Keeping Emulation Environments Portable

FP7-ICT-231954

Requirement specification for the Transfer Tools Framework

Deliverable number	*D1.2a*
Nature	***R**eport*
Dissemination level	*PU*
Delivery date	**Due:** *15th October 2010* **Actual:** *15th October 2010 (updated end of M29)*
Status	*Finalised*
Work package number	*WP1*
Lead beneficiary	*DNB*
Author(s)	*Marcus Dindorf, DNB* *David Michel, TESS*

Document history

Revisions

Version	Date	Author	Changes
0.1	*04th December 2009*	*Marcus Dindorf [DNB]*	*Created initial version.*
0.2	*13th January 2010*	*Marcus Dindorf [DNB]*	*Added section 'Requirements', 'Executive Summary', 'General description' and the glossary.*
0.3	*04th February 2010*	*Marcus Dindorf [DNB]*	*Updated section 'General description'.*
0.4	*12th February 2010*	*Marcus Dindorf [DNB]*	*Added section 'System context'.*
0.5	*19th February 2010*	*Marcus Dindorf [DNB]*	*Added description of primarily scenarios & diagrams.*
0.6	*23th February 2010*	*Marcus Dindorf [DNB]*	*Updated description of primarily scenarios & diagrams.*
0.7	*23th February 2010*	*Marcus Dindorf [DNB]*	*Minor update – added missing header and footer information on several pages.*
0.8	*01st March 2010*	*Marcus Dindorf [DNB]*	*Added diagram with system context and boundaries of the TTF. Added introduction to the key elements of the TTF.*
0.9	*03rd March 2010*	*Marcus Dindorf [DNB]*	*Added chapter System Overview. Updated URD based on comments from [TESS] & [DNB].*
1.0	*05th March 2010*	*Marcus Dindorf [DNB]*	*Updated URD based on comments from [CSM].*
1.1	*09th March 2010*	*Marcus Dindorf [DNB]*	*Updated URD based on comments from KB. Updated Abbreviations and Glossary.*
1.2	*11th March 2010*	*Marcus Dindorf [DNB]*	*Updated URD based on comments from [TESS].*
1.3	*16th March 2010*	*Marcus Dindorf [DNB]*	*Re-written use-case scenarios.*
1.4	*16th March 2010*	*Marcus Dindorf [DNB]*	*Added example introduction to Scenarios.*
1.5	*17th March 2010*	*Marcus Dindorf [DNB]*	*Updated some diagrams.*
1.6	*17th March 2010*	*Marcus Dindorf [DNB]*	*Minor improvements.*
1.7	*18th March 2010*	*Marcus Dindorf [DNB]*	*Improved scenarios and diagrams with more details about the transfer process.*

1.8	*18th March 2010*	*Marcus Dindorf [DNB]*	*Updated chapter 'System Context'. Re-written chapter 5 Use-Cases. Some minor improvements. Added chapter 'References'.*
1.9	*03rd August 2010*	*Marcus Dindorf [DNB] & David Michel [TESS]*	*Updated the requirements in chapter 'Requirements'.*
2.0	*06th August 2010*	*Marcus Dindorf [DNB]*	*Updated the requirements in chapter 'Requirements'.*
2.1	*03rd September 2010*	*Marcus Dindorf [DNB]*	*Updated the Use-Cases & Scenarios in chapter 'Use-Cases'.*
2.2	*13th September 2010*	*David Michel [TESS]*	*Complete restructuring and rewrite of all sections.*
2.3	*29th September 2010*	*Marcus Dindorf [DNB]*	*Updated all sections.*
2.4	*30th September 2010*	*Marcus Dindorf [DNB]*	*Added Executive Summary.*
2.5	*01st October 2010*	*David Michel [TESS]*	*Corrected typos, grammar.*
2.6	*02nd October 2010*	*David Michel [TESS]*	*Rewrite of introduction and executive summary.*
2.7	*03rd October 2010*	*Marcus Dindorf [DNB]*	*Rewrite of executive summary, updated Chapter Use-Case Scenarios and minor improvements.*
2.8	*12th October 2010*	*Marcus Dindorf [DNB]*	*Minor improvements and updates based on comments from reviewer.*
2.9	*15th October 2010*	*Marcus Dindorf [DNB]*	*Minor corrections and improvements.*
3.0	*22nd July 2011*	*Marcus Dindorf [DNB]*	*Revised the complete document to improve the writing style, grammar, and correct typos.*
4.0	*25th July 2011*	*Marcus Dindorf [DNB]*	*Revised following final review.*

Reviews

Date	Name	Result
23th February 2010	*Bram Lohman [TESS]*	*Major changes suggested.*
23th February 2010	*Tobias Steinke [DNB]*	*Major changes suggested.*
23th February 2010	*Andreas Lange [CSM]*	*Minor changes suggested.*
24th February 2010	*Jeffrey van der Hoeven [KB]*	*Major changes suggested.*
03rd August 2010	*David Michel [TESS]*	*Complete redesign suggested.*
06th October 2010	*Bram Lohman [TESS]*	*Minor changes and additions suggested.*
08th October 2010	*Jean-Philippe Humblot [BNF]*	*Minor changes suggested.*
23rd July 2011	*Clive Billenness [UPHEC]*	*Linguistic corrections and minor changes suggested.*

Signature/Approval

Approved by (signature)	Date

Accepted by at European Commission (signature)	Date

Executive Summary

This document describes the user requirements for the Transfer Tools Framework (TTF) and establishes the various features to be implemented during development.

The main goal of the TTF is to generate media carrier images and to generate descriptive metadata and optional technical metadata from given digital media carrier. To achieve this, appropriate transfer tools (as well as supportive tools and services for analysis or metadata generation) will be integrated for use within the TTF.

On the one hand, the TTF generates descriptive metadata that characterises the media carrier. On the other hand, the support of technical metadata that describes the generated media carrier image itself is optional for the scope of the TTF and depends on available tools, services or external registries integrated within the TTF. The framework is operated by end-users using a GUI but can also be controlled via a machine interface, to allow external systems to use the services provided by the TTF directly.

However, operating the TTF via a GUI would be the most comfortable way for the end-users to benefit from the services offered by the TTF for its end-users. End-users range from librarians with no or little IT knowledge to IT administrators with more knowledge of transfer tools and media transfer. For the former, the TTF possesses integrated media transfer knowledge. Based on this knowledge, the TTF provides assistance during the generation of media carrier images. This ability of the TTF, besides enabling the generation of media carrier images from a wide range of digital media carrier highlights the versatility of the framework.

The TTF assumes the presence of appropriate hardware readers and drivers for the particular media carrier types to be transferred. The TTF is envisaged to ultimately support the transfer of the original bit-stream of any given digital media carrier type but in the initial version of the TTF it will only come with built-in support for 3 1/2” floppy disks and CD-ROMs; these have been selected as the two digital media carrier to be integrated for the proof of concept within the KEEP project.

Research has shown that large volumes of digital media carriers exist in the collections of memory institutions, many of which are at very high risk of data loss and therefore in urgent need of long-term preservation. But from a legal perspective the act of media transfer is a very difficult and sensitive topic, and legal advice received in the project has advised caution about this. The provisions of national law regarding computer media copyright in place to protect the rights of the authors is a potential barrier to cultural heritage institutions in preserving national digital cultural heritage.

One future possibility would be to provide legal exceptions for cultural heritage institutions to allow them the bypass of copy-protected media carrier to prevent loss of digital cultural heritage. However, unless the legal framework is changed, a preservation strategy will not be able to take advantage of available transfer tools which would be capable of circumventing copy-protection mechanism.

Despite this, the transfer of obsolete media carriers remains both a major challenge and also an essential requirement within long-term preservation strategies for memory institutions and the need for a transfer tool framework solution remains crucial and should not be abandoned. That said, only legal tools should be used in any implementation of the TTF unless a specific exception exists for memory institutions which grant rights to them to transfer copy-protected media within the long-term-preservation strategies of cultural heritage institutions.

The documented user requirements within this document are listed and linked to their associated use cases scenarios, ordered by priorities (mandatory, desirable, optional,

enhancement) to help with the design decisions and development of a working implementation of the TTF.

List of Related Documents

Description of Work	*DoW*
Knowledge base about different media carriers and image file formats	*I1.1*
Evaluation of existing media transfer tools	*I1.2*
Document presenting the state of legality in the field of computer media copyright	*D1.1*

Abbreviations

Bibliothèque Nationale de France	*BnF*
Deutsche Nationalbibliothek	*DNB*
Emulation Access Platform	*EAPF*
Emulation Framework	*EF*
Graphical User Interface	*GUI*
Keep Virtual Machine	*KVM*
Koninklijke Bibliotheek	*KB*
Planets Core Registry	*PCR*
Transfer Tools Framework	*TTF*
Unified Digital Formats Registry	*UDFR*
Work Package	*WP*

Table of Contents

1 Introduction

A steadily increasing amount of various digital media carrier types, often found as, but not limited to, book and periodical supplements, make their way into collections of cultural heritage institutions. The variety and the amount of the digital media carrier types are driven by the evolution and increasing popularity of digital storage media within the rapid development in Information Technology since the invention of the personal computer.

This situation challenges the cultural heritage institutions to keep pace with rapidly superseded and changing media types. Many of them require specific knowledge about required hardware and software environments, how to operate them and, last but not least, how to access their original data to be able to preserve all the information carried on the various digital media carrier in their collections.

1.1 Motivation

Information stored on digital media carrier is exposed to several threats which can lead to either inaccessible data which is not accessible anymore due to media obsolescence, or corruption leading to partial or complete loss of data.

Even if data stored on digital media carrier can still be accessed and read out, the knowledge about how to interpret the data could be lost, making that data valueless even if it can be read.

Cultural heritage and memory institutions have large collections of various digital media carrier types at risk of:

1. Degradation - Chemical instability of the recording layer leading to:
    1. Deterioration of recording
    2. Mechanical failure
2. Obsolescence - The medium becomes inaccessible due to shifts in technology:
    1. Hardware obsolescence
    2. Software obsolescence (operating system, application, file format)

Cultural heritage institutions need help in handling the variety of digital media carriers in order to be able to transfer their stored original information into their long-term preservation archives before the media carrier becomes inaccessible due to degradation or obsolescence.

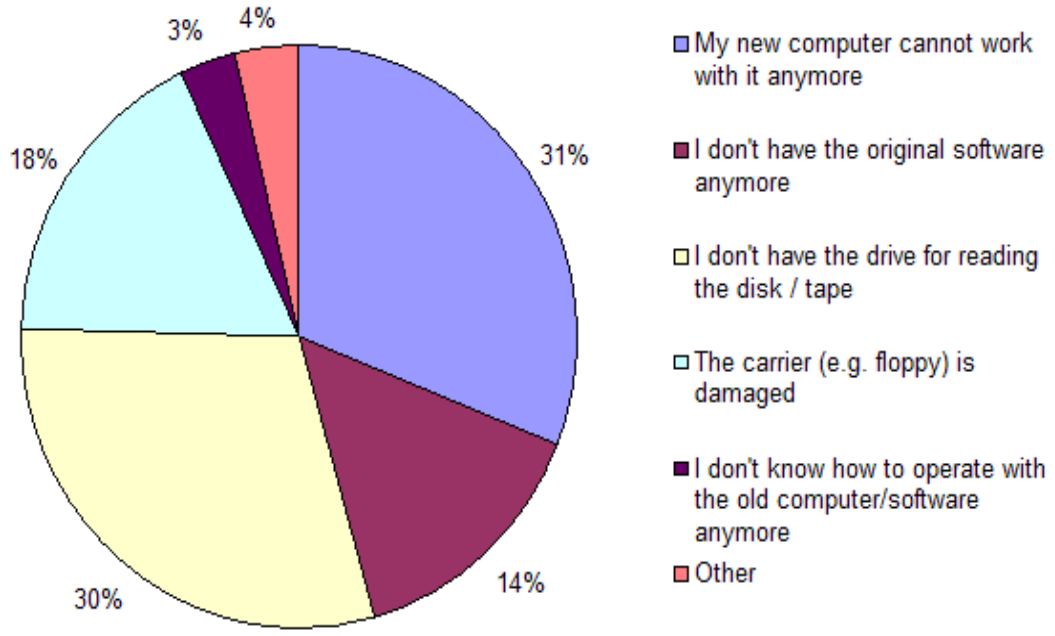

Figure 1: Result of the library user survey in the KEEP project asking: 'Do you know why you can't access these files or programs anymore?'

Figure 1 shows the result of the library user survey undertaken in the KEEP project including combined results from national libraries in the Netherlands, France and Germany for the question: 'Do you know why you can't access these files or programs anymore?'

The survey shows the problems typical users experience with their digital media carrier and required hard- and software environments. But before such problems can arise it is first of all of crucial importance to be able to access and transfer the original content of the physical media carrier into an appropriate image format onto a more reliable storage medium to enable long-term preservation planning and actions take place.

This will not only require a software solution for the transfer of the original bit-stream from the digital media carrier but also appropriate hardware reader(s).

1.2 Objectives and scope

This document specifies and describes the gathered user requirements including all functional features and non-functional requirements such as licensing and documentation of the TTF (Transfer Tools Framework). The TTF shall be able to incorporate existing and future transfer tools in such a way that it will allow the secure transfer of digital information contained in a variety of digital media carrier types into their appropriate image file formats.

This document only targets the TTF and does not address the other components of the EAP (Emulation Access Platform), namely the EF (Emulation Framework) and the KVM (KEEP Virtual Machine). Furthermore, the TTF assumes that specific hardware readers and drivers are available.

Legal, financial or organisational aspects are not covered within this document. However, the legal framework of France, the Netherlands and Germany regarding computer media copyright is being carried out and documented separately.

The conceptual architectural design based on the documented user requirements within this document is described in deliverable D1.2b which is the second half of the D1.2 deliverable bundle.

1.3 Outline

This document starts with an introduction; followed by the overview of how the TTF fits in the Emulation Access Platform (EAP).

It then presents a high-level view of the TTF and its dependencies with external systems and how the different actors can interact with the system is presented. Then, the identified Use-Case scenarios are described.

Finally, the gathered user requirements based on the identified Use-Case scenarios and the undertaken research within cultural heritage institutions in KEEP are listed, split into functional and non-functional requirements with given priorities.

This document closes with a glossary including definitions for terms used in this document.

2 System Context

The TTF is part of the EAP, which, together with the Emulation Framework (EF) and the Keep Virtual Machine (KVM), constitutes the long term preservation approach developed in the KEEP project. The Emulation Framework takes care of the rendering of digital objects addressing software obsolescence by providing a framework for execution of various emulation solutions; but to do so the digital information to be rendered has to be extracted successfully from a digital media carrier in such a way that the original bit-stream is preserved without being altered. However, before going into more details about the TTF itself, a general introduction into the context and its related issues is presented.

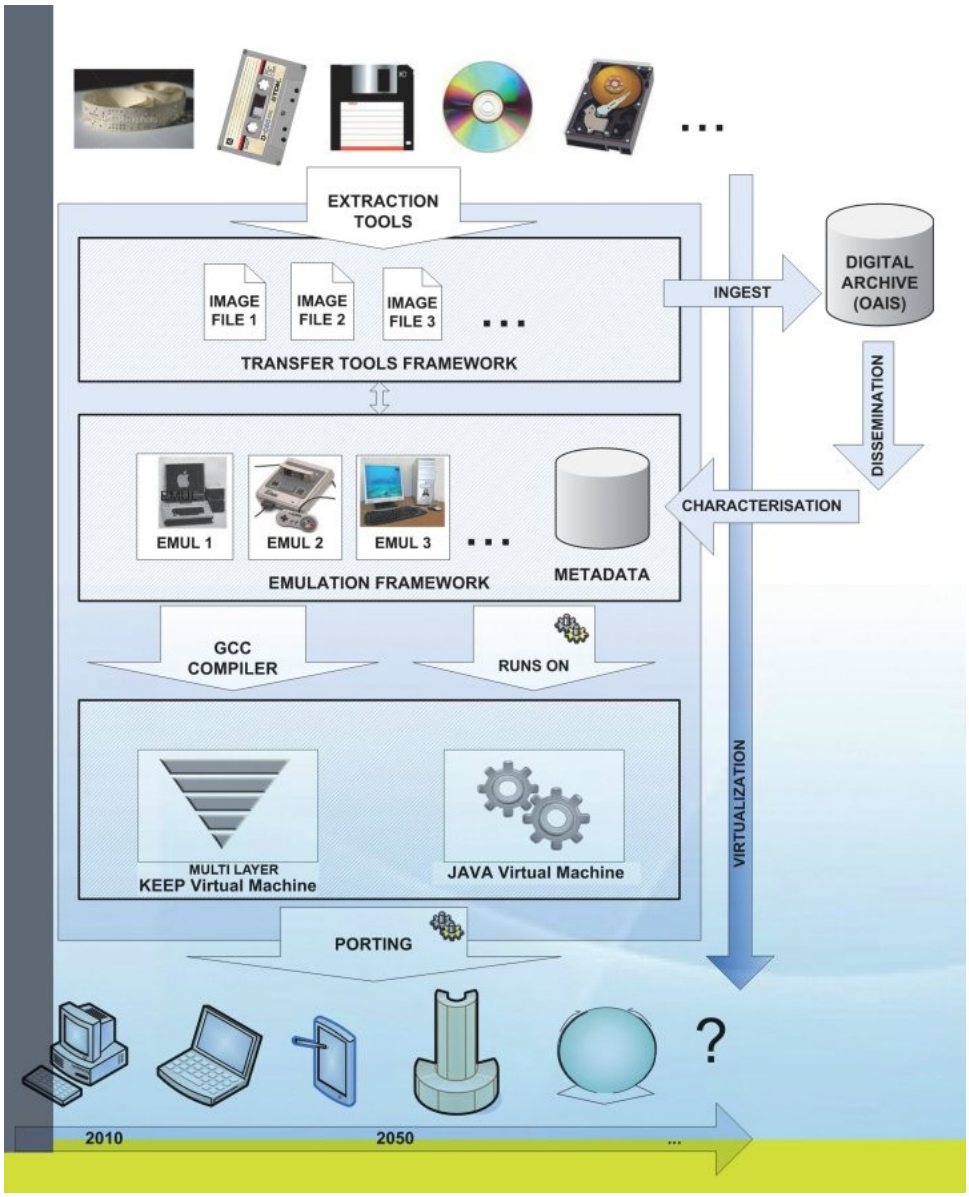

Figure 2: The Transfer Tool Framework within the Emulation Access Platform.

2.1 Media transfer

2.1.1 Media types

Figures provided by the national libraries of France, the Netherlands and Germany as well as from the Computerspiele-Museum have shown that CD-ROMs and 3 ½" floppy disks are the most popular and common supplementary materials in their collections.

Therefore, support for the transfer workflows of CD-ROMs and 3 ½" floppy disks have been chosen to be supported by the TTF for the proof of concept within the KEEP project.

However, support of further transfer workflows for digital media carrier by the TTF shall be able to be extended for future enhancement of the framework to add support for more digital media carrier types after the life-time of the KEEP project easily.

2.1.2 Image formats

There are many different common image formats having evolved in parallel for a variety of common hardware platforms. Most of the times, there are many different proprietary image formats for a single hardware platform as these formats were developed over many years by different communities for use with different software tools. Creating a new universal image format could be tempting in order to standardise them all - one format to rule them all. But this would only add complexity to an already complex topic. Moreover, the existing tools would have to implement the support for this new format and the transfer tool developers have to give reason as the existing ones work well enough.

Also, the support of image formats is the responsibility of transfer tools. The role of the TTF shall be to assist the user in choosing the appropriate image format once the related media carrier type has been identified. Therefore the TTF shall hold knowledge about digital media carrier and relations to the appropriate image formats in such a way that the TTF assists the user during the media transfer in choosing the best image format for its media carrier type.

2.1.3 Copy-protected media and the legal framework

Many of the recent commercial media contains some kind of copy-protection mechanism (i.e. ProtectCD, SafeDisc, etc.), which prevent the content of the media from being copied by unauthorized users. Copy-protected media is a huge problem which could be addressed with support of the huge community which is working on circumventing copy-protection mechanisms through methods developed by reverse-engineering or trial-and-error and present workarounds to the public via the internet. Therefore transfer tools for use within the TTF ideally shall be capable of dealing with common copy-protection mechanisms installed on the digital media carrier to circumvent these but this is prohibited by the legal framework.

However, a preservation strategy cannot not take advantage of these methods unless and until a specific exception to the current laws is put in place, for example for digital preservation at memory institutions that have as goal the retention of these items in their collections.

As of now, no such EU-wide law is in place. We must, therefore, create a system that is legal in all EU countries, by using only the legal tools available.

2.1.4 Media transfer tools

Transfer tools are the obvious mandatory tools that the TTF needs to be able to accommodate. Many tools are available, whereas some are better suited than others for certain cases. The most of these tools offer control via their own GUI and only some of the available transfer tools offer control via a command-line.

However, for the TTF only transfer tools which are able to be operated via a command-line are of interest to be integrated as it makes no sense to control a GUI via another GUI.

A comprehensive evaluation of some of the common existing transfer tools solution in hard and software has been carried out within KEEP.

For the available tools it is important to have 'smart' handling when it comes to read repetitions when transferring the contents of a media carrier which is already in bad condition. Read repetition support is a mandatory capability when it comes to transfer tools as often the information read from the digital media carrier may be either appear on first read to be unstable or actually is. One reason for such instability or read failure may be the deterioration of the media carrier but often such read failures are due to copy-protected mechanisms in place on the digital media carrier faking the disk condition to prevent copying.

Analysis tools are also a potential candidates of tools for the TTF since they could be used for checking, e.g. the integrity of the produced image file and/or metadata generation or for any other post-image generation process that would be beneficial to support the typical user during transfer workflows.

2.1.5 Technical registries

Finally, with the increasing popularity of web-services, technical registries such as global file format registries as PRONOM[1], PCR[2] and UDFR[3], as well as Mediapedia[4] which hold information for characterization could become useful sources for accessing certain information for transfer tools and image formats.

Therefore the TTF shall support integration with available web-services to assist the user with the interpretation of the transferred data.

[1] The technical registry PRONOM, http://www.nationalarchives.gov.uk/PRONOM/Default.aspx, Last retrieved: 22.07.2011

[2] Planets Core Registry, http://corereg.arts.gla.ac.uk/PlanetsCoreRegistry/welcome.html, Last retrieved: 22.07.2011

[3] Unified Digital Format Registry (UDFR), http://www.udfr.org, Last retrieved: 22.07.2011

[4] Mediapedia: Physical Format Carrier Resource, http://mediapedia.nla.gov.au, Last retrieved: 22.07.2011

2.2 Overview of the Transfer Tool Framework

Figure 3 shows an overview of the TTF clearly indicating its main goal: producing image files from a digital media carrier.

Grey elements represent components outside the scope of the TTF implementation and green and light green represents the main component of the TTF and its boundary, respectively.

Appropriate hardware readers as well as existing transfer tools, analysis tools and technical registries are greyed out as they do not need to be developed (pre-existence is assumed) although the latter are integral part of the resulting TTF as a whole; without the appropriate hardware reader and transfer tools, the TTF would not have the ability to carry out its aim.

Descriptive metadata must accompany a newly produced image file in order to characterise the media carrier and technical metadata which is optional within the scope of the TTF is required when the content of a previously generated media carrier image shall be further processed by other systems such as the EF (i.e. for rendering). The metadata model is developed within the work package WP3 of the KEEP project. [1]

User interaction is envisioned to be performed via a GUI for user-friendly interaction with the TTF while a machine interface can directly communicate with the core system of the TTF, allowing it to interact with other existing systems.

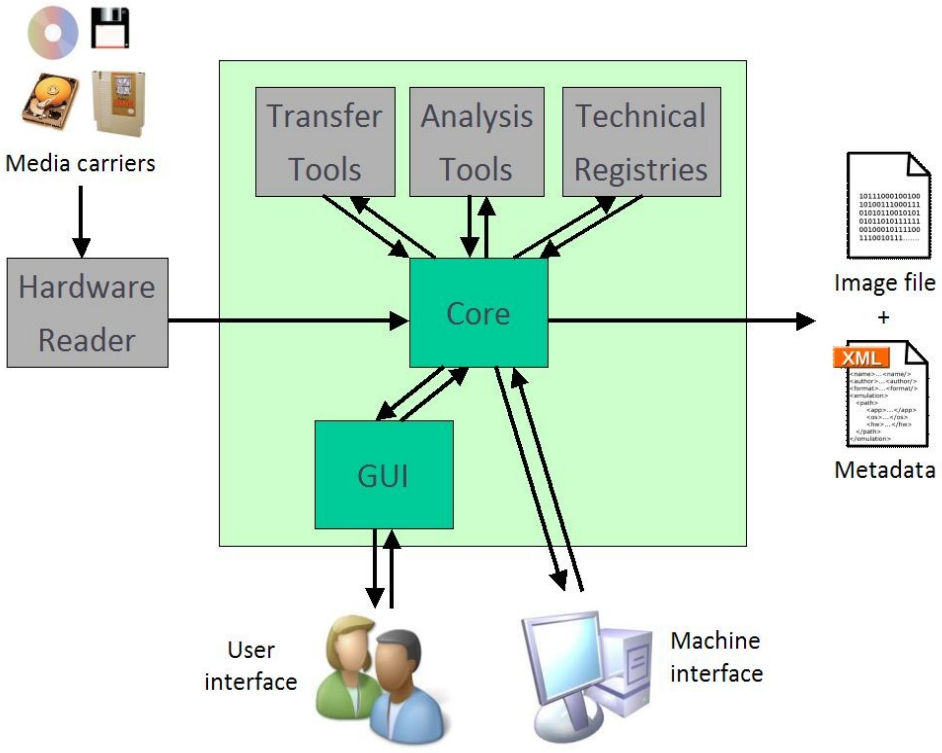

Figure 3: Overview of the Transfer Tool Framework system context and boundaries.

[1] "Annex I - "Description of Work" version 2011", page 35, Task 3.1 – 'Determination of standardized meta-data for digital objects', University of Portsmouth, Portsmouth, Hampshire, KEEP Description of Work, revision 1.0, 25th February 2011

2.3 User Role Definition

Potential users of the TTF mainly comprise:

- Librarians wanting to securely transfer the content of a media carrier into the archive of their institution for preservation purposes.
- Gamers, researchers/visitor of archive/museum/library, willing to use an emulator to access the content of an obsolete media carrier (e.g. a game cartridge or an old 5¼'' floppy disk) where an appropriate image has to first be created in order to be read by the chosen emulator.
- External (automated) systems. Although most of the user interaction will be done by humans via the GUI, it should be possible for other system to call and interact with the TTF in order to incorporate it into existing tool chains.

From this, different roles can de defined:

- A basic end-user who has limited knowledge on running transfer tools and therefore needs the TTF to reduce some of the complexity for them.
- A more advanced user ('administrator') who has more in-depth insight into the configuration and running of transfer tools and management of the resulting image files and metadata. This user will require access to more advanced parameterization and options.
- A machine 'user' which covers the machine-to-machine interaction i.e. any non-human interaction with other systems.

These 3 roles have to be taken into account when defining the requirements and use cases and they cover the 3 main ways of accessing and using the TTF.

3 Use-Case Scenarios

This section describes in detail the possible scenarios that the 3 different user roles as defined in 2.3 can face when using the TTF and forms the basis of the requirements listed in chapter 4.

Scenario 1: Transfer of a single media carrier

A librarian wants to store the supplementary material of a computer science book in a long-term preservation archive. This material is given in the form of a single 3-½" floppy disk that needs to be transferred into an appropriate image file. Metadata has to be generated to support and ensure the long-term accessibility of the transferred original bit-stream (e.g. via the Emulation Framework or another system).

To achieve this, the librarian starts the TTF and selects the drive where the floppy disk is inserted and selects the appropriate format of the media carrier from a presented list of media carrier types currently supported within the TTF (e.g. Amiga 3 ½" floppy disc). Based on this selection the TTF then characterizes the media carrier resulting in a first set of descriptive metadata which will be stored together with the media carrier image later and which might be expanded during the transfer workflow with technical metadata. What follows is the selection of an appropriate transfer workflow from those for the selected media type which are suggested to the user, ordered by feedback score (see scenario 3).

Once selected, the details of this workflow are displayed: the transfer tool and the sequence of optional post-transfer analysis/verification tools are shown. The user has the possibility to then click on each tool/step of the workflow to change the presented default parameters for each tool/step of the workflow, if required. The user is then asked for an output directory to specify the destination where the image file and metadata that will be generated by the TTF.

When "run" is clicked, the selected transfer workflow is launched and a progress bar is displayed informing the user of the progress of the job. Once finished, a link to the output directory is given so that the user can view the generated files.

When the image file and optional technical metadata creation is successfully completed, the system should perform an integrity check and verify that the content of the generated image file is identical to the original data captured from the physical media carrier.

In the case where an error occurred during the transfer, the user will be informed about the error. Then the user shall be able to decide if they want to retry, cancel or continue the execution of the initiated transfer. In any case, once the transfer workflow is completed, the system will present a summary report containing results of the performed media transfer.

Scenario 2: The administrator adds a new transfer/analysis tool or edits or removes an existing transfer/analysis tool

To add a new transfer or analysis tool to the framework, a package, containing the tool itself and metadata describing it (i.e. supported commands, media carrier, image file formats, its copy protected media capabilities, etc.) must first be created. Such a package is required by the TTF to be created for each transfer/analysis tool that is to be integrated into the framework. The administrator can add one or more packages to the TTFs internal database.

The metadata schema for such a package shall be developed during task 1.6 and shall comply with the metadata model as developed within the KEEP project. Tools that shall be integrated within the TTF must follow and implement this metadata schema in order to

provide all the required metadata necessary to be successfully integrated within an implemented transfer solution based on the TTF. [1]

Each package will be checked for integrity and verified by the system (i.e. presence of an executable, validation of accompanying metadata file against the developed metadata schema). If the package passes the validation, the package will be added to the internal database of the TTF and will be ready for use within. Each transfer/analysis tool can be removed or edited once added to the framework, to accommodate future required changes.

Scenario 3: The administrator sets up a new transfer workflow, edits or removes an existing transfer workflow

The TTF provides the function for the administrator to specify and add new transfer workflows to the framework. In order to specify a new transfer workflow the administrator has to choose a transfer tool and one or more analysis tools. The administrator then set up the sequence in which the tools shall be executed within this workflow and sets the default parameters/options for these tools (i.e. tool choices, configurations and running sequence). Once completed, the newly created transfer workflow is added to the TTFs internal database.

The administrator can select an already defined transfer workflow from a list of specified and stored workflows and edit it in an analogous way as previously described for new transfer workflows.

Scenario 4: The user provides feedback for a completed transfer workflow.

Once a transfer workflow has completed, a comprehensive report will be presented to the user informing them about failure or success of the operation. This involves checking if all tools which are set up in the chosen transfer workflow have run to completion without any errors and by checking that the expected image files and metadata file have been created successfully.

Subsequently, the user is encouraged to optionally leave their personal feedback for relevant characteristics as there might be speed, quality, ease of use, etc.. This feedback is then processed internally and will be translated into a score that will be accumulated over time every time this transfer workflow is used. Even if the user does not leave feedback, an automatic basic feedback is produced based on the success of each processed transfer workflow. The success of a transfer workflow will be internally determined as the accumulated results of each executed transfer workflow execution steps. Both feedback data sources- external: (provided by the user) and internal: (obtained from system routines) - will be used to produce an overall feedback score for the processed transfer workflow.

Scenario 5: Transfer of a set of media carriers

Typically, a user will have a collection of media carriers of the same type (i.e. a group of DOS formatted 3 ½" floppy disks) to transfer. A batch mode which supports the transfer of several media carrier in sequence, using the same transfer workflow is then enabled which requires that each carrier is of the same type and format. After being prompted for the quantity of carrier to be transferred, the first carrier is inserted and the appropriate transfer workflow is worked out via media carrier characterisation as previously described in scenario 1.

[1] "Annex I - "Description of Work" version 2011", page 32, Task 1.6 – 'Specification, implementation and testing of the transfer tool framework', University of Portsmouth, Portsmouth, Hampshire, KEEP Description of Work, revision 1.0, 25th February 2011

Once the first workflow has completed, the user is prompted for inserting the next carrier, and the transfer is repeated (using the same workflow) until all media carriers have been transferred. After successfully completing each single transfer, the system should perform an integrity check and verify that the content of the generated image file is identical to the original data captured from the physical media carrier.

In case of an error during the transfer in batch mode, the user will be informed about the error. Then the user shall be able to select whether they wish to retry, cancel or continue with the execution of the setup transfers in batch mode. In any case when completed the system will present an overview with results of all the performed media transfers.

Scenario 6: The user wants to view/edit the generated metadata

Once a workflow has completed successfully, the generated image file and its associated metadata will be stored in the specified output directory. At this stage, a manual inspection of the metadata generated by integrated tools and services might be required in order to correct any inadequate or erroneous metadata.

For this purpose, the user can launch a metadata viewer/editor that can be either external or internal to the TTF. The user is then able to verify the validity of the metadata, edit it manually, or even delete parts or the complete set of generated metadata if required.

Scenario 7: An external system calls the TTF to automate image generation and verification

In this scenario, an external program is directly calling routines of the TTFs core using functionalities offered by the TTF via the machine interface. Here the external program takes on the role of the user by using the TTF as a service provider. To this end, the machine interface, rendered public by the TTF and used by the external system, allows the TTF to be fully embedded in existing or future tool chains. It is foreseen that all functions, normally available to a human user, should also be available via the machine interface.

Scenario 8: Verification and test of generated image files

After a transfer workflow has completed, the content of the generated image file will be checked for integrity and it will be verified in order to ensure that the data within the generated image file matches the original data on the physical media carrier.

Scenario 9: Export of the current settings for transfer/analysis tools and services

This scenario addresses the use-case where the end-user wants to save a transfer workflow setting. The end-user selects a transfer workflow from a list of available workflows in the TTF. They can change the default settings and options of the individual tools and services used for that workflow. Once the transfer workflow settings are edited, the end-user can export these new settings into a specific location for later re-use.

Scenario 10: Import of saved settings for transfer/analysis tools and services

This scenario addresses the use-case when the end-user wants to re-use a transfer workflow setting that has been previously saved rather than the recommended settings. The end-user selects a transfer workflow from a list of available workflows in the TTF. They can change the default settings and options of the individual tools and services used for that workflow, if

required. Once the transfer workflow settings are edited, the end-user can export these new settings into a specific location for later reuse as previously described in scenario 9.

4 Requirements

The following two sections describe the identified requirements. The first section addresses all functional requirements; the second section non-functional requirements. All requirements have been categorised and marked prioritised:

M - Mandatory requirement: *this feature must be built into the final product.*

D - Desirable requirement: *this feature should be built into the final product unless its cost is too high.*

O - Optional requirement: *this feature can be built into the final product at the Project Manager's discretion.*

E - Possible future enhancement: *this feature is recorded here so that the idea is not lost, but it will not be included in this product.*

4.1 Functional requirements

The list of requirements in the tables below distinguish between requirements for the transfer tool framework itself and an transfer tool platform application which is planned to be developed in task 1.6[1] during the KEEP project to show the proof of concept.

The requirements for the transfer tool framework are listed in table 4.1.1 Core Framework and the ones for the transfer tool platform application are listed in table 4.1.2 Application front-end (GUI).

4.1.1 Core Framework

Requirement	Scenario	Description	Priority
TTF-CF-R1	S1, S5, S7	The system shall be able to generate a carrier image from a digital media carrier.	M
TTF-CF-R2	S1, S2, S5, S7	The system shall be able to support the transfer workflow to generate a media carrier image of at least 2 carriers.	M
TTF-CF-R3	S1, S5, S8	The system shall be able to check that the contents of the generated carrier image are the same as the original data on the physical media carrier.	M

[1] "Annex I - "Description of Work" version 2011", page 32, Task 1.6 – 'Specification, implementation and testing of the transfer tool framework', University of Portsmouth, Portsmouth, Hampshire, KEEP Description of Work, revision 1.0, 25th February 2011

TTF-CF-R4	S1, S5, S7	The system shall be able to generate descriptive metadata describing the original physical media carrier from which the original content shall be transferred into a media carrier image.	M
TTF-CF-R5	S1, S5, S7	The system shall be able to generate technical metadata describing the contents of the generated media carrier image using integrated analysis tools (i.e. JHOVE, DROID, FITS, etc.).	D
TTF-CF-R6	S1, S5, S7	The system shall be able to retrieve additional technical metadata from external sources via technical registries (i.e. PRONOM or similar web services) to allow the generated technical metadata to be updated and expanded.	O
TTF-CF-R7	S1, S5, S6	The system shall allow the generated metadata to be updated/edited	D
TTF-CF-R8	S1, S5, S6, S7	The system shall be able to output the generated carrier image including metadata into a specific location.	M
TTF-CF-R9	S1, S5, S7	The system shall be able to be completely controlled via a machine interface.	M
TTF-CF-R10	S1, S5, S7	The system shall be able to be completely controlled via a GUI.	M
TTF-CF-R11	S2	The system shall be able to register, deregister and configure transfer tools and additionally required services.	M
TTF-CF-R12	S3	The system shall allow the management (i.e. create new, configure, remove, etc.) of transfer workflows, based on the available transfer/analysis tools and services.	M
TTF-CF-R13	S3	The system shall allow the sequence of tools/services within a workflow to be edited.	M
TTF-CF-R14	S1, S5	The system shall be able to suggest appropriate transfer workflows for a given media carrier type based on their rating/score stored within the TTF.	M
TTF-CF-R15	S1, S3	The system shall allow the user to edit the default parameterization of transfer tools and additional services within a transfer workflow.	M

TTF-CF-R16	S9, S10	The system shall be able to export the current transfer workflow configuration (i.e. settings for each tools used in that workflow) to a local disk for later user	D
TTF-CF-R17	S10	The system shall be able to load previously exported settings for a particular transfer workflow	D
TTF-CF-R18	S1	The system shall be able to run transfer workflows automatically.	M
TTF-CF-R19	S5	The system shall be able to run several transfer workflows in a batch mode, unattended.	M
TTF-CF-R20	S1	The system shall be able to provide information about the progress of running transfer workflows.	D
TTF-CF-R21	S1	The system shall be able to operate in such a way that it recognises automatically the format of a given media carrier.	D
TTF-CF-R22	S4	The system shall be able to store ratings for each transfer workflow.	M
TTF-CF-R23	S4	The system shall be able to update the rating for each transfer workflow automatically based on the results of executed transfers.	M
TTF-CF-R24	S4	The system shall allow the user to provide feedback about executed media transfers (to update the internal rating of transfer workflows).	D

4.1.2 Application front-end (GUI)

Requirement	Scenario	Description	Priority
TTF-GUI-R1	S6	The GUI shall support update of the metadata accompanying a generated media carrier image via an integrated editor.	D
TTF-GUI-R2	S1, S8	The GUI shall present the results of the validity check of the generated carrier image to the user.	M
TTF-GUI-R3	S2	The GUI shall be able to decide if the presented information shall be editable by the end-user or not based on the different granted rights of the user roles admin / basic user.	M

TTF-GUI-R4	S2	The GUI shall allow only the admin to be able to manage (add / remove) transfer/analysis tools and services.	M
TTF-GUI-R5	S1, S3	The GUI shall allow only the admin to be able to manage (add / remove) transfer workflows.	M
TTF-GUI-R6	S1, S3	The GUI shall allow all end-users to update the default options of tools and services used within a transfer workflow.	M
TTF-GUI-R7	S1	The GUI must display a list of the available physical hardware drives installed on the host machine, each duplicated for any known and supported media carrier type that could be possibly inserted within to enable the user to select the proper hardware drive together with the appropriate media carrier type and format at once.	M
TTF-GUI-R8	S1	The GUI must be able to display (with capabilities for sorting the list by name, version, rating, etc.) a list of available and suggested transfer workflows.	M
TTF-GUI-R9	S1, S3	The GUI must be able to display (with capabilities for sorting the list by name, version, rating, etc.) a list of available and suggested transfer/analysis tools and services.	M
TTF-GUI-R10	S1	The GUI must be able to display metadata information about a transfer/analysis tool or service.	M
TTF-GUI-R11	S1	The GUI should be able to display a list of all options for a given tool/service for them that are also configurable by the user.	D
TTF-GUI-R12	S4	The GUI shall provide a way for enable the user to provide feedback to the system about a completed transfer workflow.	D
TTF-GUI-R13	S1, S5	The GUI shall enable end-users to choose between a managed (assisted / guided) and manual operation mode.	M

4.2 Non-functional requirements

4.2.1 Documentation

Requirement	Description	Priority
TTF-NF-R1	Provide an instruction manual for the user.	M
TTF-NF-R2	Provide an instruction manual for the admin.	M

4.2.2 Licensing

Requirement	Description	Priority
TTF-NF-R3	The system shall be subject to a software licence enabling free distribution.	M

5 Glossary

Analysis Tools – Software to analyse media carrier images.

Descriptive Metadata – Metadata describing the content of the digital media carrier image.

Media carrier image, carrier image, image – An exact replica of the contents of a digital media carrier (i.e. floppy-disk drive, CD-ROM, etc.) stored on a second storage device (i.e. hard-disk).

Metadata – Data about data.

Technical Metadata – Metadata about the digital media carrier itself.

Technical registry, external technical registry – External registry providing technical data about (i.e. file formats, etc.) which can be retrieved via web services.

Transfer, Imaging – The process of media carrier image generation.

Transfer Tools – Software to generate media carrier images.

Transfer Workflow – Customizable workflow which covers required steps to transfer and verify digital media carrier images using integrated transfer / analysis tools and maintained knowledge about media transfer from an internal database.

Web services – Web services describes a standardized way of integrating Web-based applications using the XML, SOAP, WSDL and UDDI open standards over an Internet protocol backbone. XML is used to tag the data, SOAP is used to transfer the data, WSDL is used for describing the services available and UDDI is used for listing what services are available.[1]

[1] Webopedia, http://www.webopedia.com/TERM/W/Web_Services.html, Last retrieved: 15.10.2010