, update and delete scripts that we provide

3. Write software to process updates

We provide CSV and Pipe-Delimited update files so that you can write your own software to process updates. This is the best option if you already have software that uses a different table structure than the one we provide.

In addition to the above, our Update Guide provides additional hybrid strategies that can save you a lot of time.