



## **Catalog Ingest Tool v.1.0.0**

**for the Planetary Data System**



# Table of Contents

---

<b>1</b>	<b>Catalog Ingest Tool Guide</b>	
1.1	Overview	1
1.2	Release Notes	2
1.3	Concept	4
1.4	Installation	7
1.5	Operation	9
1.6	Appendix A - UNIX Setup Options	20
1.7	Appendix B - Windows Setup Options	21
1.8	Appendix C - Referential Integrity Checking	23
1.9	Appendix D - Report Examples	26
1.9.1	Compare Report	27
1.9.2	Validate Report	29
1.9.3	Ingest Report	31



## 1.1 Overview

---

### About Catalog Ingest Tool

The Catalog Ingest Tool (CITool) is intended to streamline the ingestion process. The legacy process involves a myriad of tools and manual steps in order to ingest catalog files submitted by the nodes into the PDS central catalog. CITool offers three functions:

- Compare a Catalog Submission

The comparison function is intended for both Discipline Node and Engineering Node Data Engineers to determine the differences between previous submissions and a current submission.

- Validate a Catalog Submission

The validation function is also intended for both Discipline Node and Engineering Node Data Engineers to validate a current submission including checks for referential integrity.

- Ingest a Catalog Submission

The ingestion function is intended for the Engineering Node Operator to ingest the current submission into the PDS central catalog.

Please send comments, change requests and bug reports to the [PDS Operator](mailto:pds_operator@jpl.nasa.gov) at [pds\\_operator@jpl.nasa.gov](mailto:pds_operator@jpl.nasa.gov).

## 1.2 Release Notes

---

### Release Notes

The purpose of this section is to provide a description of a Catalog Ingest Tool release including any impact that the new or modified capabilities will have on the Discipline Nodes or the PDS user community. A somewhat itemized list of changes for each release can be found on the [Release Changes](#) page. If viewing this document in PDF form, see the appendix for details.

### Release 1.0.0

This is the official release of the Catalog Ingest Tool to the PDS community. The major changes for this release include:

- Interface and General Functionality  
Added the capability to recursively traverse a target directory as well as the ability to turn this capability off. Modified the compare report content by ordering files alphabetically, removing severity levels and not following pointers in compare mode.
- Referential Integrity Checking  
Implemented a couple of additions/fixes related to checking REFERENCE\_KEY\_ID. Added support for multiple values.
- Ingestion  
Implemented several fixes associated with ingestion of the catalog file content into the PDS catalog database. Also added the capability to ingest new standard values as part of the ingestion step.

The liens for this release are as follows:

- Report lines greater than 80 characters  
This capability will be incorporated when CITool is integrated with the new Product Tools library.
- Support for like-named Elements and Objects  
The PDS Data Dictionary currently has elements and objects with identical names. This capability will be incorporated when CITool is integrated with the new Product Tools library.
- Support for multi-valued pointers  
This capability will be incorporated when CITool is integrated with the new Product Tools library.

- Detect changes in like-named Objects

The tool does not correctly detect and report a change within an object that appears multiple times (same name) within the same catalog file.

## **Release 0.4.0**

This release of the Catalog Ingest Tool, is intended for beta testing by the PDS Technical Staff. The minor changes in this release include some updates to the command-line options and message severities.

## **Release 0.3.0**

This release of the Catalog Ingest Tool, is intended for beta testing at the Engineering Node. The major changes in this release include the capabilities to find new standard values and perform a complete referential integrity check within a set of catalog files. Other changes include various, minor fixes to reference validation and the compare mode functionality based on the previous beta testing period.

## **Release 0.2.0**

This release of the Catalog Ingest Tool, is intended for beta testing at the Engineering Node. This release includes the capabilities to validate catalog files, perform referential integrity checking for references and ingest the catalog files into the PDS central catalog database.

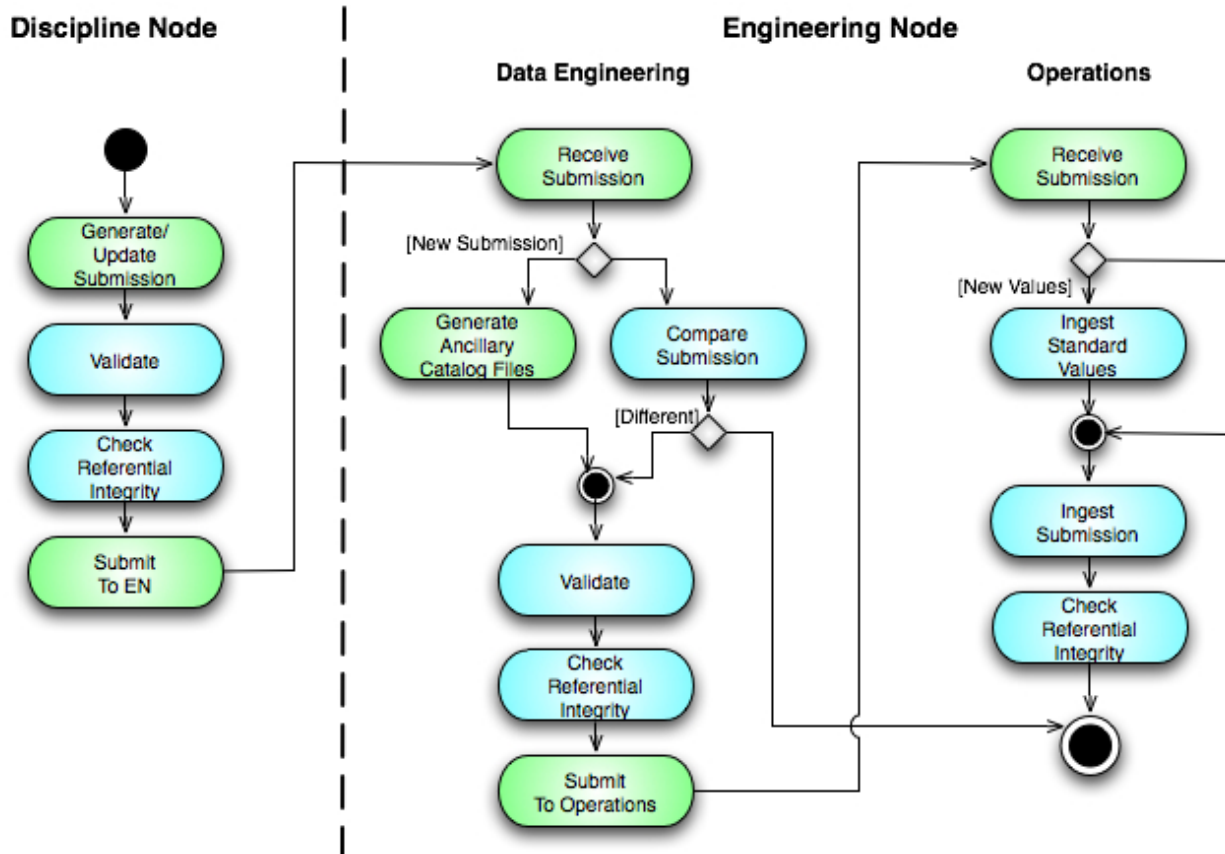
## **Release 0.1.0**

This is the first release of the Catalog Ingest Tool, intended for beta testing at the Engineering Node. This release includes the capability to compare a catalog file with another catalog file or a directory of catalog files with another directory.

## 1.3 Concept

### Concept

The Catalog Ingest Tool (CITool) has been designed to replace several tools that have been utilized over the years to support the catalog submission process. Besides simplifying the process by consolidating tools, it also provides consistency in handling of catalog files because the application is built on top of the Product Tools Library which provides reusable APIs for parsing and validating PDS product labels. The following diagram details the flow of the catalog submission process:

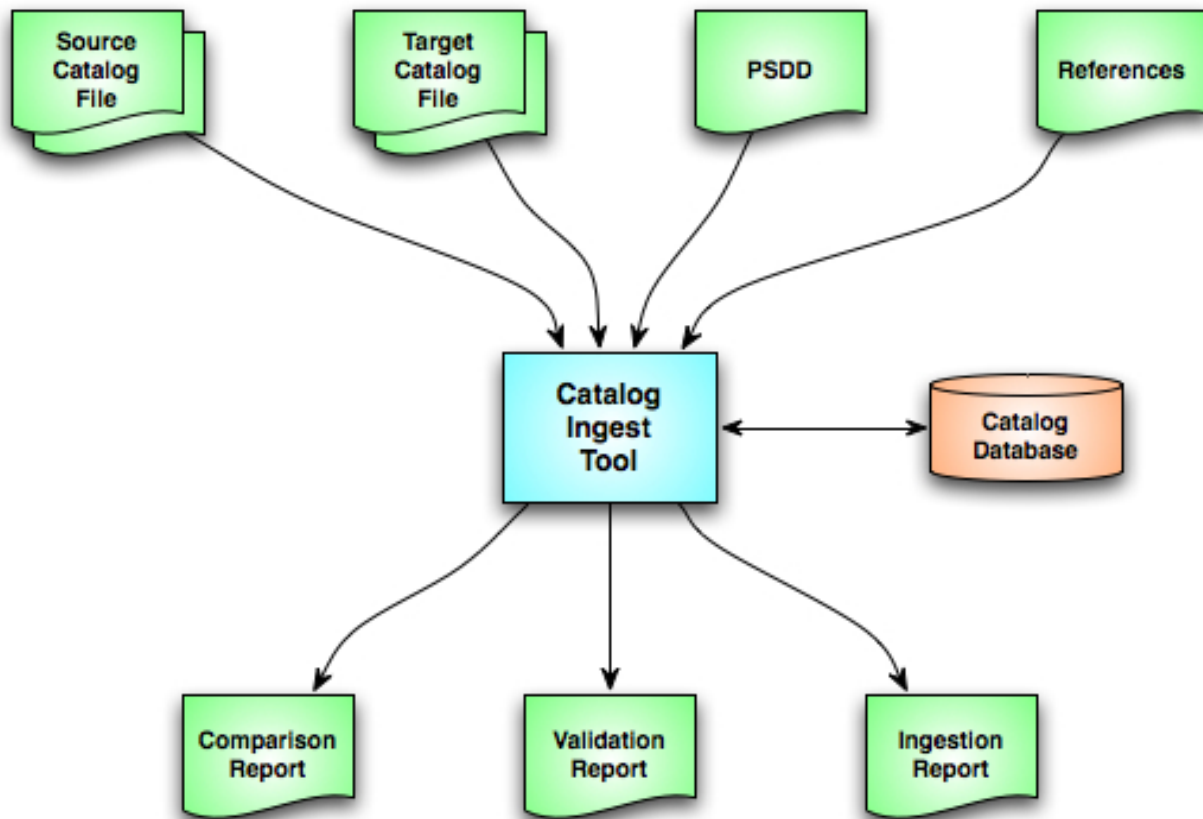


The catalog submission process is divided into three steps:

### Inputs/Outputs



The following diagram details the inputs and outputs for CITool:



The inputs and outputs for the tool are described below.

### Inputs

The application allows the specification of a target or targets for each of the modes. In the *Compare* mode, the first target specified is compared against the second target. In the *Validate* and *Ingest* modes, the specified targets are taken as a whole and considered the catalog submission.

The application also allows the user to import multiple PDS compliant data dictionaries for use during the *Validate* mode. The assumption is that the user will load one instance of the Planetary Science Data Dictionary (PSDD) and any number of local data dictionaries appropriate for the design session. In this case, the tool will validate the contents of the data dictionary files. The latest PSDD can be downloaded from the [Data Dictionary](http://pds.nasa.gov/tools/dictionary.shtml) page on the PDS web site at <http://pds.nasa.gov/tools/dictionary.shtml>.

In order to perform referential integrity checking against references contained in the catalog files, the application allows the user to import the Allrefs file. This file contains the listing of all references contained in the PDS catalog database. The latest version of this file is also available for download from the [Data Dictionary](#) page on the PDS web site. The link to the file can be found in the menu bar on the left hand side

of the page.

### **Outputs**

Each mode of the application (Compare, Validate, Ingest) produces a report in a slightly different format. The [Operation](#) section provides details on these reports and links to example reports. If viewing this document in PDF form, see the appendix for the actual examples.

The other major output of this application is the ingestion of the catalog submission into the PDS catalog database. This interface utilizes the Java Database Connectivity (JDBC) API and supports the proprietary database schema of the PDS catalog database.

## 1.4 Installation

---

### Installation

This section describes how to install the Catalog Ingest Tool (CITool) contained in the *citool* package. The following topics can be found in this section:

- [System Requirements](#)
- [Unpacking the CITool Package](#)

### System Requirements

The following sub-sections detail the system requirements for the tool.

#### Java Requirement

The Catalog Ingest Tool was developed using Java and will run on any platform with a supported Java Runtime Environment (JRE). The tool was specifically developed under Sun Java version 1.5, so the tool will execute correctly under versions 1.5 or 1.6.

Since the tool was developed using Sun's Java, this is the preferred Java environment for operation. The Sun Java package can be obtained from the [Sun Java](#) web site. Other Java environments are relatively compatible with Sun's Java.

#### Data Dictionary Requirement

Release *1r64* or later of the Planetary Science Data Dictionary (PSDD) is required for the tool to function properly. Release *1r66* of the PSDD supports the validation of explicit FILE objects. The latest version of the PDS data dictionary can be retrieved from the [PDS Data Dictionary](#) web page.

### Unpacking the CITool Package

Download the *citool* package from the [Tools & Documentation](#) web page. The binary distribution is available in identical zip or tar/gzip packages. Unpack the selected binary distribution file with one of the following commands:

```
% unzip citool-1.0.0-bin.zip
```

```
or
% tar -xzf citool-1.0.0-bin.tar.gz
```

Note: Depending on the platform, the native version of *tar* may produce an error when attempting to unpack the distribution file because many of the file paths are greater than 100 characters. If available, the GNU version of *tar* will resolve this problem. If that is not available or cannot be installed, the zipped package will work just fine in a UNIX environment.

The commands above result in the creation of the *citool-1.0.0* directory with the following directory structure:

- **README.txt**

A README file directing the user to the available documentation for the project.

- **LICENSE.txt**

The copyright notice from the [California Institute of Technology](#) detailing the restrictions regarding the use and distribution of this software. Although the license is strictly worded, the software has been classified as Technology and Software Publicly Available (TSPA) and is available for *anyone* to download and use.

- **bin/**

This directory contains batch and shell scripts for executing the tool.

- **doc/**

This document directory contains a local web site with the Catalog Ingest Guide, javadoc, unit test results and other configuration management related information. Just point your favorite browser to the *index.html* file in this directory.

- **lib/**

This directory contains the dependent jar files for the tool along with the executable jar file (citool-1.0.0.jar) containing the Catalog Ingest Tool software.

## 1.5 Operation

---

### Operation

The Catalog Ingest Tool (CITool) has three functions:

- Compare a Catalog Submission  
Comparisons of catalog submissions, either file to file or directory to directory, can be performed with a report detailing the differences.
- Validate a Catalog Submission  
Validation specific to catalog files can be performed with a report detailing the results. In addition, referential integrity checks are performed within the set of catalog files. Details on the types of checks performed can be found in the [Referential Integrity Checking](#) section. If viewing this document in PDF form, see the appendix for details.
- Ingest a Catalog Submission  
Ingestion into the central catalog database can be performed assuming the user has appropriate permissions.

The following topics can be found in this section:

- [Tool Setup](#)
- [Tool Execution](#)
- [Report Formats](#)

Note: The command-line examples in this section have been broken into multiple lines for readability. The commands should be reassembled into a single line prior to execution.

### Tool Setup

In order to execute CITool, the user's environment must first be configured appropriately. This section describes how to setup the user environment on UNIX-based and Windows machines.

#### UNIX-Based Setup

This section details the environment setup for UNIX-based machines. The preferred method is to specify the shell script, *CITool*, on the command-line. Setting the *PATH* environment variable to the location of the

script, enables the shell script to be executed from any location on the user's machine.

The following command demonstrates how to set the *PATH* environment variable, by appending to its current setting:

```
% setenv PATH ${PATH}:%HOME/citool-1.0.0/bin
```

The tool can now be executed via the shell script as demonstrated in the following example:

```
% CITool <command-line arguments>
```

Additional methods for setting up a UNIX-based environment can be found in the [UNIX Setup Options](#) section. If viewing this document in PDF form, see the appendix for details.

## Windows Setup

This section details the environment setup for Windows machines. The preferred method is to specify the batch file, *CITool.bat*, on the command-line. Setting the *PATH* environment variable to the location of the file, enables the batch file to be executed from any location on the user's machine.

The following command demonstrates how to set the *PATH* environment variable, by appending to its current setting:

```
C:\> set PATH = %PATH%;C:\citool-1.0.0\bin
```

The tool can now be executed via the batch file as demonstrated in the following example:

```
C:\> CITool <command-line arguments>
```

Additional methods for setting up a Windows environment can be found in the [Windows Setup Options](#) section. If viewing this document in PDF form, see the appendix for details.

## Tool Execution

CITool can be executed in various ways. This section describes how to run the tool, as well as its behaviors and caveats.

## Command-Line Options

CITool can be run in three modes: compare, validate and ingest. The following tables describe the command-line flag options available when each mode is enabled.

### Compare Mode

Setting the *m* flag option with *compare* will enable the tool to run in compare mode. In the compare mode, the following flag options are valid:

Flag	Description
-m, --mode compare	Specifying <i>compare</i> runs the tool in compare mode.
-t, --target <catalogs,URLs,dirs>	Explicitly specify two targets (catalog files, directories, and URLs) to compare. Targets can also be specified implicitly (example: CITool OLDDATASET.CAT, NEWDATASET.CAT).
-r, --report-file <file>	Specify the report file name. Default is standard out.
-L, --local	Do not perform directory recursion on a target directory.
-v, --verbose <1 2 3>	Specify the severity level and above to include in the report: (1=Info, 2=Warning, 3=Error). Default is warning and above (level 2).
-c, --config <file>	Specify a configuration file to set the default values.
-h, --help	Display CITool usage.
-V, --version	Display CITool version.

### Validate Mode

Setting the *m* flag option with *validate* will enable the tool to run in validate mode. In the validate mode, the following flag options are valid:

Flag	Description
-m, --mode validate	Specifying <i>validate</i> runs the tool in validate mode.
-d, --dict <.full file(s)>	Specify the Planetary Science Data Dictionary full file name and any local dictionaries.
-I, --include <path(s)>	Specify paths to search for files referenced by pointer statements in a catalog file. Separate each path with a comma. Default is to always look in the directory of the catalog file, then search the specified directory paths.
-a, --alias;	Enable aliasing. Allows the tool to properly handle object and element names defined as aliases in the Planetary Science Data Dictionary.

Flag	Description
-A, --allrefs <allrefs file>	Specify the allrefs dictionary support file or URL.
-r, --report-file <file>	Specify the report file name. Default is standard out.
-L, --local	Do not perform directory recursion on a target directory.
-t, --target <catalogs,URLs,dirs>	Explicitly specify the target (catalog file, directory, or URL). The target can also be specified implicitly (example: CITool DATASET.CAT).
-v, --verbose <1 2 3>	Specify the severity level and above to include in the report: (1=Info, 2=Warning, 3=Error). Default is warning and above (level 2).
-c, --config <file>	Specify a configuration file to set the default values.
-h, --help	Display CITool usage.
-V, --version	Display CITool version.

## Ingest Mode

Setting the *m* flag option with *ingest* will enable the tool to run in ingest mode. In the ingest mode, the following flag options are valid:

Flag	Description
-m, --mode ingest	Specifying <i>ingest</i> runs the tool in ingest mode.
-t, --target <catalog,URLs,dirs>	Explicitly specify target (catalog file, directory, and URLs) to ingest. Targets can also be specified implicitly (example: CITool -m ingest TEST.CAT).
-u, --dbuser <username>	Username of the database.
-p, --dbpass <password>	Password of the database.
-s, --dbserver <servername>	Name of the database server.
-n, --dbname <db name>	Name of the database.
-r, --report-file <file>	Specify the report file name. Default is standard out.
-v, --verbose <1 2 3>	Specify the severity level and above to include in the report: (1=Info, 2=Warning, 3=Error). Default is warning and above (level 2).
-L, --local	Do not perform directory recursion on a target directory.
-c, --config <file>	Specify a configuration file to set the default values.
-h, --help	Display CITool usage.
-V, --version	Display CITool version.

## Running CITool



This section demonstrates some of the ways that the tool can be run using the command-line option flags.

## Compare Mode

In compare mode, the tool can run in the following ways:

- Comparing Two Catalog Files

The following command demonstrates how to compare a source catalog file

*\$HOME/DIR1/DATASET.CAT* with a target catalog file *\$HOME/DIR2/DATASET.CAT*:

```
% CITool $HOME/DIR1/DATASET.CAT, $HOME/DIR2/DATASET.CAT -m compare
```

- Comparing Two Catalog Directories

The following command demonstrates how to compare a source directory, *\$HOME/DIR1*, containing a set of catalog files with a target directory, *\$HOME/DIR2*, containing another set of catalog files.

```
% CITool $HOME/DIR1, $HOME/DIR2 -m compare
```

In this example, the tool will look for files with matching file names between the source and target directory before doing a comparison. So, if a *DATASET.CAT* file is found in directory *\$HOME/DIR1*, a *DATASET.CAT* file needs to be found in directory *\$HOME/DIR2*, and so forth.

- Writing a Compare Report to File

In the first two examples above, the output report is written to standard out. The following command demonstrates how to write the report to a file named *report.txt*:

```
% CITool $HOME/DIR1/DATASET.CAT, $HOME/DIR2/DATASET.CAT -m compare -r report.txt
```

## Validate Mode

In validate mode, the tool can run in the following ways:

- Validate a Single Catalog File

The following command demonstrates how to validate a single catalog file *DATASET.CAT* against a data dictionary *pdsdd.full*:

```
% CITool DATASET.CAT -m validate -d pdsdd.full
```



- Validate a Directory of Catalog Files

The following command demonstrates how to validate a directory *\$HOME/DIR* containing a set of catalog files against a data dictionary *pdsdd.full*:

```
% CITool $HOME/DIR -m validate -d pdsdd.full
```

- Checking for Referenced Files in Different Locations

If a catalog file contains a pointer statement that references a file, the tool will always assume it is in the same location as the catalog file. If it cannot be found there, then the tool will look for that referenced file in the paths specified by the include directories option.

The following command demonstrates the validation of a catalog file that contains pointer statements to files located in a directory called *CATALOG*:

```
% CITool VOLDESC.CAT -m validate -d pdsdd.full -I $HOME/CATALOG
```

- Perform Referential Integrity Check with the Allrefs File

The allrefs dictionary support file ensures that the reference citations in the *REFS.CAT* file are consistent with what is currently in the PDS database.

The following command demonstrates how to specify an allrefs dictionary support file, *allrefs.out*, to validate and do a complete referential integrity check on references from a set of catalog files in a directory *\$HOME/DIR*.

```
% CITool $HOME/DIR -m validate -d pdsdd.full -A allrefs.out
```

- Writing the Validation Report to File

The following command demonstrates how to write a validation report to a file named *report.txt*:

```
% CITool $HOME/DIR -m validate -d pdsdd.full -A allrefs.out -r report.txt
```

## Ingest Mode

In ingest mode, the tool can run in the following ways:

- Ingest a Single Catalog File

The following command demonstrates how to ingest a single catalog file *DATASET.CAT* as a user *tempuser* with a password *temppass* into the database server *starsyb* and the database name *tempdb*.

```
% CITool DATASET.CAT -m ingest -u tempuser -p temppass -s starsyb -n tempdb
```

- Ingest a Directory of Catalog Files

The following command demonstrates how to ingest a directory *\$HOME/DIR* containing a set of catalog files as a user *tempuser* with a password *temppass* into the database server *starsyb* and the database name *tempdb*.

```
% CITool $HOME/DIR -m ingest -u tempuser -p temppass -s starsyb -n tempdb
```

## Changing Tool Behaviors Using a Configuration File

A configuration file can be passed to the tool to change its default behaviors. This provides a way to use the tool with a single flag. For more details on how to setup the configuration file see the [Using a Configuration File](#) section.

The following command demonstrates how to run the tool using a configuration file:

```
% CITool -c config.cfg
```

## Using a Configuration File

A configuration file is used to set the default behaviors of the tool. It consists of a text file made up of keyword/value pairs. The configuration file follows the syntax of the stream parsed by the Java `Properties.load(java.io.InputStream)` method.

Blank lines and lines which begin with the hash character "#" are ignored.

Values may be separated on different lines if a backslash is placed at the end of the line that continues below.

Escape sequences for special characters like a line feed, a tabulation or a unicode character, are allowed in the values and are specified in the same notation as those used in Java strings (e.g. `\n`, `\t`, `\r`).

Since backslashes (`\`) have special meanings in a configuration file, keyword values that contain this character will not be interpreted properly by CITool even if it is surrounded by quotes. A common example would be a Windows path name (e.g. `c:\VTT_EN_1-1\target`). Use the forward slash character instead (`c:/VTT_EN_1-1/target`) or escape the backslash character (`c:\\VTT_EN_1-1\\target`).

Note: Any flag specified on the command-line takes precedence over any equivalent settings placed in the configuration file.

The following table contains valid keywords that can be specified in the configuration file when running in compare mode:

Property Key	Associated Flag	Valid Value(s)
<code>citool.mode</code>	<code>-m</code>	Specify <i>compare</i> to run the tool in compare mode.
<code>citool.targets</code>	<code>-t</code>	Specify two targets (catalog files, directories, and/or URLs) to compare.
<code>citool.report</code>	<code>-r</code>	Specify the report file name. Do not specify this property key if writing to standard out.
<code>citool.recurse</code>	<code>-L</code>	Set to 'false' to disable directory recursion on a target directory. Set to 'true' otherwise or do not specify this property key.
<code>citool.verbose</code>	<code>-v</code>	Specify the severity level and above to include in the report (1=info, 2=warning, 3=error). Default is warnings and above (level 2).

The following table contains valid keywords that can be specified in the configuration file when running in validate mode:

Property Key	Associated Flag	Valid Value(s)
<code>citool.mode</code>	<code>-m</code>	Specify <i>validate</i> to run the tool in validate mode.
<code>citool.targets</code>	<code>-t</code>	Specify a target (catalog file, directory, or URL) to validate.
<code>citool.dictionaries</code>	<code>-d</code>	Specify the Planetary Science Data Dictionary full file name and any local dictionaries.
<code>citool.includepaths</code>	<code>-l</code>	Specify paths to search for files referenced by pointer statements in a catalog file.
<code>citool.recurse</code>	<code>-L</code>	Set to 'false' to disable directory recursion on a target directory. Set to 'true' otherwise or do not specify this property key.

Property Key	Associated Flag	Valid Value(s)
citool.alias	-a	Set to <i>true</i> to enable aliasing.
citool.allrefs	-A	Specify the ' <i>allrefs</i> ' dictionary file or URL.
citool.report	-r	Specify the report file name. Do not specify this property key if writing to standard out.
citool.verbose	-v	Specify the severity level and above to include in the report (1=info, 2=warning, 3=error). Default is warnings and above (level 2).

The following table contains valid keywords that can be specified in the configuration file when running in ingest mode:

Property Key	Associated Flag	Valid Value(s)
citool.mode	-m	Specify <i>ingest</i> to run the tool in ingest mode.
citool.targets	-t	Specify targets (catalog files, directories, and URLs) to ingest.
citool.dbuser	-u	Specify the username of the database.
citool.dbpass	-p	Specify the password of the database.
citool.dbserver	-s	Specify the name of the database server.
citool.dbname	-n	Specify the name of the database.
citool.report	-r	Specify the report file name. Do not specify this property key if writing to standard out.
citool.recurse	-L	Set to ' <i>false</i> ' to disable directory recursion on a target directory. Set to ' <i>true</i> ' otherwise or do not specify this property key.
citool.verbose	-v	Specify the severity level and above to include in the report (1=info, 2=warning, 3=error). Default is warnings and above (level 2).

The following example demonstrates how to set a configuration file:

```
# This is a CITool configuration file

citool.mode      = validate
citool.targets   = ./TEST_DIR
citool.report    = report.txt
citool.dictionaries = pdsdd.full
```

This is equivalent to running the tool with the following flag options:

```
-t ./TEST_DIR -m validate -r report.txt -d pdsdd.full
```

The following example demonstrates how to set a configuration file with multiple values for a property key:

```
# This is a CITool configuration file with multiple values

citool.mode      = compare
citool.targets   = DIR
citool.dictionaries = pdsdd.full, localdd.full
```

This is equivalent to running the tool with the following flags:

```
-t DIR -m compare -d pdsdd.full, localdd.full
```

The following example demonstrates how to set a configuration file with multiple values that span across multiple lines:

```
# This is a CITool configuration file with multiple values

citool.mode      = compare
citool.targets   = DIR
citool.dictionaries = pdsdd.full, \
                    localdd.full
```

The following example demonstrates how to override a setting in the configuration file.

Suppose the configuration file *config.cfg* is defined as follows:

```
# This is a CITool configuration file

citool.mode      = validate
citool.targets   = ./TEST_DIR
citool.dictionaries = pdsdd.full
```

To use another dictionary instead like *mypdsdd.full*, then the following command demonstrates how to perform

this behavior change:

```
% CITool -c config.cfg -d mypdsdd.full
```

## Report Formats

This section describes the contents of the CITool report formats. The links below detail the results. If viewing this document in PDF form, see the appendix for the actual examples.

The tool has a different report format depending on the CITool mode.

### Compare Report

In a [compare](#) report, the location, severity, and textual representation of the differences between a source and target catalog file are shown. The differences being reported are about the target catalog file with respect to the source. A 'SAME' or 'DIFFERENT' keyword is displayed next to each target to indicate when a target file is identical or different, respectively, against its source.

### Validate Report

In a [validate](#) report, the location, severity, and textual description of each detected anomaly is reported. A 'PASS', 'FAIL', or 'SKIP' keyword is displayed next to each file to indicate when a file has passed, failed, or skipped PDS validation, respectively. In addition, anomalies against the referential integrity of the catalog files are reported.

### Ingest Report

In a [ingest](#) report, the location, severity, and textual description of each detected anomaly is reported. Completion or failure of each catalog file ingestion is reported. Detailed information is displayed when the catalog ingestion has failed.

## 1.6 Appendix A - UNIX Setup Options

---

### UNIX Setup Options

This section details a couple of options for setting up a UNIX environment for launching CITool.

#### Specify the CLASSPATH on the Command-Line

An alternative method to setting the *CLASSPATH* variable with all of the tool's dependent JAR files is to specify the *java.ext.dirs* Java property on the command-line when running the tool each time. This is done by passing the property via the Java "-D" flag as demonstrated in the following example:

```
% java -Djava.ext.dirs=$HOME/citool-1.0.0/lib \  
gov.nasa.pds.citool.CITool <command-line arguments>
```

#### Specify the JAR on the Command-Line

Another alternative method is to specify the executable JAR file on the command-line when running the tool each time. This is done by passing the JAR file specification via the Java "-jar" flag as demonstrated in the following example:

```
% java -jar \  
$HOME/citool-1.0.0/lib/citool-1.0.0.jar <command-line arguments>
```



## 1.7 Appendix B - Windows Setup Options

---

### Windows Setup Options

This section details a couple of options for setting up a Windows environment for launching CITool.

#### Specify the CLASSPATH on the Command-Line

An alternative method to setting the *CLASSPATH* variable with all of the tool's dependent JAR files is to specify the *java.ext.dirs* Java property on the command-line when running the tool each time. This is done by passing the property via the Java "-D" flag as demonstrated in the following example:

```
C:\> java -Djava.ext.dirs=c:\citool-1.0.0\lib \
gov.nasa.pds.citool.CITool <command-line arguments>
```

#### Specify the JAR on the Command-Line

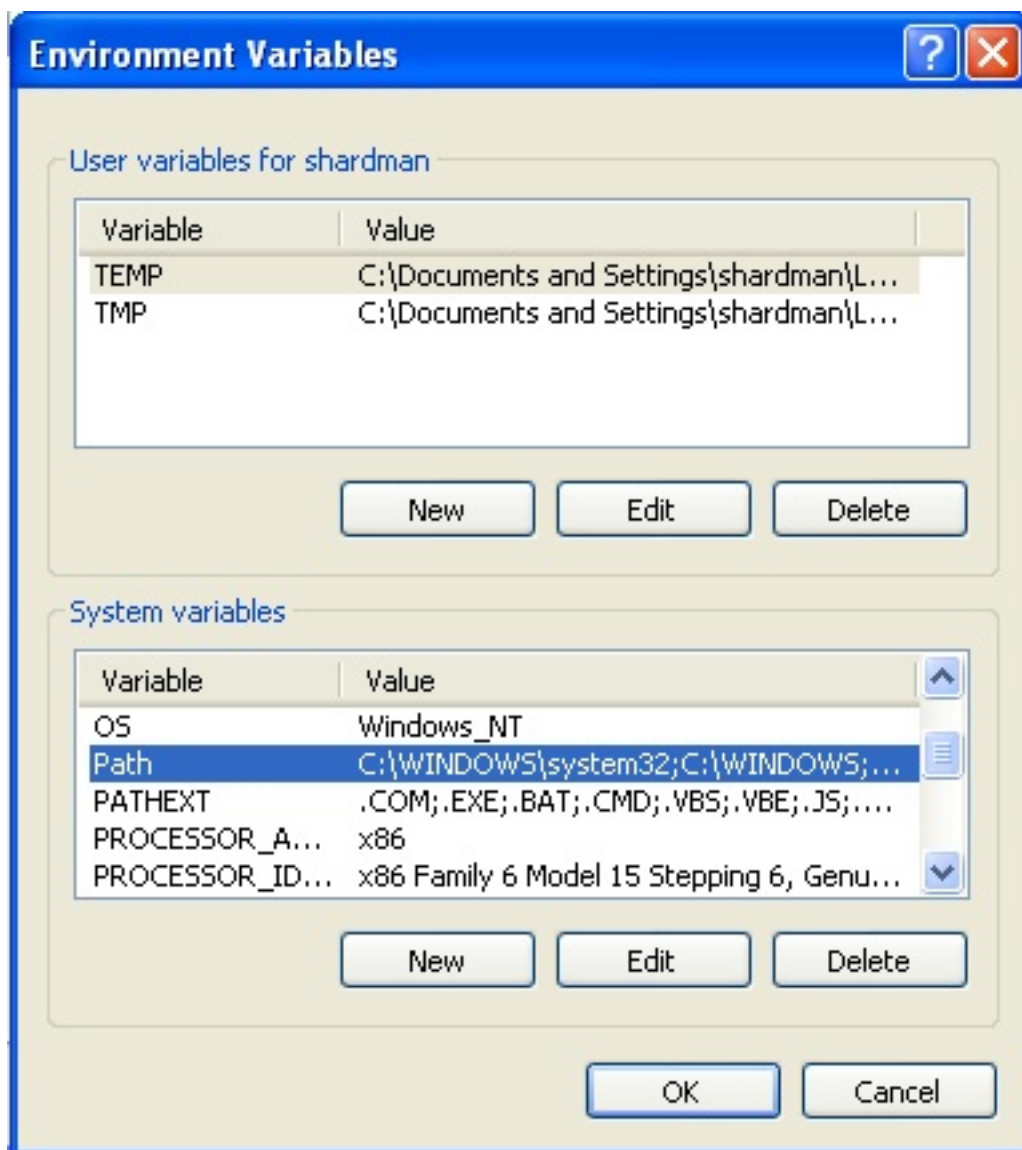
Another alternative method is to specify the executable JAR file on the command-line when running the tool each time. This is done by passing the JAR file specification via the Java "-jar" flag as demonstrated in the following example:

```
C:\> java -jar \
c:\citool-1.0.0\lib\citool-1.0.0.jar <command-line arguments>
```

#### Specify the Path in the Control Panel

The method for setting the executable path permanently for CITool is to set the *Path* environment variable via the control panel as follows:

- Right-click on *My Computer* icon on your desktop and select the *Properties* menu item.
- Navigate to the *Advanced* tab and select the *Environment Variables* button. At this point, you should now see a window like the one below:



- Highlight the *Path* variable in the System Variables list and select the Edit button.
- Append to the current contents of the variable, the path to the *bin* directory within *citool* package. Separate the package path from the current contents of the variable with a semicolon.
- Select the *OK* button when you are finished editing the *Path* variable, then select the *OK* button at the Environment Variables window to apply the changes.

Note: If you already have a DOS window open, you will need to close and re-open the window for the *Path* changes to take effect.

## 1.8 Appendix C - Referential Integrity Checking

---

### Referential Integrity Checking

This section describes how CITool checks for referential integrity among the catalog files.

Referential integrity is a database concept that ensures that relationships between the tables remain consistent. In CITool, referential integrity is performed during validation mode. The following referential integrity checks are performed:

- [Reference Integrity](#)
- [Personnel Integrity](#)
- [Target Integrity](#)
- [Instrument Integrity](#)
- [Instrument Host Integrity](#)
- [Mission Integrity](#)
- [Data Set Integrity](#)

Note, however, that not all referential integrity checks can be performed if the set of catalog files that were passed into the tool is incomplete. For example, if a *TARGET.CAT* file is not passed into the tool, then target integrity checks will not be performed.

#### Reference Integrity

In a reference integrity check, CITool will ensure that referential integrity exists in the use of *REFERENCE\_KEY\_ID* and *REFERENCE\_DESC* as follows:

- Each *REFERENCE\_KEY\_ID* value specified in a catalog file other than the *REF.CAT* file exists in the *REF.CAT* file.
- Each *REFERENCE\_KEY\_ID* value specified in the *REF.CAT* file is specified in at least one other catalog file.
- Each *REFERENCE\_KEY\_ID* value exists in the '*allrefs*' dictionary support file.
- Each *REFERENCE\_DESC* value specified in the *REF.CAT* file matches the corresponding value in the '*allrefs*' dictionary support file.

#### Personnel Integrity

In a personnel integrity check, CITool will ensure that referential integrity exists in the use of *PDS\_USER\_ID* and *NODE\_ID* as follows:

- Each *PDS\_USER\_ID* and *NODE\_ID* value specified in a catalog file other than the *PERSON.CAT* file exists in the *PERSON.CAT* file.

### Target Integrity

In a target integrity check, CITool will ensure that referential integrity exists in the use of *TARGET\_NAME* as follows:

- Each *TARGET\_NAME* value specified in a catalog file other than the *TARGET.CAT* file exists in the *TARGET.CAT* file.
- Each *TARGET\_NAME* value specified in the *TARGET.CAT* file is specified in at least one other catalog file.

### Instrument Integrity

In an instrument integrity check, CITool will ensure that referential integrity exists in the use of *INSTRUMENT\_ID* as follows:

- Each *INSTRUMENT\_ID* value specified in a catalog file other than the *INSTRUMENT.CAT* file exists in the *INSTRUMENT.CAT* file.
- Each *INSTRUMENT\_ID* value specified in the *INSTRUMENT.CAT* file is specified in at least one other catalog file.

### Instrument Host Integrity

In an instrument host integrity check, CITool will ensure that referential integrity exists in the use of *INSTRUMENT\_HOST\_ID* as follows:

- Each *INSTRUMENT\_HOST\_ID* value specified in a catalog file other than the *INSTRUMENT.CAT* file, exists in the *INSTRUMENT.CAT* file.
- Each *INSTRUMENT\_HOST\_ID* value specified in the *INSTRUMENT.CAT* file is specified in at least one other catalog file.

### Mission Integrity

In a mission integrity check, CITool will ensure that referential integrity exists in the use of *MISSION\_NAME* as follows:

- Each *MISSION\_NAME* value specified in a catalog file other than the *MISSION.CAT* file exists in the *MISSION.CAT* file.
- Each *MISSION\_NAME* value specified in the *MISSION.CAT* file is specified in at least one other catalog file.

### Data Set Integrity

In a data set integrity check, CITool will ensure that referential integrity exists in the use of *DATA\_SET\_ID* and *DATA\_SET\_ID\_COLLECTION\_ID* as follows:

- Each *DATA\_SET\_ID* value specified in a catalog file other than the *DATASET.CAT* file exists in the *DATASET.CAT* file.
- Each *DATA\_SET\_ID* value specified in the *DATASET.CAT* file is specified in at least one other catalog file.
- The *DATA\_SET\_COLL\_OR\_DATA\_SET\_ID* value specified in the *VOLUME.CAT* and *NSSDC.CAT* files is consistent with the *DATA\_SET\_ID* or *DATA\_SET\_COLLECTION\_ID* value specified in the *DATASET.CAT* file.
- The *PRODUCT\_DATA\_SET\_ID* value specified in the *INVENTORY.CAT* file is consistent with the *DATA\_SET\_ID* or *DATA\_SET\_COLLECTION\_ID* value specified in the *DATASET.CAT* file.

## 1.9 **Appendix D - Report Examples**

---

### **Report Examples**

This section details the various report formats available from the Catalog Ingest Tool.

## 1.9.1 Compare Report

---

### Compare Report

The following is an example of a compare report:

```
PDS Catalog Ingestion Tool Report

Version          1.0.0
Date             Tue, Jan 05 2010 at 08:12:46 AM

Parameter Settings:

Mode             Compare
Target(s)
  Source = file:/C:/catalog-files/diff/dir1/
  Target = file:/C:/catalog-files/diff/dir2/
Directory Recursion  true
Severity Level      WARNING

Summary:

[X] Differences Found
[ ] No Differences Found

Compare Details:

DIFFERENT: file:/C:/catalog-files/diff/dir2/DATASET.CAT
  line 2: Element RECORD_TYPE has different value than source
    Source: line 13 of file:/C:/catalog-files/diff/dir1/DATASET.CAT
    13c2
    < FIXED_LENGTH
    ----
    > STREAM

  line 3: Element LABEL_REVISION_NOTE has different value than source
    Source: line 2 of file:/C:/catalog-files/diff/dir1/DATASET.CAT
    12d13
    < 2005-07-29 RS:simpson incorporated aerobraking data
    ----

  line 16: Element DATA_SET_ID has different value than source
    Source: line 17 of file:/C:/catalog-files/diff/dir1/DATASET.CAT
    17c16
    < ODY-M-RSS-1-RAW-V1.0
    ----
    > ODY-M-RSS-1-RAW-V2.0
```

```
line 31: Element ARCHIVE_STATUS not present in source
line 28: Element DATA_SET_TERSE_DESC not present in source
line 32: Element CITATION_DESC not present in source
line 786: Object DATA_SET_TARGET not present in source
```

SAME: file:/C:/catalog-files/diff/dir2/MISSION.CAT

End of Report



## 1.9.2 Validate Report

---

### Validate Report

The following is an example of a validate report:

```
PDS Catalog Ingestion Tool Report

Version          1.0.0
Date             Tue, Jan 05 2010 at 08:24:24 AM
Dictionary version
/* Planetary Science Data Dictionary database dump */
/* Start of alias definitions */
/* Version: OPS */
/* Online Database: pdscat1r76 */
/* Generated: Wed Nov 11 07:39:27 2009 */

Parameter Settings:

Mode             Validate
Target           file:/C:/catalog-files/goodcatfiles/
Directory Recursion true
Dictionary File(s) [c:\pdsdd\pdsdd.full]
Allrefs File     file:/C:/catalog-files/goodcatfiles/allrefs_1r74.out
Severity Level   WARNING
Report File      report-validate.txt

Summary:

5 of 5 validated, 0 skipped
5 of 5 passed

4 referential integrity check(s) made, 0 skipped
2 of 4 passed

0 new standard value(s) found

Validation Details:

PASS: file:/C:/catalog-files/goodcatfiles/MK88-120DS.CAT

PASS: file:/C:/catalog-files/goodcatfiles/MK88-120DS_REF.CAT
WARNING line 17: Could not find year at the end of the citation. Found 557-578.
WARNING line 24: Could not find year at the end of the citation. Found
```

```

http://www.ifa.hawaii.edu/88inch/manuals/user.pdf.
  WARNING line 24: REFERENCE_DESC value does not start with UHGUIDE
  WARNING line 31: Could not find year at the end of the citation. Found #2059.

PASS: file:/C:/catalog-files/goodcatfiles/MK88-120_INM.CAT

PASS: file:/C:/catalog-files/goodcatfiles/MK88_INST.CAT

PASS: file:/C:/catalog-files/goodcatfiles/MK88_INSTHOST.CAT

Referential Integrity Details:

FAIL: Data Set Integrity
  Parent File(s): [MK88-120DS.CAT]
  Begin checking children
  End checking children
    line 7: MK88-120DS.CAT: DATA_SET_ID = MK88-L-120CVF-3-RDR-120COLOR-V1.0 is not
specified in a non DATASET.CAT.

FAIL: Reference Integrity
  Parent File(s): [MK88-120DS_REF.CAT]
  line 17: Description for PIETERS1986 does not match description in allrefs
dictionary.
    Source: allrefs_lr74.out
    < Pieters, C. M. (1986), Composition of the Lunar Highland Crust from
Near-Infrared Spectroscopy, Reviews of Geophysics, Vol. 24, No. 3, 557-578.
    ----
    > Pieters, C. M. (1986), Composition of the Lunar Highland Crust from
Near-Infrared Spectroscopy Reviews of Geophysics, Vol. 24, No. 3, 557-578.

    line 8: MCCORDETAL not found in allrefs dictionary.
  Begin checking children
    line 102: MK88-120DS.CAT: REFERENCE_KEY_ID = MCCORDETAL1981 is not found in a
REF.CAT.
    line 33: MK88_INST.CAT: REFERENCE_KEY_ID = MCCORDETAL1981 is not found in a
REF.CAT.
  End checking children

PASS: Instrument Integrity
  Parent File(s): [MK88_INST.CAT]
  Begin checking children
  End checking children

PASS: Instrument Host Integrity
  Parent File(s): [MK88_INSTHOST.CAT]
  Begin checking children
  End checking children

End of Report

```

## 1.9.3 Ingest Report

---

### Ingest Report

The following is an example of ingest report:

```
PDS Catalog Ingestion Tool Report

Version          1.0.0
Date             Tue, Jan 05 2010 at 08:39:32 AM

Parameter Settings:

Mode             Ingest
Target           file:/C:/catalog-files/goodcatfiles/
Directory Recursion true
Severity Level   WARNING
Report File      report-ingest.txt

Summary:

Catalog Ingestion is completed.
  Number of processed files: 5
  Number of successful ingestion to the table: 20
  Number of failed ingestion to the table: 2
  Number of new standard values ingested: 1

Catalog Ingestion Details:

WARNING: DSINFO - ARCHIVE_STATUS: null
Processed: file:/C:/catalog-files/goodcatfiles/MK88-120DS.CAT

Processed: file:/C:/catalog-files/goodcatfiles/MK88-120DS_REF.CAT

Processed: file:/C:/catalog-files/goodcatfiles/MK88-120_INM.CAT

Processed: file:/C:/catalog-files/goodcatfiles/MK88_INST.CAT

ABORT: HOSTD - the keyword INSTRUMENT_HOST_DESC was not found.
          - INSTRUMENT_HOST_ID: MK88
ABORT: HOSTINFO - one or more of the required keywords was not found.
          - INSTRUMENT_HOST_ID: MK88
          - INSTRUMENT_HOST_NAME: null
          - INSTRUMENT_HOST_TYPE: null
          - LABEL_REVISION_NOTE: S. SLAVNEY, 1999-07-27
Failed to ingest one or more catalog object(s):
file:/C:/catalog-files/goodcatfiles/MK88_INSTHOST.CAT
```

End of Report