Government Notices • Goewermentskennisgewings

DEPARTMENT OF HIGHER EDUCATION AND TRAINING

NO. 709 **14 AUGUST 2015**

STANDARDS FOR SUBMISSION, PROCESSING AND ARCHIVING: DHET007

I, Bonginkosi Emmanuel Nzimande, MP, Minister of Higher Education and Training, in terms of section 41B (1) and (4) of the Continuing Education and Training Act, 2006 (Act No. 16 of 2006); read with section 3 of the Higher Education Act, 1997 (Act No. 101 of 1997); section 8(2)(b) of the National Qualifications Framework Act, 2008 (Act No. 67 of 2008); section 5(1)(a)(i) of Skills Development Act, 1998 (Act No. 97 of 1998) and the Higher Education and Training Information Policy published in Notice No.832, Government Gazette No.36973 of 01 November 2013, hereby publish DHET 007: Submission, Processing and Archiving Standard as set out in the Schedule.

DR BE NZIMANDE, MP
MINISTER OF HIGHER EDUCATION AND TRAINING
DATE: 21/07/2015

Version: 1.00

Standard for the submission, processing and archiving of data submissions received by non-legacy Department of Higher Education and Training Management Information Systems Version 1.00

Acronyms

DHET	Department of Higher Education and Training
DHET MIS	Department of Higher Education and Training Management Information System developed after 2012
DMZ	Demilitarized zone
HETMIS	Higher Education and Training Management Information System
LAN	Local Area Network
SSH	Secure Shell

Definition of Terms

"archiving of data" – the process of moving data that is no longer actively used to a separate data storage device for long-term storage

"archive server" – a server on which data is archived

"batch processing data files" - input data collected into batches of files and processed in batches

"data supplier" – any entity or organization that is permitted to submit data to the relevant recipient DHET MIS

"data validation utility" – software programme that validates data to ensure that the data conforms to the minimum data requirements of the recipient DHET MIS

"database server" – a server that hosts a database

"staging server" – an interim or quality assurance server that is separate from the development or production environment of the system

"SSH server" – a server which uses the secure shell protocol to accept connections from remote computers

1   Purpose of the standard

The purpose of this standard is to provide a framework for the overall submission, processing and archiving of batch processing data files received by Department of Higher Education and Training Management Information Systems (DHET MISs) developed after 2012 that form part of the overall implementation of the Higher Education and Training Management Information System (HETMIS).

The standard will ensure that the submission to, and processing and archiving of all batch processing data file data submissions submitted to DHET MISs is completed in a standardized manner that ensures the overall integrity, security and business continuity of the DHET MIS.

# 2   Scope and applicability of the standard

This standard applies to all DHET MIS that form part of the overall implementation of HETMIS.

3   Submission, processing and archiving of data submission

Data tables contained within the DHET MISs will be populated by batch processing data files submitted by data suppliers. The following standardized submission, processing and archiving process will be implemented for all DHET MISs:

3.1   Extracting, testing and submission of data from the data supplier

The following section describes the processes completed by the data supplier when submitting data to a DHET MIS.

Figure 3.1.1
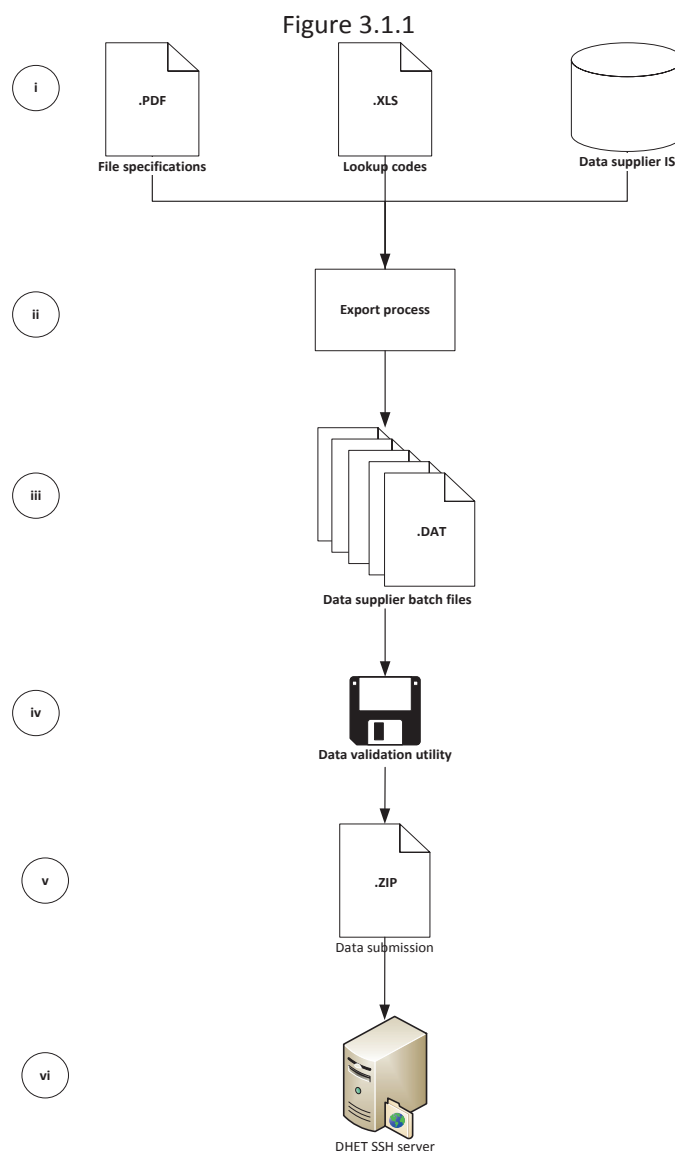
i.   The data supplier has an information system that allows the data supplier to capture and store information related to the DHET MIS. The data supplier may opt to purchase a commercial information system or to develop their own information system. The type/nature of the

information system at the data supplier is of no concern to the Department of Higher Education (DHET). The data supplier must however ensure that the information system is able to extract the relevant data in the required format as specified in the file specifications for the recipient DHET MIS.

ii.　The data supplier information system must have an export functionality which exports the data required by the recipient DHET MIS in the format described in the file specification for the DHET MIS.

iii.　The data supplier information system export functionality creates the required files, with the required naming convention for submission to the recipient DHET MIS.

iv.　The user at the data supplier validates the data submission using a data validation utility. The data validation utility is provided to the data supplier free of charge by the DHET.

v.　If the data submission completely conforms to the layout, content and business rules of the DHET MIS file specification the data validation utility completes the following steps:
   a.　Creates an aggregated and detailed report which describes the results of the data validation processes.
   b.　Creates a version file which describes the version of the data validation software used.
   c.　Creates a contact file which describes the user and the e-mail address at which the user can be contacted.
   d.　Creates a compressed copy of the data submission files, the data validation reports, the version file and the contact file. The compressed data submission file is encrypted with a password that is not known to the user of the data validation software.
   e.　Indicates to the user that the submission is ready for transfer to the DHET SSH server. If the user indicates that the submission may be copied to the DHET SSH server then the data validation software copies the submission using the SCP protocol. The data validation software connects to the server using a login and password. The login and password of the DHET SSH server are unknown to the data validation utility user. The data validation utility provides feedback to the user in regard to the success of the transfer of the submission to the DHET SSH server.
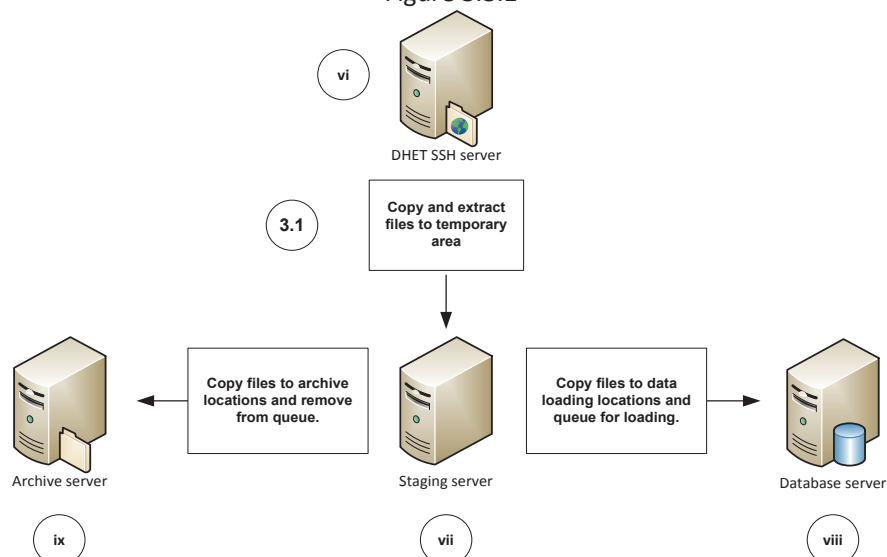
3.2　　DHET SSH server

Figure 3.2.1

　　　　　　　　　　3
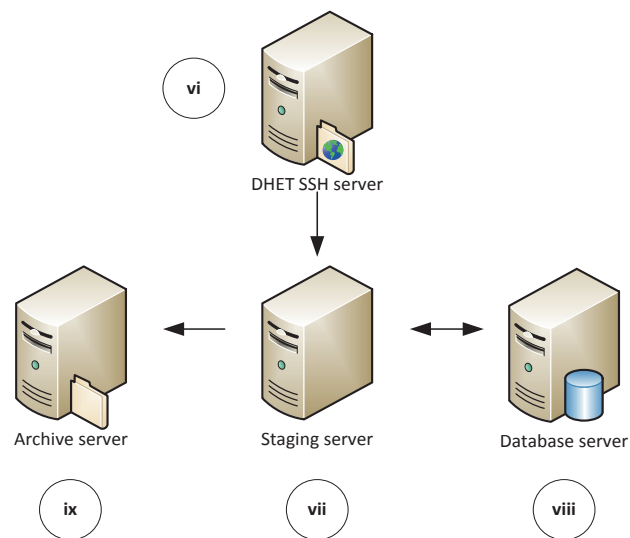
vi.   The DHET SSH server is a server that is situated in the DMZ of the DHET LAN. The server is used only as a temporary storage area for data submissions that are transmitted via the data validation utility to the server. The server is configured to accept SCP protocol transfers of data submissions from a single user account per DHET MIS which is password protected. The logins and passwords used to transmit data submissions are known only to the administrator of the HETMIS system. The DHET SSH server is accessible from a Staging server (see Section 3.3) within the DHET LAN; however the DHET LAN is not accessible from the DHET SSH server.

3.3   Staging server and components

Figure 3.3.1

vii.   The Staging server is inside the DHET LAN. Connections to the DHET SSH server from within the DHET LAN can be made. The Staging server has a software component that completes the following steps as an automated process. This software component is executed on a daily basis as an automated task (in other words no human intervention is required after the task has been activated). The time and frequency at which the software component is executed is customizable.

All data submissions found on the DHET SSH server are copied and extracted into a temporary area on the Staging server. This process includes the following types of discrete steps:
a.   Any data submissions that do not conform to the required naming convention are automatically deleted. The software component saves the names of any data submissions that have been deleted to a log file.
b.   The data submissions are extracted using a password. Any data submissions that are not password protected or cannot be opened using the expected password are automatically deleted. The software component saves the names of any data submissions that have been deleted to a log file.
c.   If any of the extracted files have a naming convention that does not conform to the expected naming convention then all of the files from the specific submission are deleted. The software component saves the names of any files that have been deleted to a log file.

All extracted data files are copied to the relevant data storage locations and the recipient database data loading queue is updated. This process includes the following discrete steps:
a.   Copy all archive related files (item v. a, b and c) to an archive location on the Archive server.
b.   Copy all data related files (item v. d) to a data loading location on the Database server.
c.   Connect to the recipient database and update the data loading queue in regard to item b above. In addition determine all data related files that have been loaded in the previous data loading cycle.
d.   Copy all of the data files that have already been loaded (see item c above) to the archive location on the Archive server.

3.4   Database server and components

Figure 3.4.1
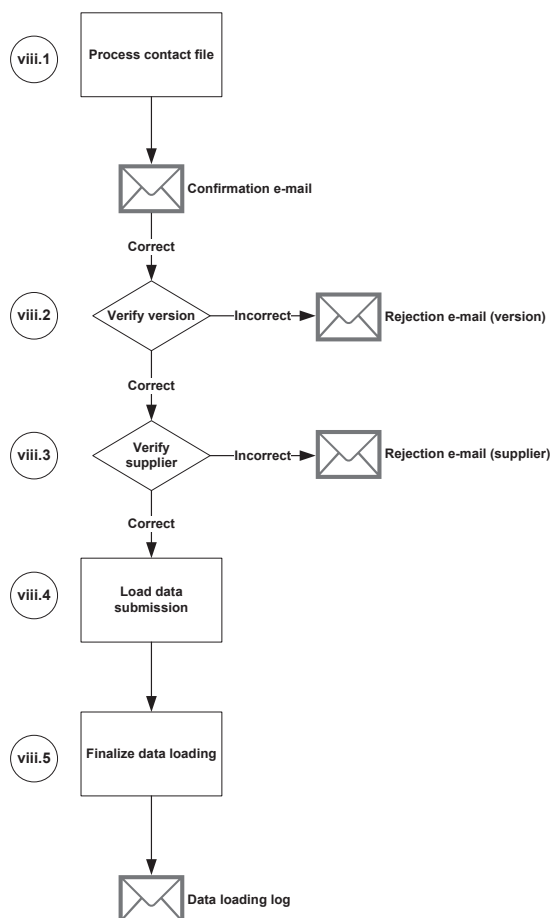
viii. The Database server is inside the DHET LAN[*]. The Database server has a software component that completes the following steps as an automated process. This software component is executed on a daily basis as an automated task (in other words no human intervention is required after the task has activated). The time at which the software component executes on a daily basis is customizable and is set to start after all of the new data submissions found on the SSH server have been copied to the Staging server and the loading queues have been updated (see Section 3.3).

The following steps are completed per data submission in logical order of date of submission and data supplier.

viii.1 The contact file (see item v. c) of the submission is processed and an e-mail is sent to the data supplier indicating that the submission is being processed.

viii.2 The version file (see item v. b) of the submission is processed and the version of the data submission is confirmed. If the data submission version is lower than the expected version then the process is aborted and an e-mail is sent to the data supplier indicating that the

---

[*] Although the Archive server, Database server and Staging server are described as logically separate objects they may in fact be the same server.

data submission was rejected. The same e-mail message is sent to the DHET MIS administrator.

viii.3   The supplier of the data submission is verified to ensure that the data supplier may submit data to the specific information system. If the data supplier is not recognized then the process is aborted and an e-mail is sent to the data supplier indicating that the data submission was rejected. The same e-mail message is sent to the DHET MIS administrator.

viii.4   The data submission is loaded into temporary data tables after which a final validation of the data is completed and the data is loaded into the data tables of the DHET MIS. A log is created which contains the following types of log records:

   a.   Detailed log records for all records that generate a log message. The log record contains the following type of information:

      a.   Data submission name – which contains the data submission filename
      b.   Data supplier code – which contains the code of the data supplier
      c.   Log Type – which contains values like High, Medium or Low
      d.   Log Message – which provides a detailed message
      e.   Description – which provides additional information for the message
      f.   Log Date – the date and time at which the log message was generated.

   b.   A summary record that details the number of records rejected, inserted, updated and unchanged, and the total number of records received in the data submission. The Log Type for this record is "Low".

viii.5   The load for the specific submission is finalized by completing the following steps:

a.   All temporary tables are deleted.

b.   The data loading queue is updated.

c.   The data loading log file is sent to the data supplier. The same e-mail message is sent to the DHET MIS administrator.

Steps viii.1 to viii.5 are repeated until the data loading queue has been cleared.