



FBI National Data Exchange (N-DEx)

Data Integration (Mapping) Guide

Version 2.0

10/21/2013



CJIS Document Number – N-DEx-DOC-09202-2.0

Prepared by:
Criminal Justice Information Services Division
National Data Exchange Program Office

CHANGE DESCRIPTION FORM

Revision	Change Description	Created/Changed by	Date	Approved By
1.0	Initial Version	Brenda Corder	9/23/2011	Sherri Shreves
2.0	Revision	Julia Minnocci	7/23/2013	Sherri Shreves
2.0	Revision	Julia Minnocci	8/23/2013	Sherri Shreves
2.0	Revision	Julia Minnocci	10/21/2013	Sherri Shreves

National Data Exchange (N-DEx) Program Office

Data Integration (Mapping) Guide

1.0 Document Purpose

The purpose of this guide is to document the data integration process to N-DEx. This process includes data mapping, integration development, and initial test data validation and verification. This document is intended to supplement guidance through the N-DEx Liaison Specialist and serve as a roadmap for submitting data to the N-DEx system. The intended audience for this document is agency officials and technical personnel.

2.0 Process

An N-DEx Liaison Specialist will assist the agency during the integration process. The process is as follows: determining the type(s) of data to be submitted to N-DEx and who will complete the work involved (see N-DEx Data Contribution Checklist). Currently, the N-DEx Program Office offers mapping and data integration process assistance for data submission.

2.1 Data Types

The N-DEx Information Exchange Package Documentation (IEPD) has been posted to the Justice Standards Clearinghouse located at the left side of the www.it.ojp.gov web page. The N-DEx IEPDs currently available is as follows:

N-DEx Incarceration/Booking/Probation/Parole 2.1.1 is the N-DEx submission standard for agencies submitting booking, incarceration, probation and parole, e.g., correctional facilities, prisons, jails, presentence and pretrial, supervised release and holding reports data.

N-DEx Incident-Arrest IEPD 2.1.2 is the N-DEx submission standard for agencies submitting incident, arrest, service call, and missing person report information. This IEPD also enables N-DEx to provide the National Incident-Based Reporting System (NIBRS) extract to the Uniform Crime Reporting (UCR) Program on behalf of NIBRS certified agencies requesting this capability.

If a data-item owner wants a NIBRS extract created from their submission, the incident or arrest report must include the NIBRS Originating Agency Identifier (ORI) element, the NIBRS Report Category Code element with corresponding value (indicating what type of submission should be extracted to NIBRS) and an indication of which NIBRS entities to incorporate in the extract (including the Distribution Text element where

the value is 'NIBRS' as part of the desired entities). This extract will satisfy reporting requirements.

NIBRS participating agencies, who do not wish to utilize the N-DEx extract services for NIBRS reporting, should not provide NIBRS ORI, NIBRS Report Category Code, or Distribution Text elements. Additional details for creating a NIBRS extract will be provided by the N-DEx Liaison Specialist.

2.2 Data Integration Development

The Data Integration phase is comprised of two major tasks. The first task is mapping the agency data elements to the N-DEx IEPD. This task determines where each element from the agency data migrates into the N-DEx schema. The second task is the development of the extraction application software. This task utilizing the mapping results generates the means to extract the agency data, wrap it with N-DEx Extensible Markup Language (XML) tags, and write it to an approved readable format of XML to be digested into N-DEx. The extraction application may be developed using XSLT, JAVA, C++, or .NET.

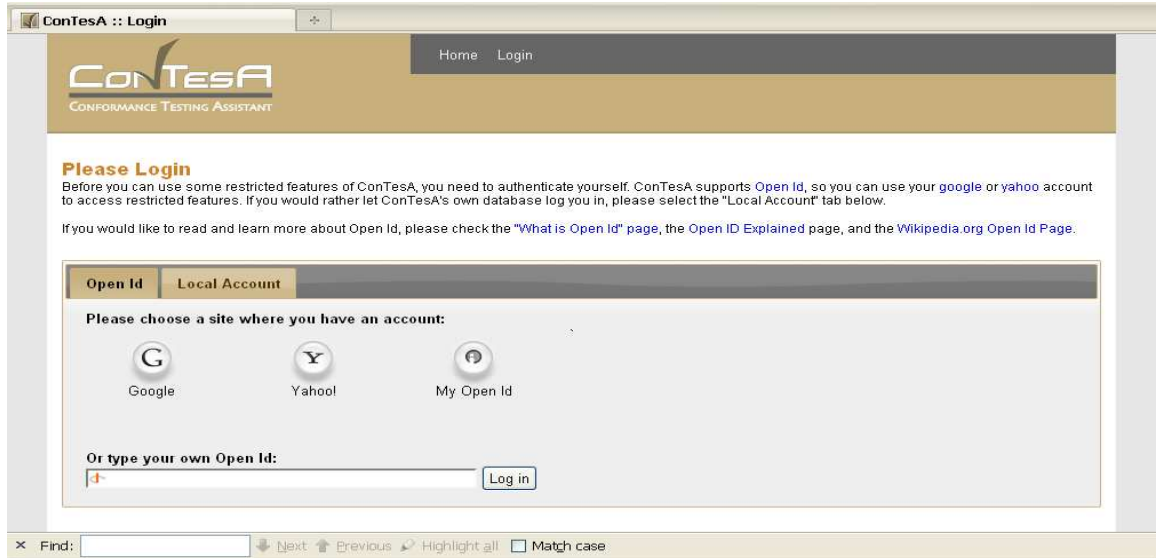
It is the agency's responsibility to assess its capabilities in performing these tasks. The agency should contact the N-DEx Liaison Specialist to determine how the N-DEx Program Office can support the agency with data integration. If the agency chooses to perform the integration, the N-DEx Program Office can offer consultants to answer questions and guide the agency during the mapping/integration development process. Currently, the N-DEx Data Management Team can perform the integration development task free of charge upon agency request.

2.3 Agency Test Data Validation

After the agency data has been extracted into the N-DEx IEPD format, the next step is to conduct XML structural verification checks for National Information Exchange Model (NIEM), Logical Entity eXchange Specifications (LEXS), and N-DEx IEPD conformance. This is accomplished by using the Conformance Testing Assistant (ConTesA). ConTesA is designed to validate an XML instance document against all current versions of the LEXS schema and the N-DEx IEPD. In addition to XML schema validation, ConTesA performs a number of N-DEx business rule validation checks. The tool provides a style sheet for viewing the data, which may help to better visualize the LEXS and N-DEx schemas and the way associations work together. ConTesA is available to participating agencies in order to expedite the data integration process.

It is highly recommended that data be tested and pass the ConTesA validation prior to submitting to the N-DEx Program Office for content

verification. The ConTesA conformance tool is located at <https://contesa.ittl.gtri.org/contesa>. If a user does not have an account one can be created. A command line version of ConTesA is also available for local execution. A sample of the account creation screen is provided below.



1. Log in to ConTesA. <https://contesa.ittl.gtri.org/contesa/>
 - a. Click File Validation
 - b. Enter Username (email address) and Password
2. The user will browse agency's local machine to find the xml instance document they wish to validate
3. Select 'Upload'
4. After validation has completed, select 'View'
5. The user will click on the arrow next to the view they wish to render, e.g., "LEXS 3.1 Detailed View"
6. Click the 'View' tab next to the package

2.4 Test Data File Specification

File Names adhere to the following naming convention:

- AAAAAAAAAA_YYYYMMDD_HHMM_NNNN.xml, where AAAAAAAAAA is the Data Submitter Agency ORI, YYYYMMDD is the 4-digit year, 2-digit month, and 2-digit day, HHMM is the 2-digit hour (24-hour clock) and 2-digit minute, and NNNN is a 4-digit sequence number. The date and time in the file name should be the creation date and time of the xml file. The sequence number is used to guarantee unique file names. The agency system is responsible for making sure the files are named

uniquely. A sample filename for second file created by a Data Submitter Agency with ORI 'ORIPD0001' at 1:02 pm on March 9, 2007, is 'ORIPD0001_20070309_1302_0002.xml'.

File Size:

Note: .xml file = Xml Document Instance
XML Document sizing recommendations

- One message package per a single XML Document Instance e.g. WVNDX0000_20080130_1201_0001.xml
- 70 MB maximum file size for a single XML Document Instance
- 12 Total attachments inside a single XML Document Instance

Zip file sizes for batch ingest:

- see prerequisite XML Document Instance constraints above
- Maximum of 20,000 .xml files in a single .zip file. Each XML Document Instance (file) would have a single message and package
- 2 GB maximum file size for a single .zip file

2.5 CJIS Validation & Verification

Once the agency and/or the N-DEx Program Office has generated error-free test data through ConTesA, the validation & verification submission task may begin.

The following are methods to submit test data:

- External Media device e.g. Compact Disk (CD), Digital Video Disk (DVD), Hard drive, or other media. The external media device should be sent via Federal Express to the N-DEx Program Office.
- Electronic submission, including e-mail and Secure File Transfer Protocol (SFTP) over the Internet.

The below steps describe the process to be followed for submission test data:

1. Submit no more than 100 incident/arrest submissions (ConTesA conformance tested and passed) to the N-DEx Program Office.
2. The N-DEx Program Office will create and provide a report that includes the following:
 - The data verification will include the N-DEx Business Rules found in the IEPD Master Documentation which includes a review of the filename structure.
 - The data verification will also include those elements found in the IEPD Component Mapping Template spreadsheet regarding mandatory elements.
 - Errors encountered will be noted in a report.
 - The secondary reviewer will confirm the results of the report.

3. The N-DEx Program Office will provide the report to the agency and answer agency questions.
4. This process will continue until no errors are found during the test data verification process.
5. The agency will submit 10,000 incident/arrest records. Repeat steps 2-4. The process will continue until no errors are found during the process. At that time, the N-DEx Program Office will notify the agency of the next steps for legacy data submission.

It is suggested N-DEx receive all legacy data from the data submitter and/or data owner. During the test data process, the agency will work with the N-DEx Program Office on the procedure for submitting legacy data to N-DEx. Different options exist based on the amount of legacy data being submitted. The initial dataset can be loaded onto an external media device (CD, DVD, external hard drive, or thumb drive) for manual submission to N-DEx. The external media device should be sent via Federal Express to the N-DEx Program Office. Once the initial legacy data has been ingested into the N-DEx system, the agency can submit to N-DEx via SFTP, web service, web portal or physical media device (CD, DVD, hard drive, etc.)