

TransUnion^{tu}



Data Onboarding File Requirements

July 6, 2023

Version 1.0



© 2023 Neustar LLC
All Rights Reserved

No part of this publication may be reproduced or distributed in any form or by any means, electronic or otherwise, now known or hereafter developed, including, but not limited to, the Internet, without the express written permission of Neustar. This document is protected by U.S. and International copyright laws. This document and the subject matter contained herein is Neustar proprietary and confidential information, and may not be shared or used for any purposes other than the purpose for which it was provided by Neustar, without the express written permission of Neustar. By using this document, you are agreeing that you will not attempt, directly or indirectly, to reverse engineer, decompile, or disassemble any Neustar services or service information, any confidential or proprietary criteria developed or used by Neustar relating to services you receive from Neustar, or any of the Neustar confidential and proprietary information contained herein. The entire right, title and interest in and to this document and Neustar services, and all copyrights, patents, trade secrets, trademarks, trade names, and all other intellectual property rights associated with any and all ideas, concepts, techniques, inventions, processes, or works of authorship including, but not limited to, all materials in written or other tangible form developed or created by Neustar, shall at all times vest exclusively in Neustar. You acknowledge that any misuse, misappropriation or threatened misappropriation of Neustar's intellectual property rights, or any breach or threatened breach of the foregoing restrictions, may cause immediate and irreparable injury to Neustar, and in such event, Neustar shall be entitled to seek injunctive relief. Nothing stated herein will be construed to limit any other remedies available to Neustar including, but not limited to suspension and/or termination of the services provided to you.

Requests for permission to reproduce or distribute any part of, or all of, this publication should be mailed to:

Law Department
Neustar
555 West Adams
Chicago, Illinois 60661

The "tu" logo, Neustar, and other trademarks, service marks, and logos (the "Trademarks") used in this publication are registered or unregistered Trademarks of Neustar LLC or their respective owners. Trademarks may not be used for any purpose whatsoever without the express written permission of the Trademark owner.

The content of this document is furnished for informational use, is subject to change without notice, and should not be construed as a commitment by Neustar. Neustar assumes no responsibility or liability for any errors or inaccuracies that may appear in the content of this guide. This publication may include technical or other inaccuracies or typographical errors. Changes are periodically added to the information herein. Neustar may make improvements and/or changes in the services or facilities described in this publication at any time. Neustar provides this documentation without warranty, term, or condition of any kind, either implied or expressed, including, but not limited to, the implied warranties, terms or conditions of merchantability, satisfactory quality, and fitness for a particular purpose. Neustar, its employees and agents will not be responsible for any loss, however arising, from the use of, or reliance on this information.

[Neustar.com](https://www.neustar.com)

Table of contents

- Overview..... 4**
- Prerequisites..... 4**
- Account Setup 4**
- Data Upload..... 5**
- Data Requirements 5**
 - File Name Patterns and Attributes5
 - Replace Vs. Incremental Files6
 - Data Structure6
 - Campaign Files.....6
 - Data Feeds6
 - Column Headers for Identifiers7
 - Hashing.....9
 - Segment Data Types9
 - File Size Recommendations10
- Neustar Data Retention Policy 10**
 - Data Retention of Neustar SFTP Server.....10
 - Data Retention on Neustar Internal Server.....11
- Revision History 12**

Overview

The following procedure details the file requirements when data is uploaded to the data onboarding service. This document is intended for an engineer or data analyst within the organization that serves as the technical point of contact with Neustar. They are also required to send additional information, such as a taxonomy or destination integration details, as needed.

Neustar supports:

- Campaign lists, where all identifiers in a file are associated with a single attribute.
- Data feeds where a file either has transactional data being passed for use in analytics or contains multiple taxonomized segments.

Campaign lists are typically used for ad-hoc campaign activations. Data feeds are used to provide monthly/weekly updates for segment groups or pass on transactional data for analytics. The data provider must contact its support team on how to best support the application.

Prerequisites

None.

Account Setup

Once Neustar receives the required onboarding information, it sets up the data onboarding account and provides the information to start the onboarding process. Neustar supports pulling data from a client's third-party secure cloud storage (i.e AWS, GCP). SFTP access credentials are sent automatically from Neustar's servers (with the sender address as onboarding-noreply@neustar.biz) through two separate emails. Additional configurations are required for onboarding portal support.

The following time frames apply regarding data onboarding:

- Once data is sent to Neustar, there is at least a 10-business day turnaround to compile the information into a new data source.
- The initial onboarding process typically takes 30 days once the contract is signed.
- Subsequent file processing takes up to five business days.
- If the onboarding portal (<https://onboardingportal.neustar.biz>) is used for campaign lists, it can be activated within one business day (if configured destinations have a properly-formatted file with standard headers when submitted).

Data Upload

1. Once the initial data files are uploaded to the specified SFTP location, the data provider sends its support team a confirmation email.
 - a. Files larger than 1GB must use the SFTP client to send files to SFTP://onboarding.neustar.biz (use Port 22). SSH key authentication is also supported for automated file transfers.
 - b. Files smaller than 1GB use the SFTP web interface (<https://onboarding.neustar.biz>). Uploaded files are transferred to a Neustar server and a receipt is created in the **Receipts** folder.
2. Neustar's team reviews the transaction to verify the process is correctly set up.

Data Requirements

There are several requirements when uploading files to Neustar:

File Name Patterns and Attributes

Upload file names in the following format: **{ClientName} – {DataSet} – {YYYYMMDDHHMMSS}.text**

- **{ClientName}** is the company providing the data.
- **{DataSet}** is a unique ID to differentiate multiple uploads from each another.
 - If the file is a campaign list for syndication designated for a single destination platform, the **{DataSet}** should include both the name of the destination platform and the external segment ID for the platform.
- **{YYYYMMDDHHMMSS}** is the date format when the file uploads to Neustar's SFTP server.

An example of this filename format is: **AcmeCorp-MyPlatforms485-20200118.txt**

- The following file attributes are required for uploading files:
Tab or pipe delimited text files with one line per record.
- UTF-8 character encoding.
- PGP encrypted and compressed using Neustar's public key:
<https://onboarding.neustar.biz/custom/DataOnboarding-pub.asc>

Note

ZIP and Excel files are not supported.

Replace Vs. Incremental Files

Neustar supports the ability to either completely refresh or provide incremental updates to a campaign list or data feed. This setting is managed by the support team and is discussed when setting up a new data feed. The difference between the two is:

- **Replace** allows advertisers to send Neustar a complete update of all users and related audiences and replaces the existing users/segment information.
- **Incremental** allows advertisers to send Neustar updates for individual users without having to onboard the entire audience each time. A lookback window of up to 180 days allows the removal of a user. Support for removing a user is outline in the **Updating Audience Segment Membership** guide.

Note

User removal is only for matched Neustar identifiers.

Data Structure

Campaign Files

All records in a file are associated to a single segment as a campaign list and contain input identifiers that use standard column headers. The format is used to simplify the onboarding and syndication process. Campaign files uploaded to the SFTP are activated through the onboarding portal.

Data Feeds

Data feeds uploaded to Neustar contain input identifiers and either transaction for DSDK data normalization, or taxonomized segments for syndication.

To support files that include taxonomized segments, Neustar recommends structuring data using two delimiters:

- One that separates each input identifier column, and another column of segment IDs.
- One that separates each individual segment ID within the column of segment IDs.

All segment IDs are treated as the enumerated data type; this requires alternate file types with one column for each segment, so the use of other data types is supported. Each enumerated data type segment ID (i.e. A123456) maps to a segment name (i.e. "Buying Power > Very High"), and the list of acceptable segment IDs with corresponding segment names and categories must be provided to Neustar ahead of time.

The following are examples for sending taxonomized segments (with input identifiers in **RED**). **Segments** is a column used to capture all segments that only support Enumerated Data Types.

DSDKID | SEGMENTS

ed7d790e24b8bb4b92bb75b364881e7961aa39d | A123456, B987654

FNAME | LNAME | STREET | CITY | STATE | POSTAL | EMAIL | PHONE | SEGMENTS

John | Doe | 123 Ridge Ln | Gilroy | CA | 93530 | john.doe@gmail.com | 5555555555 | A123456, B987654

Note

JSON blobs are not a supported data type within file attributes.

Column Headers for Identifiers

The column headers for input identifiers in the files must use the names in the following table for each identifier to ensure correct file processing. If the data provided is not associated with an identifier, it does not require a column.

Individual onboarding requires either an email name and address, or a mobile device ID column. Additional headers may be available for specific applications.

Note

Updates to existing files require consistent column headers.

Name	Description/Requirements
CUSTID	A string of characters that uniquely identify the client. A file with only PII that is associated with a CUSTID can be provided as a base file (additional files are provided with a CUSTID and taxonomized segments only). CUSTIDs must be provided with other identifiers to support match back reports.
FNAME	Individual’s first name. Submit Address, Email, or Phone with a name.
LNAME	Individual’s last name. Submit Address, Email, or Phone with a name.
STREET	House number, street pre-direction, street name, street type, and/or post-direction.
SECONDARY	Apartment or suite number.
CITY	Postal city name associated with the street address.
STATE	Two-letter state abbreviation.

POSTAL	Five-digit U.S. postal code associated with the address. Treat as a string of digits with any leading zeroes present.
EMAIL	Individual’s valid email address. Multiple emails can be provided using the same column header in multiple columns. All email addresses must be lower case without spaces.
PHONE	Individual’s ten-digit U.S. phone number. Do not include any dashes, parentheses, or spaces.
MOBILEID	Raw MAID formatted as a 36-character string with dashes (8-4-4-4-12 format) with all upper- or lower-case characters.
IPADDRESS	Dot-decimal notation of IPv4 and/or IPv6 addressees with one IP address per line (i.e. 156.154.81.54 non-proxy IP addresses observed within the past 30 days must be sent).
DSDKHHID	A Neustar Household Level 3 ID in the Federated Identify Graph. These IDs can only be onboarded if they originally came from client’s PII and were onboarded to their DSDK.
DSDKINDVID	A Neustar Household Level 0 ID in the Federated Identify Graph. These IDs can only be onboarded if they originally came from client’s PII and were onboarded to their DSDK.
AKUID	Neustar User ID (ak_user_id) from the DSDK. Also known as User ID. This anonymous identifier ties to Level IDs in the Federated Identify Graph. Files with this type of identifier must not include any other identifier. Match back reports are not supported, and CUSTIDs are not required.
TUID	Browser-based Neustar Cookie ID synced via a pixel. Files with this type of identifier must not include any other identifier. Match back reports are not supported, and CUSTIDs are not required.
SEGMENTS	List of taxonomized Enumerated Data Type segments in one column of the file. Each individual attribute must be separated by a list separator within the Segments column. The list separator must be different from the file delimiter.

Hashing

Send input identifiers as plain text. Email and phone as follows:

- Sent hashed using MD5, SHA1, or SHA256.
- Email address are lower case with no spaces; phone numbers are in U.S. 10-digit format with no dashes, parentheses or spaces before hashing.

Name	Description
HPIIFNLNZIP	This is for hashed PII in the following format: First initial, last name, and five-digit zip code as the whole sha256 hashed values (i.e. sha256 ((fname+lname+zip5).lower())).
HPIIFNLN9ZIP	This is for hashed PII in the following format: First initial, last name, and nine-digit zip code as the whole sha256 hashed values (i.e. sha256 ((fname+lname+zip9).lower())).

To identify the hashed values, append the hash type before the standard header (i.e. MD5EMAIL, SHA1EMAIL).

Submit hashed mobile IDs under the **MobileID** header but expect match rates to be negatively impacted.

Segment Data Types

The data types in the following table are supported in data feeds (campaign files do not require segments defined in the file as it is treated as a segment). If using **Number** or **Date** data types, dedicated columns are required for each value.

Data Type	Description
Enumerated (ENUM)	It is recommended to include all ENUMs in a single column "SEGMENTS". ENUMs are one of a set of expected text values (i.e. the set of acceptable values for "Previous Vehicle Purchases" could be "C4325542" and "C43255423").
Number	Only a single number can an attribute column support. Greater than, less than, and continuous ranges are also supported. Prior to passing number ranges in the data, ensure the format for range values is specified in the latest version of the segment layout file provided to Neustar.

Date	<p>Only a single data can an attribute column support. Dates must be mapped to weeks, quarters, years, or custom ranges. A date value must appear in one of the following formats:</p> <ul style="list-style-type: none"> • MMDDYY • MM/DD/YY • DDMMYYYY • DD/DD/YYYY • YYMMDD • YY/MM/DD • YYYYMMDD • YYYY/MM/DD
-------------	---

File Size Recommendations

File size recommendations provide guidance and flexibility during the optimization processing of the received files. Neustar can opt to either aggregate or stop processing files that fall outside of the following file size requirements:

- Max file size before part files are recommended: 20 GB
- Recommended part file size: 1 to 5 GB
- Minimum file size of part files: 200 MB
- Max number of part files: 20

Files smaller than 200 MB can be delivered provided they are not a subset of a larger dataset being delivered.

Neustar Data Retention Policy

Data Retention of Neustar SFTP Server

Files uploaded to Neustar's SFTP are transferred to an internal server, and a receipt is created in the **Receipts** folder. The transfer typically begins within an hour for properly formatted files, and files are deleted following the transfer.

Data Retention on Neustar Internal Server

All client files transferred to Neustar's internal network for processing are encrypted at rest. All further processing is done against the encrypted files as the original file upload is permanently removed from the SFTP.

Since encrypted data files are deleted 30 days from Neustar's internal servers, files must be processed within 30 days or must be resent to Neustar. The 30-day window begins the day a file is successfully uploaded to the SFTP (includes transactional data sent to Neustar). Data required to support services after the 30-day timeframe must be resent.

Neustar creates and maintains segment data for each file used at destination platforms. Once an incoming file is successfully processed, the segment data is maintained for six months. This allows Neustar to keep segments active within a destination platform.

The data retention timelines have the following impact on syndications:

- Syndications use associated post-process segment data, not client-supplied files.
- Successful processing of data files within the 30-day window results in segment data being available for syndication.
- Data files are treated either as an incremental update or a complete replace.
- All segment data more than six months old is removed from the syndication destination.
- In order to keep segments active with volume, data files need to be resent with the affiliated segments before the six-month window closes.

Revision History

Reviewed date	Description of change	Reviewed by
June 30, 2023	Initial Release	Andrew Blair