

Data loading tools

Overview over all known ETL tools and there capabilities for loading data into the tranSMART database.

Overview of the ETL tools

The [tranSMART Project Management Committee](#) has decided to focus our efforts on transmart-batch and tmDataLoader going forward.

ETL tool	Open Source	tranSMART version (Development Roadmap)	Download	Manual
transmart-data	✓	17.1 and earlier (for data dictionaries and database management)	16.2 and earlier: https://github.com/tranSMART-Foundation/transmart-data 17.1: https://github.com/tranSMART-Foundation/transmart-core/tree/master/transmart-data	Loading data with transmart-data
tmDataLoader	✓	16.2 and earlier (17.1 in development)	https://github.com/Clarivate-LSPS/tmDataLoader	https://github.com/Clarivate-LSPS/tmDataLoader/wiki https://drive.google.com/file/d/0ByehpOFihEbadDZiT1VvYW5ERnM/view
transmart-batch	✓	17.1 and earlier	16.2 and earlier: https://github.com/tranSMART-Foundation/transmart-batch 17.1: https://github.com/tranSMART-Foundation/transmart-core/blob/master/transmart-batch	16.2 and earlier: https://github.com/tranSMART-Foundation/transmart-batch/tree/master/docs 17.1: https://github.com/tranSMART-Foundation/transmart-core/blob/master/transmart-batch/docs/
Integrated Curation Environment (ICE)	✓	16.2 and earlier	https://github.com/transmart/transmart-ICE	https://drive.google.com/file/d/0B8lizkKDeaKhMWZBwnlnODVEQW8/view
Kettle	✓	16.2 and earlier	https://github.com/transmart/transmart-ETL	Loading data with Kettle (Step by step tranSMART ETL Guide)

Overview of the loading capabilities (state of 18.10.2016)

Data type / ETL tool	transmart-data		tmDataLoader		transmart-batch		Integrated Curation Environment (ICE)	
	Support	HDD	Support	HDD	Support	HDD	Support	HDD
Clinical	Y	---	Y	---	Y	---	Y	---
Dictionary	Y	---	N	---	N	---	N	---
aCGH / CNV *	Y	Y	Y	Y	Y	Y	N	N
Metabolomics	Y	Y	Y	Y	Y	Y	Y	Y
miRNA	Y	Y	Y	Y	Y	Y	Y	Y
mRNA	Y	Y	Y	Y	Y	Y	Y	Y
Proteomics	Y	Y	Y	Y	Y	Y	Y	Y
RNASeq	Y	Y	Y	Y	Y	Y	Y	Y
SNP	N	N	Y	Y	N	N	Y	Y
VCF	Y	Y	Y	Y	N	N	N	N

* aCGH data is called CNV data in the ETL tool transmart-batch



Columns:

- **Support:** The column tells if an import is generally possible
- **HDD:** The column tells if an import as an highdimensional data is possible (for explanation of HDD see here: [Supported Data Types](#))

Values:

- **Y:** the ETL tool is able to load the specified data
- **N:** the ETL tool is NOT able to load the specified data

Further information

- [Loading data with transmart-batch \(Cell Line Use Case dataset\)](#)
- [Loading data using tMDataLoader](#)
- [Loading data with Kettle \(Step by step tranSMART ETL Guide\)](#)
- [Loading data using Integrated Curation Environment \(ICE\)](#)
- [Loading data with transmart-data](#)
- [TranSMART Guide for Manual Data Deleting](#)

- [Loading data with transmart-batch \(Cell Line Use Case dataset\)](#)
- [Loading data using tMDataLoader](#)
- [Loading data with Kettle \(Step by step tranSMART ETL Guide\)](#)
- [Loading data using Integrated Curation Environment \(ICE\)](#)
- [Loading data with transmart-data](#)
- [TranSMART Guide for Manual Data Deleting](#)