PROVINCIAL-TERRITORIAL FGP-TBS XML-JSON BUILDER WORKSPACE PROCESS FLOW STANDARD *Each flow chart container represents a workspace bookmark, colours and titles are representative of how each bookmark appears in the workspace*

Standard Published Parameters

* ACTIVATE\_GEO: Selects geomatics data for processing.
* ACTIVATE\_NON\_GEO: Selects non-geomatics data for processing.
* ACTIVATE\_TRANSLATION: Turns the AWS\_TRANSLATE tool on or off. When turned off will add proxy data to the required translated output.
* CATALOGUE\_READER\_SELECT: Selects the P/T API as data extraction source. When turned off, for testing/development purposes, will access a local FFS file with the API data already extracted.
* FORCED\_PYCSW\_URL: URL of a local PyCSW used for development. This is accessed by the METADATA\_DELTA\_FINDER and XML\_PUBLISHER if optioned.
* IN\_CSV\_LOOKUP\_TABLES\_DIR: The directory location of CSV lookup tables accessed in the transformation process.
* IN\_FFS\_TESTING\_FILE: The directory location of the FFS file used in place of the API extraction.
* IN\_GMD\_XML\_TEMPLATE\_DIR: The directory location of the XML templates used to write output XML files. For geomatics data output.
* IN\_TBS\_JSON\_TEMPLATE\_DIR: The directory location of the JSON templates used to write output JSON files. For non-geomatics data output.
* LOCAL\_SOURCE\_METADATA\_DELTA\_FINDER: When selected, the METADATA\_DELTA\_FINDER will access the local directory instead of the PyCSW. For testing and development purposes.
* LOCAL\_WRITER: When selected, will output extracted data to a local directory instead of the PyCSW. For testing and development purposes.
* LOG\_FILE: Directory location the runtime log of the workspace.
* OUT\_JSON\_LOCAL\_DIR: The local directory location of non-geomatics JSON files, when LOCAL\_WRITER is selected.
* OUT\_XLS\_NOTIFICATION\_DIR: Directory location of the PROCESS\_REPORT\_MANUAL\_TASKS xls file used to record anomalies in the transformation process.
* OUT\_XML\_FAILED\_DIR: Directory location of XML files that failed to write to the PyCSW.
* OUT\_XML\_LOCAL\_DIR: Directory location of XML files when LOCAL\_WRITER is selected.
* OUT\_XML\_PASSED\_DIR: Directory location of XML files that successfully published to the PyCSW.
* P-T\_ABBR: The ISO 3166-2 code for the provincial/territorial entity data being transformed.
* PYCSW\_OVERWRITE: Will overwrite all existing datasets in the PyCSW if selected.
* SAMPLE\_SELECT: Will write a specific quantity of data records when selected. For testing and development purposes.
* SAMPLE\_SIZE: The specific quantity of data records to write when SAMPLE\_SELECT is selected.

ETL PROCESS INITIATION

Starts the process and sets date/time for process reports

Transformers:

* Creator transformer to initiate process
* DateTimeStamper
* VariableSetter

DATA EXTRACTION

Reads the API and exposes required attributes.

Filters geo/non-geo data when attribute testing allows for selection

Key Custom Transformers:

* Catalogue\_Reader (universal or P/T specific)
* LICENSE\_FILTER
* GEOPORTAL\_WEBLINK\_ADDER
* P/T specific data reader for attribute exposure/renaming

PRE METADATA MAPPING TRANSFORMATIONS

Contains universal custom transformers that require processing prior to metadata mapping

Key Custom Transformers:

* As required

P/T SPECIFIC TRANSFORMATIONS

Contains custom transformers that are specific to a province or territory due to the unique nature of the metadata and cannot be addressed with a universal transformer.

Key Custom Transformers:

* Any P/T specific CT

Connector Key

All Data

Data Flow | Spatial Management Required | No Spatial Management Required

Geo Data

Nongeo Data

SPATIAL DATA TYPE MANAGEMENT

Will perform one or more of the following tasks; Validates geospatial data and maps spatial data types. Filters geo/non-geo data when it can’t be completed in the DATA EXTRACTION process.

Key Custom Transformers (where required):

* GEOSPATIAL\_DATA\_VALIDATOR
* SPATIAL\_TYPE\_MAPPER
* MANUAL\_GEOSPATIAL\_SETTER

GEO SAMPLER

Outputs a sampling of data for testing purposes. Option selectable in SAMPLE\_SELECT published parameter.

Key Transformers:

* Sampler

METADATA MAPPING

Contains universal custom transformers that utilize lookup tables to map data types to HNAP standards

Key Custom Transformers:

* METADATA\_VALUE\_MAPPER(S)
* METADATA\_FORMAT\_MAPPER
* METADATA\_VALUE\_MAPPER\_ERROR\_MANAGER

PUBLISHING MANAGEMENT

Contains universal custom transformers that determine the URL of the PyCSW Repository, and compares existing repository data with the current run time data to identify records that are new, obsolete or updated.

Key Custom Transformers:

* PYCSW\_URL\_MAPPER
* METADATA\_DELTA\_FINDER

LANGUAGE TRANSLATION

Universal custom transformer that leverages the AWS translation service to translate specific data items to English or French as required.

Due to the fees required of the AWS, there is an option to bypass the translation service and enter proxy items in the required data fields for testing purposes by turning off the ACTIVATE\_TRANSLATION published parameter.

Key Custom Transformer:

* AWS\_TRANSLATE

POST METADATA MAPPING TRANFORMATIONS

Contains universal custom transformers to fine tune and customize metadata prior to publishing

Key Custom Transformers:

* MORE\_INFO\_MANAGER
* REMOVE\_BROKEN\_URL\_WMS\_ESRI\_REST
* WMS\_REST\_LANGUAGE\_FORMATTER
* DUPLICATE\_SERVICE\_REMOVER
* \*GEOPORTAL\_WEBLINK\_ADDER
* GMD\_SECTION\_DATA\_EXTRACTION
* URL\_HTTPS\_MAKER
* Others as required

POST METADATA MAPPING TRANSFORMATIONS – NONGEO

Contains universal custom transformers to fine tune and customize metadata prior to publishing

Key Custom Transformers:

* TBS\_DEFAULT\_KEYWORD\_TOPIC\_SUBJECT
* RESOURCE\_NAME\_UNDERSCORE\_REMOVER
* Others as required

NONGEO SAMPLER

Outputs a sampling of data for testing purposes. Option selectable in SAMPLE\_SELECT published parameter.

Key Transformers:

* Sampler

JSON CREATION

Inserts processed data into a series of JSON templates that are assembled into a single JSON file for each data record, to be written to a local folder

Key Custom Transformer:

* JSON\_PUBLISHER

METADATA OUTPUT

Local repository locations to write XML files. These can be used to validate XML files in the schematron.

Destination Folders:

* XML\_LOCAL: for items bypassing the PyCSW repository when LOCAL\_WRITER published parameter is selected.
* XML\_PASSED: copies of items successfully published to the PyCSW repository.
* XML\_FAILED: copies of items that failed to publish to the PyCSW repository.

XML CREATION & PUBLISHING

Inserts processed data into a series of XML templates that are assembled into a single XML file for each data record, then published to PyCSW repository.

Outputs a copy of each XML file to a local folder and filters records that failed to publish.

For development/testing, the PyCSW repository can be bypassed to write XML to a local folder only by selecting the LOCAL\_WRITER option in published parameters.

Key Custom Transformer:

* XML\_PUBLISHER

LANGUAGE TRANSLATION – NONGEO

Universal custom transformer that leverages the AWS translation service to translate specific data items to English or French as required.

Due to the fees required of the AWS, there is an option to bypass the translation service and enter proxy items in the required data fields for testing purposes by turning off the ACTIVATE\_TRANSLATION published parameter.

Key Custom Transformer:

* AWS\_TRANSLATE

PRE METADATA MAPPING TRANSFORMATIONS - NONGEO

Contains universal custom transformers that require processing prior to metadata mapping

Key Custom Transformers:

* Any P/T specific CT
* RESOURCES\_VALIDATOR
* Others as required

METADATA MAPPING - NONGEO

Contains universal custom transformers that utilize lookup tables to map data types to HNAP standards

Key Custom Transformers:

* METADATA\_VALUE\_MAPPER(S)
* METADATA\_VALUE\_MAPPER(S)\_ONE2MANY
* METADATA\_FORMAT\_MAPPER
* METADATA\_VALUE\_MAPPER\_ERROR\_MANAGER

METADATA OUTPUT – NONGEO

Local repository location to write JSON files.

Destination Folders:

* JSON\_\_LOCAL