



Project Number	IST-2006-033789
Project Title	Planets
Title of Deliverable	Core Registry V3: Software Requirements Document
Deliverable Number	PC3-D20
Contributing Sub-project and Work-package	PC/3, PA/3
Deliverable Dissemination Level	External
Deliverable Nature	Software Requirements Document
Contractual Delivery Date	30 th September 2009
Actual Delivery Date	21 st December 2009
Author(s)	Tessella, TNA, KB

Contributors

Person	Role	Partner
Pauline Sinclair	Author	Tessella
Rob Sharpe	Reviewer	Tessella
Tim Gollins	Reviewer	TNA
Lynne Montague	Reviewer	TNA
Sara van Bussel	Reviewer	KB-NL

Document Approval

Person	Role	Partner
Awaiting Approval by partners		

Distribution

Person	Role	Partner

Revision History

Issue	Author	Date	Description
V1.R1.M0	Pauline Sinclair	10-Jul-2009	Re-written from the PRONOM 7 SRD to correct mistakes and clarify the requirements.
V1.R1.M1	Rob Sharpe	05-Aug-2009	Reviewed.
V1.R1.M2	Pauline Sinclair	20-Sep-2009	Added in requirements from Testbed and Plato and clarifications from TNA and KB-NL. Added in a high-level business view of PCR and an intermediate-level data model.
V1.R1.M3	Pauline Sinclair	09-Nov-2009	Updated after review by Lynne Montague (TNA) and Sara van Bussel (KB-NL).
V1.R2.M0	Pauline Sinclair	21-Dec-2009	Updated after review by Lynne Montague (TNA) and Sara van Bussel (KB-NL) to contain just the requirements for PCR 2.1.

Requirement Sources

Ref.	Document	Date	Version
[P3SD]	PRONOM 3 System Description	22-Aug-2003	03
[P4URD]	PRONOM 4 User Requirements	13-July-2004	0f
[P4IM]	PRONOM 4 Information Model	29-June-2004	0d
[Current system]	Undocumented requirements but implicit from the current PRONOM3 implementation.	-	-
[Meeting 1]	Meeting at Kew		
[Clarification 1]	Response to e-mailed questions	26-Nov-2004	
[Byte]	PRONOM 4 Byte Sequence Modelling	21-Dec-2004	
[Clarification 2]	Response to e-mailed questions	08-Dec-2004	
[PD]	Tessella PRONOM 4a Project Description	27-Sept-2004	V1.R2.M0
[Review]	Review meeting	16-Dec-2004	
[Feedback1]	E-mailed feedback from Adrian Brown to V1.R2.M1 of this document	06-Jan-2005	
[13-Jan-2005]	Teleconferences on 11 th and 13 th January to discuss file format identifications	13-Jan-2005	

Ref.	Document	Date	Version
[27-Jan-2005]	E-mails and telephone calls 24-27 January to iron out details of some requirements	27-Jan-2005	
[10-Feb-2005]	E-mail to clarify search & reporting requirements for software process types.	10-Feb-2005	
[11-Feb-2005]	E-mail to clarify search requirements for vendors.	11-Feb-2005	
[18-Feb-2005]	E-mail to clarify displaying provenance information in detailed reports.	18-Feb-2005	
[23-Feb-2005]	E-mail to clarify the behaviour of the lifecycles searches with respect to null release and withdrawn dates.	23-Feb-2005	
[DemoProgress]	Minutes of PRONOM 4 progress review Meeting on 10 March 2005.	10-Mar-2005	
[14-Apr-2005]	Phone call to agree requirements for support period searches, with respect to null dates.	14-April-2005	
[8-Apr-2005]	Conversation with Adrian Brown on 8 April 2005.	8-April-2005	
[PRONOM5a]	PRONOM 5a requirements Tessella Ref: NPD/4826/CL/CSC/2006FEB06/17:46:49	06-Feb-2006	
[5aSRD]	PRONOM 5a Software Requirements Document	13-Feb-2006	V1.R4.M0
[URD]	Preservation & Maintenance Project ITT Annex B Ref: SF/PM/07/DOS	13-Apr-2006	Version 1.0
[MEET 4]	PRONOM Meeting Tessella Ref: NPD/4950/CL/MIN/2006AUG16/11:36:51	15-Aug-2006	
[TWSRD]	Technology Watch Software Requirements Document	21-Sep-2006	V1.R6.M0
[PRO6-SRD]	PRONOM 6 SRD V1R6M0.doc PRONOM 6 Software Requirements Document Tessella Ref: NPD/4950/SR/SRD/PRONOM 6 SRD V1R6M0.doc	21-Feb-2007	V1.R6.M0
[DROID3]	DROID SRD V3.R0.M0.doc Requirements for the DROID file identification tool.	05-Feb-2008	V3.R0.M0
[PC3/D4]	PC3_D4.doc Requirements for the Characterisation Registry for Planets.	25-Jan-2008	V0.3
[PA-tool-registry]	Requirements_PA_tool_registry-v11.doc Requirements for the Preservation Action Tool Registry for Planets.	04-Nov-2008	V1.5
[PRONOM7]	PRONOM 7 SRD DRAFT_FOR_REVIEW.doc PRONOM 7 Software Requirements Document	16-Jun-2008	V1.R1.M0
[MEETING-WIN/08/08/05]	Meeting between KB, TNA and Tessella. Present: Andrew King (TSS), Lewis Jardine (TSS), Rob Sharpe (TSS), Caroline van Wijk (KB), Sara van Bussel (KB), Adrian Brown (TNA), Matt Palmer(TNA)	5-Aug-2008	
[EMAIL-08-07-17]	Tessella Ref: NPD/4678/CL/CSC/2008AUG06/18:11:24 Email correspondence between Andrew King Caroline.vanWijk@KB.nl Frank Houtman" Frank.Houtman@KB.nl Re: XML export	17-July-2008	
[EMAIL-080724]	Tessella Ref: NPD/4678/CL/CSC/2008AUG06/20:22:50 Email correspondence between Lynne.Montague@nationalarchives.gov.uk "Brown, Adrian" Adrian.Brown@nationalarchives.gov.uk Lewis.Jardine@tessella.com Andrew King Re: XCDL schema files and XCEL files	24-July-2008	
[EMAIL-080728]	Tessella Ref: NPD/4678/CL/CSC/2008JUL28/09:15:55 Email correspondence between Andrew King Lynne.Montague@nationalarchives.gov.uk "Sara van Bussel" Sara.vanBussel@KB.nl "Brown, Adrian" Adrian.Brown@nationalarchives.gov.uk Caroline.vanWijk@KB.nl Frank Houtman" Frank.Houtman@KB.nl Julie.Stevenson@tessella.com Lewis.Jardine@tessella.com Re: licences	28-July-2008	
[PC3/D6]	PC/3 D6v2 Registry Iteration 3 Design - User Requirements Report	29-Aug-2008	V2
[Planets Review]	Planets review meeting discussions, 23-24/7/2009 in Vienna. Requirement suggestion came from one of the reviewers, Mark Fresco.	23/24-July-2009	
[MEETING-10-08-2009]	Tessella Ref: NPD/4678/CL/MIN/2009AUG18/09:38:59 Meeting at TNA to discuss the requirements for PCR version 3.	10-Aug-2009	

Ref.	Document	Date	Version
[MEETING-19-08-2009]	Tessella Ref: NPD/4678/CL/MIN/2009SEP03/15:34:12 Meeting at TNA to look at the usability of PCR version 2 (aka Pronom 7).	19-Aug-2009	
[MEETING-26-08-2009]	Tessella Ref: NPD/4678/CL/MIN/2009SEP01/11:46:05 Meeting held at TNA with Sara van Bussel and Lynne Montague to discuss auditing and other requirements.	26-Aug-2009	
[TESTBED-AUG-2009]	Tessella Ref: NPD/4678/CL/MIN/2009SEP02/18:25:28 Meeting held at TNA to discuss the Testbed requirements for PCR 3.	26 & 27 Aug 2009	
[EMAIL-04-09-2009]	Tessella Ref: NPD/4678/CL/CSC/2009SEP04/17:24:31 Email correspondence between Pauline Sinclair, Christoph Becker, Rob Trickey, Hannes Kulovits, Andreas Rauber and Tim Gollins	04-Sep-2009	
[WORKSHOP]	PCR Programmers' Workshop held at the British Library, St. Pancras	05-Feb-2009	
[TNA-Review]	Review of V1R1M2 of the SRD for the Core Registry V3 Tessella Ref: NPD/4678/CL/CSC/2009OCT26/08:33:05	23-Oct-2009	

Other References

Ref.	Document	Date	Version
[Planets Conceptual Data Model]	Planets Conceptual Data Model	29-July-2008	V1.R5.M4
[PRONOM_7_Data_Model]	Pronom Data Model.doc PRONOM 7 Data Model document Tessella Ref: NPD/4678/D/SMG/2009MAY18/17:34:26	6-May-2009	V1.R1.M0
[PRONOM_7_ADD]	PRONOM_7_ADD_v7r1m2.doc PRONOM 7 Architectural Design Document Tessella Ref: NPD/4678/AD/ADD/2008OCT10	10-Oct-2008	V1.R1.M2
[PLANETS_GLOSSARY]	http://www.planets-project.eu/private/pages/wiki/index.php/Glossary	15-Apr-2008	
[GDFR_FACETS]	Global Digital Format Registry (GDFR) Classification http://www.gdfr.info/docs/GDFR-Classification-1_0_5.pdf	09-Nov-2007	1.0.5
[INSPECT_DICTIONARY]	InSPECT Significant Properties Data Dictionary	05-Aug-2008	0.5

EXECUTIVE SUMMARY

This document details the software requirements for the Planets Core Registry versions 2.1. The core registry incorporates the functionality of both the Preservation Action Tool Registry and the Preservation Characterisation Registry. Note that the previous release of the Core Registry (V2) was also known as Planets Core Registry V2, the Planets Technical Registry and PRONOM 7.

TABLE OF CONTENTS

1.	Introduction	9
1.1	Purpose of this Document	9
1.2	Scope of this Document	9
1.3	Context of this Issue	9
1.4	Overview	9
1.5	Definition of Terms	9
2.	General Description	11
2.1	Business Overview	11
2.1.1	Version 3	11
2.2	Information View	12
2.3	Functional View	12
2.4	Environmental Considerations	13
2.4.1	Planets instance	13
2.4.2	TNA's system	13
2.4.3	Web browsers	13
3.	Logical Model	14
3.1	Overview	14
3.2	The Planets Core Registry	15
3.3	Definition of information held	16
3.3.1	File Format	16
3.3.2	Software	18
3.3.3	Computer Hardware	19
3.3.4	Storage Medium	21
3.3.5	Technical Environment	22
3.3.6	Character Encoding	22
3.3.7	Compression Technique	23
3.3.8	Property	23
3.3.9	Transformation Pathway	23
3.4	Note on Component Manifestations	23
3.5	A More Detailed View of the Data Model	26
4.	Issues and Discussion	27
4.1	Clarifications of Specific Issues	27
4.1.1	Software Packages	27
4.1.2	Classification	28
4.1.3	XCDL/XDEL files	28
4.1.4	PUIDS	29
4.1.5	Testbed Results Submission	30
4.1.6	InSPECT Significant Properties Data Model	30
4.2	Outstanding Issues and Enhancements	32
4.2.1	Invocation Details	32
4.2.2	Advanced Software Search	32
4.2.3	Version Control, Rollback and Synchronisation	33
5.	Data Requirements	34
5.1	General Requirements	34
5.1.1	Main information	34
5.1.2	Families	35
5.1.3	Types	35
5.1.4	Other information	36
5.1.5	PUID Assignment	38
5.2	Core Entities	39
5.2.1	File Format Requirements	40
5.2.1.1	File format information	40
5.2.1.2	External signatures	42
5.2.1.3	Internal signatures	42
5.2.1.4	Reference Files	43
5.2.2	Software Package Requirements	43
5.2.2.1	Software Components	45

5.2.2.2	Software Package Tools	45
5.2.2.3	Software Package Licences.....	46
5.2.3	Hardware Requirements	46
5.2.4	Character Encoding Requirements.....	48
5.2.5	Compression Technique Requirements.....	49
5.2.6	Storage Medium Requirements	50
5.3	Subsidiary Entity Requirements	51
5.3.1	People and Organisations (Agents)	51
5.3.2	Documentation	53
5.3.3	Intellectual Property Rights	53
5.3.4	Technical Environments.....	54
5.3.5	Processes.....	54
5.3.6	Pathways.....	56
5.3.7	Properties for File Formats and Components	58
5.3.8	Provenance information	59
5.3.9	External identifiers.....	60
5.3.10	Classification Facet Values	61
5.4	System Information.....	62
5.4.1	Audit Entries	62
5.4.2	Version Control, Rollback and Synchronisation.....	63
5.4.3	Technology Watch Alerts	63
5.4.4	Quarantine Area for Testbed Results Storage.....	64
Legacy search interface Requirements		65
6.1	General Requirements	65
6.2	Search and Reporting Requirements	66
6.2.1	Search layout	66
6.2.2	Searches for formats by extension, name or risk (format tab).....	66
6.2.3	Searches for software packages capable of processing formats (format tab).....	67
6.2.4	Search for format by PUID (PUID tab)	69
6.2.5	Search for software (software tab).....	69
6.2.6	Search by vendor name (vendor tab).....	69
6.2.7	Search by support period (lifecycles tab)	70
6.2.8	Search by release dates (lifecycles tab)	71
6.2.9	Storage media search (storage media tab).....	72
6.2.10	Simple search (simple search tab)	72
6.2.11	Detailed format report.....	73
6.2.12	Common requirements for detailed core entity reports	75
6.2.13	Detailed software report	75
6.2.14	Detailed agent report	76
6.2.15	General search results	76
6.2.16	General detailed reports.....	77
6.2.17	Migration Pathways Search.....	78
6.3	File Format Comparison.....	79
6.4	Help system Requirements	79
6.5	Non-functional Requirements.....	79
7.	Administration Requirements	80
7.1	General Requirements	80
7.2	Security.....	80
7.3	Repository Reader Functionality	81
7.3.1	Core Entities	81
7.3.1.1	Common.....	81
7.3.1.2	File Formats	82
7.3.1.3	Software Packages	83
7.3.1.4	Hardware.....	84
7.3.1.5	Character Encoding	85
7.3.1.6	Compression Technique	85
7.3.1.7	Storage Media.....	85
7.3.2	Subsidiary Entities.....	86
7.3.2.1	Common.....	86
7.3.2.2	Agents	86
7.3.2.3	Intellectual Property Rights	87

7.3.2.4	Documents	87
7.3.2.5	Technical Environments.....	88
7.3.2.6	Processes	88
7.3.2.7	Pathways.....	89
7.3.2.8	Properties.....	90
7.3.2.9	Component Manifestation Types	90
7.3.2.10	External Identifiers	90
7.3.2.11	Images	91
7.3.2.12	External Signatures.....	91
7.3.2.13	Internal Signatures.....	91
7.3.2.14	Reference Files.....	92
7.3.3	Policy Information.....	92
7.3.3.1	Pathway Policy Information.....	92
7.3.3.2	Process Policy Information	92
7.3.3.3	File Format Risk Policy Information	93
7.3.4	Search Functionality.....	93
7.4	Repository Administrator Functionality.....	94
7.4.1	Core Entities	94
7.4.2	Subsidiary Entities.....	94
7.4.3	Audit Trail	95
7.4.4	Lists of Values	96
7.4.5	DROID Signature File Generation.....	96
7.4.6	Technology Watch Alerts	97
7.4.7	Review Testbed Results	97
7.4.8	PUID Range Assignment	97
8.	Interfaces for External Systems	98
8.1	REST Interface	98
8.2	SOAP Web Services	98
8.2.1	Characterisation	98
8.2.2	Preservation Planning	99
8.2.2.1	Preservation Planning – File Formats.....	101
8.2.2.2	Preservation Planning – Software	101
8.2.2.3	Preservation Planning – Migration Pathways	104
8.2.3	Redaction	106
	Digital Repository System Access Control	106
8.2.5	DROID.....	106
9.	System Wide Requirements	108
9.1	Platform Requirements.....	108
9.2	Documentation Requirements.....	108
9.3	Installation Requirements.....	109
9.4	Security Requirements	109
9.5	Quality, Reliability and Maintainability Requirements	109

1. Introduction

1.1 Purpose of this Document

This document describes the software requirements for the Core Registry V2.1 system.

It is based on the requirements described in the existing Core Registry V2 (aka PRONOM 7) Software Requirements document [PRONOM7], supplemented by the requirements which were uncovered during the review of version 2.1 of the Registry.

It formed the context for the design phase and was used to derive the system test script. It was also used as the contractual description of the delivered system throughout the project.

1.2 Scope of this Document

This document updates and supersedes all previous Core Registry (or PRONOM) requirements documents. It includes the requirements for all parts of the Core Registry and covers those requirements for the Planets preservation action registry and preservation characterisation registry, which are both incorporated into the Core Registry.

1.3 Context of this Issue

This release has been made at the end of the development of version 2.1 of the Core Registry.

1.4 Overview

This document consists of a number of sections.

The first three sections (after this introduction) are discursive to enable readers to get an overview of the Registry:

- Section 2 gives a general description of the informational and functional view of the system together with an overview of environmental constraints on its installation and use.
- Section 3 describes the logical model of the information that the Registry has to hold.
- Section 4 describes some of the discussions that have been made during the requirements gathering process for versions 2, 2.1 and 3 of the Core Registry.

The remaining sections are a formal statement of the requirements in a form that enables them to be tested at the end of the development of the Registry:

- Section 5 covers the data requirements
- Section 6 covers the requirements that are specific to the legacy search system (which is an extension of TNA's current PRONOM web interface).
- Section 7 covers the administrative requirements
- Section 8 covers the web service interfaces
- Section 9 covers the environment requirements including those related to installation and documentation.

1.5 Definition of Terms

SRD The Software Requirements Document specifies the behaviour of the software system.

URD The User Requirements Document catalogues the users' requirements for the system.

ADD The Architectural Design Document describes the high-level design document for the entire system.

SDB Tessella's Safety Deposit Box digital preservation software, as used by TNA.

TNA The National Archives of England, Wales and the United Kingdom

PCR Planets Core Registry

Specific requirements status flags:

- 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.
- X Deleted requirement. Gravestone retained for reference.

2. General Description

2.1 Business Overview

The Planets Core Registry (PCR) comprises the Planets characterisation registry and the Planets preservation action tool registry and forms part of Planets' network of digital preservation services, where its role is to support the digital preservation activities of the other applications in the network. As such it provides persistent, unambiguous technical information about the representation and significant properties of generic object types, including file formats, compression algorithms, character encoding schemes, operating systems, application software, and hardware environments. It also implements the first ever scheme of persistent unique identifiers for object formats.

The registry also contains descriptions of characterisation tools which support the validation of preservation actions by measuring the significant properties of preserved objects. This enables the registry to support the automated deployment of characterisation tools, which are available as services within the Planets interoperability framework. The core registry is supplemented by the Testbed services registry, which contains further information about Planets services, including how to invoke them.

To aid the characterisation of digital objects, the Planets characterisation sub-project has developed two XML-based languages for describing digital objects and extracting information from them: the Extensible Characterisation Definition Language (XCDL), and the Extensible Characterisation Extraction Language (XCEL). Therefore, as part of its characterisation support role, the registry will store and expose XCDL and XCEL descriptions of formats.

In the context of the registry, a preservation action tool is a software program that performs a specific action on a digital object to ensure the continued accessibility of this digital object. This action could result in a transformation of the object (format migration) or a (re)creation of the technical environment required for rendering the object (emulation), or result in a combination of these two. The way that tools and services can be applied is described in a *pathway*. A *pathway* is a predefined set of one or more preservation actions operating on a specific input file format and resulting in a specified output format or target environment. Therefore, to support preservation action, the registry stores descriptions of preservation action tools, pathways and technical environments, which consist of combinations of operating systems and hardware.

The Planets preservation planning tool (PLATO) will make use of the registry for the discovery of preservation actions applicable to a set of objects and the evaluation of these actions to create a preservation plan. In particular, in the future it will use the aggregated results of experiments and evaluations run within the Planets Testbed application and stored in the registry and the outputs of the risk assessment service incorporated into the registry.

The registry provides an administration interface for maintenance of its data, and an application programming interface (API) to allow external systems, and the other Planets applications in particular, to access data stored in the registry. In addition to its support role within the Planets services network, the core registry will also serve as a source of information for preservation action, characterisation and planning for general users, such as employees from organisations that are concerned with digital preservation.

2.1.1 Version 3

The core registry is being developed in three iterations. The first two iterations have introduced core support for Planets characterisation and preservation action processes, including the description of characterisation tools, persistent unique identifiers, the capability to describe the inherent and instance properties of formats, the ability to document preservation action pathways and support for preservation planning.

The scope of the third iteration has been reduced to fit the available time and budget. The third iteration of the core registry will be version 2.1, which consolidates the functionality included in the first two iterations and fixes a number of bugs. It is intended that version 3 will be developed at some point in the future.

It is expected that version 3 of PCR will extend the range of preservation planning information made available to Plato. It will enable the storage of aggregated results from Testbed experiments and their retrieval by Plato and it will also support the technology watch service which will monitor

and identify changes in the core registry that potentially have an impact on preservation planning. Finally it will provide evolutionary enhancements to the registry and its interfaces.

It is thought that in version 3 the registry will look to a potential future as one of the nodes in a global Unified Digital Formats Registry (UDFR). As such it will start to move away from a view of generic objects that is based on families, which it inherited from PRONOM, to a view that is based on Global Digital Format Registry (GDFR) facets, which is more extensible.

2.2 Information View

The Core Registry is a repository of technical data that exists to support the long-term preservation of electronic records. As such, it stores information on file formats and their support by applications. It can also report on the ability of software to read and write given file formats, and the degree to which content is altered by software applications other than those originally used to create the electronic records.

In addition to information on file formats, the Core Registry stores technical information on other key entities from the digital preservation domain. These include hardware, software packages, technical environments, compression techniques, character encodings, storage media, pathways, and properties.

The Core Registry is intended to fulfil the dual roles of being the Preservation Characterisation Registry and the Preservation Action Registry within Planets. It is built upon the PRONOM system developed by The National Archives (TNA) in the UK.

2.3 Functional View

There are three main interfaces onto the Registry:

- A Java, web-based interface which acts as both the administration interface, where it is used for the maintenance of data, both factual and policy, and a public interface which allows users to browse and search the data. To access the administration functionality users have to log in; the browse and search functionality is accessible to all without logging in.
- The Legacy search interface. This is a .NET, public, web-based, user interface to search the data held in the repository. It is an extension of the PRONOM web interface currently deployed atTNA.
- A number of web services through which it communicates with other systems such as DROID, and Planets services and tools like Plato and Testbed.

These coincide with the three main classes of user of the system:

- Standard Users interact with the Core Registry via a web browser, to request and view reports.
- Administrators also interact via a web browser, for the same access as Standard Users plus the ability to add, edit or delete data from the system.
- External systems, such as DROID, interact with the Core Registry via web service calls or a REST interface. They can use this interface to access data held by the Registry; this is a purely data-level interface with no presentation of the data.

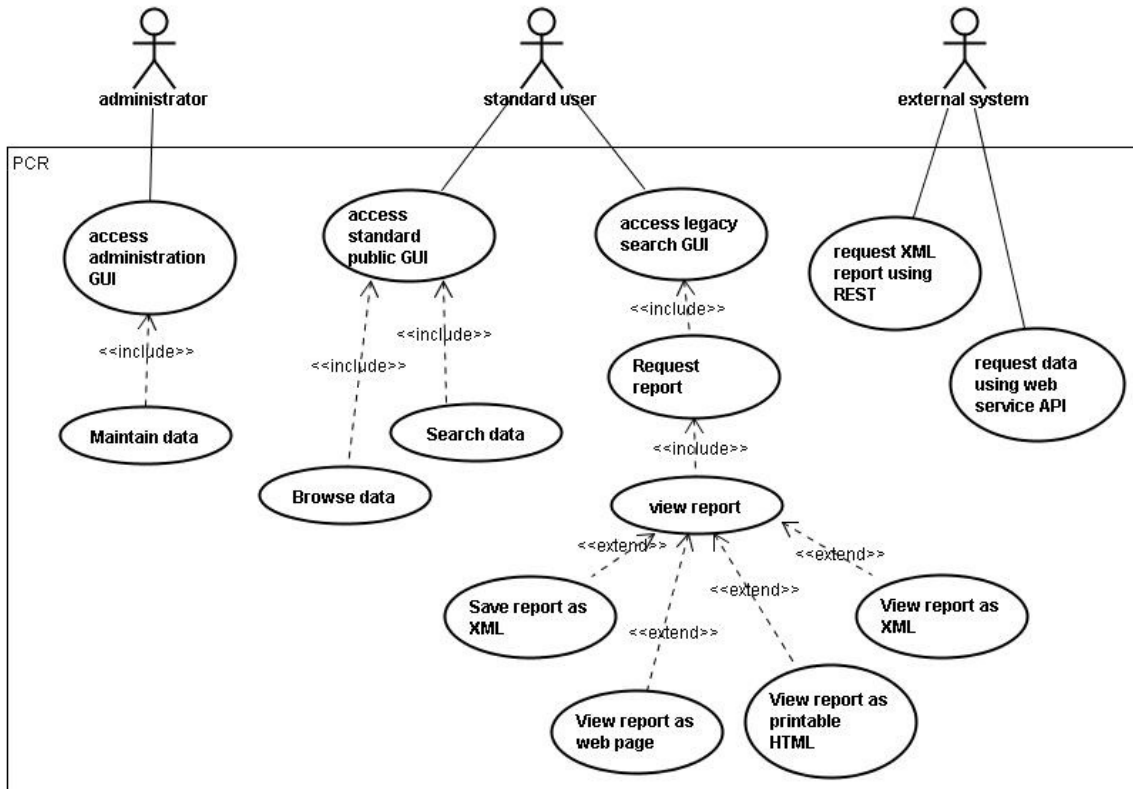


Figure 1: Main uses of the system.

2.4 Environmental Considerations

2.4.1 Planets instance

The Planets instance of the Core Registry will run in a Java application server (in fact only a web server is needed, so it will run on Tomcat alone). A database engine is also required but it is possible to choose almost any database technology (since the system uses Hibernate). Tests have been run on SQL Server 2005, Oracle 10g and Derby.

2.4.2 TNA's system

TNA will maintain the Core Registry in the same way as other users, that is by using the Java web-based administration application. However, to use the legacy search system a SQL Server 2005 database is required, which, if deployed, constrains the choice of database engine for the Java application. The legacy search system runs on the TNA preferred platform of Windows 2003 Server, and Microsoft Internet Information Server 6.

2.4.3 Web browsers

Both web-based GUIs are targeted at browsers that support XHTML1.0 Transitional and CSS 2, in particular those listed in section 9.1 of this document, which details the platform requirements.

3. Logical Model

3.1 Overview

According to the Conceptual Data Model [Planets Conceptual Data Model], a registry is needed to hold information about **File Formats and Environments**:

- **File Formats** are used not only to help decide what type of digital objects have been received but also to determine which tools to use for preservation activities, which properties to measure and to see which transformation pathways are possible.
- An **Environment** consists of a **Community** and a **Technical Environment**:
 - The Community can consist of information on **Producers, Consumers** and **Policy Factors** that might influence preservation planning.
 - A Technical Environment is a combination of **Software** and **Hardware**:
 - The Software that needs to be specified can include **Tools, Application Software, Operating Systems** and **Storage Management Software**. Tools can be classified as **Characterisation Tools, Migration Tools** or **Emulation Tools**. Characterisation Tools can be further sub-divided into **Format Identification Tools, Format Validation Tools, File Property Extraction Tools, Embedded Bytestream Extraction Tools, Component Discovery Tools** and **Component Property Measurement Tools**.
 - The hardware to be specified can include **Computer Hardware, Peripherals** and **Storage Media**.

In addition the Registry should hold information on **Properties** and **Transformation Pathways**:

- **Properties** can be classified as **Collection Properties, Deliverable Unit Properties, Component Properties** and **ByteStream Properties**. The latter are associated with a file format and can be subdivided into **Inherent Properties** (i.e. a property that all files in that format will share, e.g., the ubiquity of the format) and **Instance Properties** (i.e. a property that varies from file to file, e.g., image height or number of pages).
- **Transformation Pathways** can be classified as **Migration Pathways** (associated with a Migration Tool) or **Emulation Pathways** (associated with an Emulation Tool). Each Transformation Pathway can be verified by comparing Properties before and after transformation. The set of **Transformation Pathway Property Verifications** needed can also be recorded in the Registry.

The following class diagram illustrates the relationships between these entities:

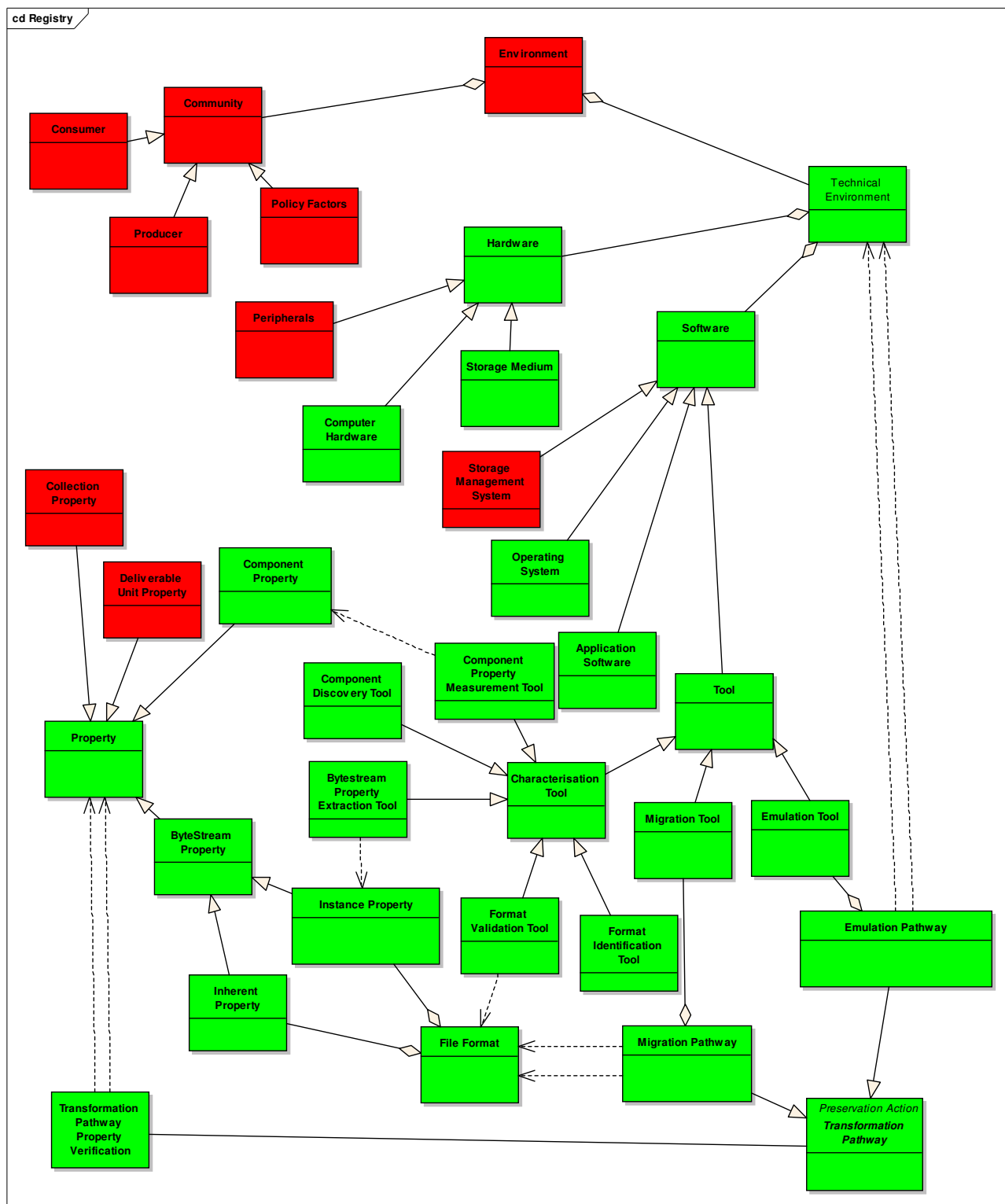


Figure 2: Planets Conceptual Data Model Class Diagram.

3.2 The Planets Core Registry

The Core Registry does not hold information about all of the above entities (and it also holds additional information).

The following table shows the main entities held (those coloured green in Figure 2):

Entity	Description
File Format	This describes a file format (a defined structure for the storage, display and processing of data).
Software (including Application Software, Operating System and all classes of Tools)	Software package tools are contained inside a software package, and represent functionality offered by that software. They are wrappers used to invoke a software package to perform a specific task.
Computer Hardware	This describes a physical computer system.
Storage Medium	This represents a physical medium on which digital material can be stored.
Technical Environment	This is an environment (a combination of items of hardware and software) in which software can run.
Character Encoding	Information about the different methods by which character data is stored on binary media.
Compression Technique	Information concerning the different methods by which binary data can be reduced in size to save space.
Property	This represents a property which file formats or components can possess.
Transformation Pathway	This is a conceptual entity that describes how to perform a transformation from one file format to another using software.

Hence, note that the Core Registry does **NOT** hold information about the following entities (coloured red in Figure 2):

Entity	Comment
Communities (including Consumers, Producers and Policy Factors)	This type of information is hard to describe and is concerned with the general environment of operation rather than the Technical Environment.
Storage Management Systems	No requirement to hold this information has been specified yet (but could be added easily)
Peripherals	No requirement to hold this information has been specified yet (but could be added easily)
Collection Properties and Deliverable Unit Properties	The list of available properties are more likely to be held in cataloguing systems or documented in descriptive metadata schemas like MODS or Dublin Core

3.3 Definition of information held

The following sections describe each of the main entities in turn.

3.3.1 File Format

The following table describes the information held about file formats:

Field	Description	Comment
PUID	This is the PRONOM Persistent Unique IDentifier for this file format.	
Name	This is the name of this file format	
Version	This is the format version.	
Description	This is a brief description of this file format.	

Field	Description	Comment
Release Dates	These are the dates that this format was released to the various global markets.	
Notes	This holds any other information concerning this file format.	
Provenance information	This holds details about where the information about the file format came from.	
Byte order	This defines the Endianness of files stored in this format.	Little-endian, big-endian or both
Disclosure levels	This represents the level to which this file format's specification is publicly disclosed.	Full, partial, none or unknown
Orientation	This is the byte orientation of files of this format.	Text or binary
Inherent properties	These are the properties that all file formats possess.	
Risk score (based on the information for inherent properties)	This is a dimensionless quantity which can be used to compare the risk of file formats. Its use is purely comparative and has no absolute meaning.	This can be calculated via a wizard using the risks associated with specific values of the file format's inherent properties.
Component manifestation types	Where a file format has components, these represent the physical manifestation of these components in this file format.	For example, PDF 1.4 has a document component, and the Component Manifestation Type is PDF.
Assessments	These are technical assessments of this file format.	
External signatures	These are external signatures that files of this format typically possess.	File extension is an example.
Instance properties	These are properties that files of this format possesses.	
Internal signatures	These are internal signatures (i.e. sequences of bytes) that can be used to identify files of this type.	
Reference files	These are example files of this file format.	
Technical environments	These are technical environments that this file format can execute natively in.	
XCDL schema files	These are the schemas for any PLANETS XCDL documents describing this format.	
XCEL files	These are any PLANETS XCEL documents describing this format.	
Processes	These are the software processes that interact with this file format.	The type can be Create, Render, Identify, Validate, or Extract Metadata..
Agents	These are agents who have links to this file format.	The role can be Support, Developer, Information Provider or Unknown.
Aliases	These are the aliases (i.e. other names) by which this file	

Field	Description	Comment
	format is also known.	
Documents	These are any documents concerning this file format.	
External identifiers	These are any external identifiers that can be used to recognise this file format.	
IPRs	These hold details of any IPR information concerning this file format.	
Relationships	These are relationships to other entities. They are directional, and hence come in two forms: source and target.	The relationship can be one of Is a subsequent version, Is a previous version, Is a subtype of, Is a supertype of, Can contain, Can be contained by, Has priority over, Equivalent to, Has lower priority than, Other, Requires, Required by, or Is default for.
Entity Types	These are used to classify file formats.	It can be Presentation, Text (mark-up), Text (word-processed), Audio, Spreadsheet, Image (Vector), Page Description, Video, Text (unstructured), Email, Database, Text (structured) or Image (raster)
Withdrawn details	These indicate whether or not the file format has been (or will be on a specific date) withdrawn by its producers and the date on which this occurred, if known.	
Facets	These are references to the GDFR facets.	

3.3.2 Software

Field	Description	Comment
PUID	These are the PRONOM Persistent Unique IDentifiers.	
Name	This is the name of this software package.	
Version	This is the software version	
Description	This is a brief description of this software package.	
Release Dates	These are the dates that this software was released to the various global markets.	
Notes	This holds any other information concerning this software package.	
Provenance information	This holds details about where this information about this software package came from.	
Source availability	This describes the public availability of the source code of this software.	Public, restricted, not available or unknown.
Package availability	This describes the public availability of the binaries for this software.	Public, restricted, not available or unknown.

Service pack level	This contains any information concerning service pack levels for this software.	
Associated tools	Tools describe the functionality offered by this software package.	
Processes	Processes are a fine-detail look at how this software interacts with file formats, technical environments or components.	
Software package locations	These describe places where one can obtain this software.	
Software package languages	These are the human languages that this software is available in.	
Software package interfaces	These are the ways in which a user can interface with this software.	Command line, GUI, Web service, API or other.
Software package components	These are any components that make up this software.	e.g. dlls
Hosting technical environments	These are any technical environments that this software package can run under.	
Target technical environments	These are the technical environments that this software is a part of.	
Software package licences	This describes any licences that cover this software.	
Software package images	These are any pictures of the software package	
Programming languages	These are any programming languages that were used to write this software.	
Agents	These are agents who have links to this software package.	The role can be Support, Developer, or Unknown.
Aliases	These are aliases by which this software package is also known.	
Documents	These are any documents concerning this software package.	
External identifiers	These are any external identifiers that can be used to recognise this software package.	
IPRs	These hold details of any IPR information concerning this software.	
Relationships	These are relationships to other entities. They are directional, and hence come in two forms: source and target.	The relationship can be one of Is a subsequent version, Is a previous version, Is a subtype of, Is a supertype of, Can contain, Can be contained by, Equivalent to, Other, Uses, Used by, Requires, Required by, or Available on.
Entity Types	These are used to classify software packages.	Possible types include operating system, calendar Spreadsheet, Video editing, Email, Database, desktop

		publishing, graphics, web design etc.
Withdrawn details	These indicate whether or not the software package has been (or will be on a specific date) withdrawn by its producers and the date on which this occurred, if known.	
Facets	These are references to the GDFR facets.	

3.3.3 Computer Hardware

Field	Description	Comment
PUID	These are the PRONOM Persistent Unique IDentifiers.	
Name	This is the name of this hardware.	
Version	This is the manufacturer's designated version of the hardware.	
Description	This is a brief description of this hardware.	
Release Dates	These are the dates that this hardware was released to the various global markets.	
Notes	This holds any other information concerning this hardware.	
Provenance information	This holds details about where this information came from.	
Components	These are any components that make up this hardware.	
Images	These are any pictures of this hardware.	
Agents	These are agents who have links to this hardware.	The role can be Support, Developer, or Unknown.
Aliases	These are other names by which this hardware is also known.	
Documents	These are any documents concerning this hardware.	
External identifiers	These are any external identifiers that can be used to recognise this hardware.	
IPRs	This holds details of any IPR information concerning this hardware.	
Relationships	These are relationships to other entities. They are directional, and hence come in two forms: source and target.	The relationship can be one of Is a subsequent version, Is a previous version, Is a subtype of, Is a supertype of, Can contain, Can be contained by, Equivalent to, Other, Uses, Used by, Requires, or Required by.
Entity Types	These are used to classify hardware.	
Withdrawn details	These indicate whether or not the hardware has been (or will	

	be on a specific date) withdrawn by its manufacturers and the date on which this occurred, if known.	
Facets	These are references to the GDFR facets.	

3.3.4 Storage Medium

Field	Description	Comment
PUID	These are the PRONOM Persistent Unique IDentifiers.	
Name	This is the name of this storage medium.	
Version	This is the manufacturer's designated version of this storage medium.	
Description	This is a brief description of this storage medium.	
Release Dates	These are the dates that this medium was released to the various global markets.	
Notes	This holds any other information concerning this medium.	
Provenance information	This holds details about where this information came from.	
Images	These are any pictures of this storage medium	
Agents	These are agents who have links to this storage medium.	The role can be Support, Developer, or Unknown.
Aliases	These are alternative names by which this storage medium is also known.	
Documents	These are any documents concerning this storage medium.	
External identifiers	These are any external identifiers that can be used to recognise this storage medium.	
IPRs	These hold details of any IPR information concerning this storage medium.	
Relationships	These are relationships to other entities. They are directional, and hence come in two forms: source and target.	The relationship can be one of Is a subsequent version, Is a previous version, Is a subtype of, Is a supertype of, Can contain, Can be contained by, , Equivalent to, , Other, Uses, Used by, Requires, Required by.
Entity Types	These are used to classify storage media.	
Withdrawn details	These indicate whether or not the storage medium has been (or will be on a specific date) withdrawn by its manufacturers and the date on which this occurred, if known.	
Facets	These are references to the GDFR facets.	

3.3.5 Technical Environment

Field	Description	Comment
PUID	These are the PRONOM Persistent Unique IDentifiers.	
Name	This is the name of this technical environment.	
Description	This is a brief description of this technical environment.	
Documents	These are any documents about this technical environment.	
Hardware	This is the hardware which comprises this technical environment.	
Software packages	These are the software packages that comprise this technical environment.	
Facets	These are references to the GDFR facets.	

3.3.6 Character Encoding

Field	Description	Comment
PUID	These are the PRONOM Persistent Unique IDentifiers.	
Name	This is the name of this character encoding.	
Version	This is the version of this encoding.	
Description	This is a brief description of this encoding.	
Release Dates	These are the dates that this encoding was released to the various global markets.	
Notes	This holds any other information concerning this encoding.	
Provenance information	This holds details about where this information came from.	
Code Unit Width Text	This relates to the technical definition of this encoding.	
Encoding Form Width Text	This relates to the technical definition of this encoding.	
Code Page	This is the EBCDIC reference.	
Agents	These are agents who have links to this character encoding.	The role can be Support, Developer, or Unknown.
Aliases	These are alternative names by which this character encoding is also known.	
Documents	These are any documents concerning this character encoding.	
External identifiers	These are any external identifiers that can be used to recognise this character encoding.	
IPRs	These hold details of any IPR information concerning this character encoding.	
Relationships	These are relationships to other entities. They are directional, and hence come in two forms: source and target.	The relationship can be one of Is a subsequent version, Is a previous version, Is a subtype of, Is a supertype of, Can

		contain, Can be contained by, Equivalent to, Other, Uses, Used by, Requires, or Required by.
Entity Types	Used to classify character encodings.	
Withdrawn details	These indicate whether or not the character encoding has been (or will be on a specific date) withdrawn by its producers and the date on which this occurred, if known.	
Facets	These are references to the GDFR facets.	

3.3.7 Compression Technique

Field	Description	Comment
PUID	These are the PRONOM Persistent Unique IDentifiers.	
Name	This is the name of this compression technique.	
Version	This is the version of this compression technique.	
Description	This is a brief description of this compression technique.	
Release Dates	These are the dates that this compression technique was made available to the various global markets.	
Notes	This holds any other information concerning this compression technique.	
Provenance information	This holds details about where this information came from.	
Lossiness Type	This describes the nature of this compression.	Lossless, Lossy, or Unknown.
Agents	These are agents who have links to this compression technique.	The role can be Support, Developer, or Unknown.
Aliases	These are alternative names by which this compression technique is also known.	
Documents	These are any documents concerning this compression technique.	
External identifiers	These are any external identifiers that can be used to recognise this compression technique.	
Entity Families	There are families of related compression techniques, which this compression technique is a member of.	
IPRs	These hold details of any IPR information concerning this compression technique.	
Relationships	These are relationships to other entities. They are directional, and hence come in two forms: source and target.	The relationship can be one of Is a subsequent version, Is a previous version, Is a subtype of, Is a supertype of, Can contain, Can be contained by, Equivalent to, Other, Uses, Used by, Requires, or

		Required by.
Withdrawn details	These indicate whether or not the compression technique has been (or will be on a specific date) withdrawn by its producers and the date on which this occurred, if known.	
Facets	These are references to the GDFR facets.	

3.3.8 Property

Field	Description	Comment
PUID	These are the PRONOM Persistent Unique IDentifiers.	
Name	This is the name of this property.	
Description	This is a brief description of this property.	
Property class	This defines what kind of property it is.	Inherent or instance.
Component Type (Component properties only)	If the property is a component property, then this describes the type of component to which it relates.	Document, Image, Video, Audio, Spreadsheet, Presentation or Container.
Property options	If the property has options for its potential values, these are listed here.	
Facets	These are references to the GDFR facets.	

3.3.9 Transformation Pathway

Field	Description	Comment
PUID	These are the PRONOM Persistent Unique IDentifiers.	
Name	This is the name of this pathway.	
Description	This is a brief description of this pathway.	
Pathway type	This classifies the type of pathway.	File format migration, file format object extraction or technical environment emulation.
Pathway steps	These are the stages of this transformation. They can be chained to allow for complex transformations.	
Documents	These are any documents concerning this pathway.	
Pathway roles	These represent the functions for which this pathway has been approved.	The different possible roles are Presentation, Preservation and Extraction, and the approval statuses are Unknown, Current and Withdrawn.
Facets	These are references to the GDFR facets.	

3.4 Note on Component Manifestations

There is a certain amount of confusion as to what a component manifestation is and why it is necessary to store information about them; this section is an attempt to clarify what is meant by a component manifestation.

Component manifestations become important when the preservation of digital objects by migration is considered. A component (also known as record component or deliverable unit component) is a

logical entity that represents the smallest part of an archival record that can be considered for migration. That is, if any file (or more accurately a bytestream) that is part of a component becomes obsolete and needs migrating, then all files (or rather bytestreams) that make up the component will need to be included in the migration process (although some files (bytestreams) may not change during the migration). A component manifestation is the physical embodiment of a component in a particular technology or set of technologies.

In most cases there is a one-to-one mapping between components and files. However, a component may consist of several files or it may be only part of a file. Some examples may help to make this clear.

Example 1:

A component could be a web page then it is likely to consist of multiple files e.g. the HTML file containing the main content, the stylesheet, the images for the page, the header HTML file, and the footer HTML file.

Component = document

Component manifestation = web page in HTML

Files that make up the component: main HTML file, header HTML file, footer HTML file, CSS stylesheet, a number of JPEG images.

Example 2:

A component could be spreadsheet, such as an Excel spreadsheet, that has been embedded within an MS-Word document (the file).

Component = spreadsheet

Component manifestation = Excel spreadsheet

Bytestream that makes up the component: embedded Excel spreadsheet.

The types of components currently recognised by PCR are:

- Document
- Image
- Video
- Audio
- Spreadsheet
- Presentation
- Container

3.5 A More Detailed View of the Data Model

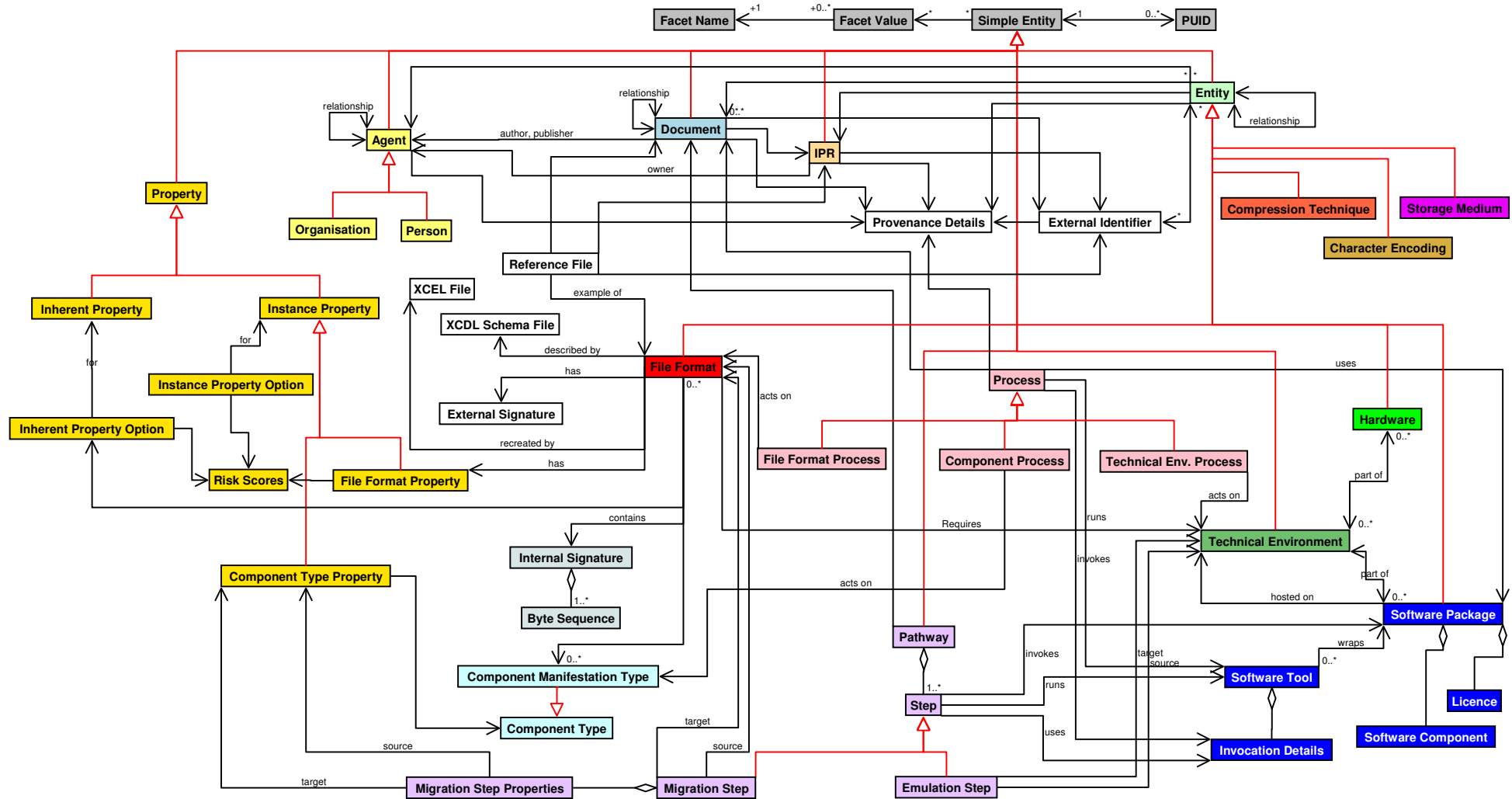


Figure 3: An Intermediate Level View of the Planets Core Registry Data Model

Figure 3 shows an intermediate level view of the data model. It does not show all the classes present in the model, but it does show the main ones and their relationships, both inheritance and associations. In particular, it does not show the 36 enumeration, or 'list of values', tables, nor does it show the system management tables which include the audit tables.

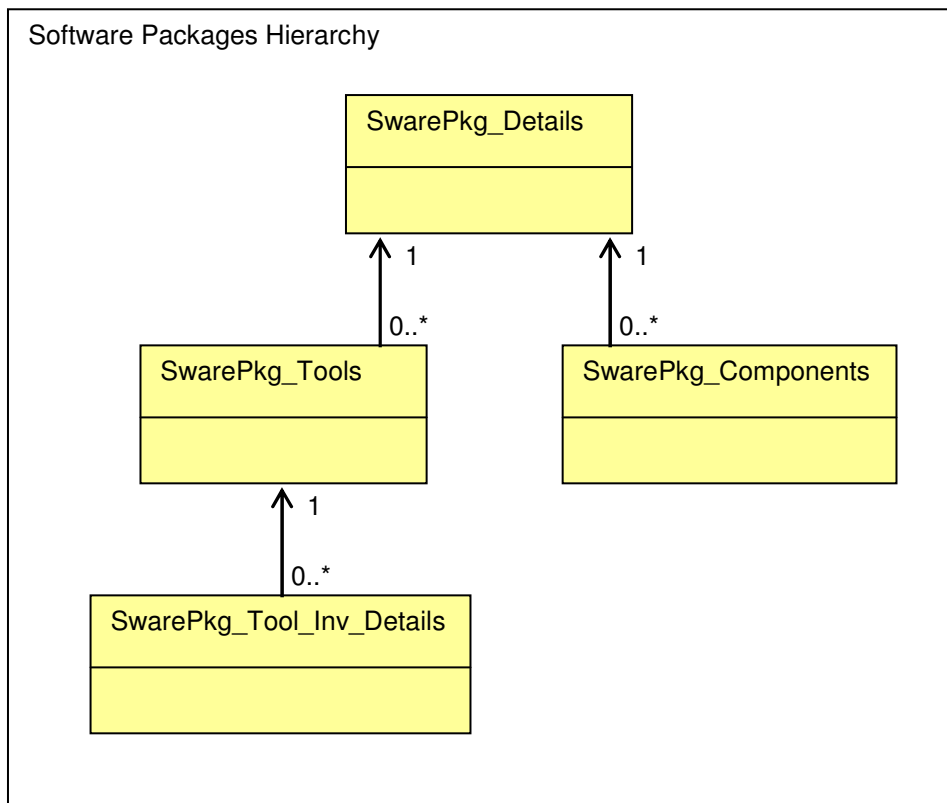
4. Issues and Discussion

4.1 Clarifications of Specific Issues

This section contains discussions on various issues that have been resolved during requirements gathering. The resolutions are discussed here to ensure that the background information is not lost.

4.1.1 Software Packages

The highest-level entity relating to software is a software package and it is used to record information that is *common* to all of the different forms of software. Below this, there is a hierarchy of associated details about software packages.



- Software packages can have multiple software package components.
 - For example, all the dlls used by Word 2007 could be listed by name as software package components for the software package “Word 2007”.
- Software packages can have multiple software tools
 - Software tools are wrappers to subsets of functionality within software packages. This wrapper provides a way of invoking the Software Package.
 - Software tools can be Planets services, which are defined as “tool[s] that can be invoked from a site hosting PLANETS related resources.” in the requirements for the Preservation Action Registry.
 - Software tools can have invocation details associated with them.

In this structure:

- Software packages are equivalent to 'software products' in the Preservation Action Registry Requirements, and 'software' in PRONOM 6.2.
- Software tools are equivalent to 'software tools' in both PRONOM 6.2 and the Preservation Action Registry Requirements, albeit with the additional constraint that a software tool must belong to a software package.
- Software tools can be defined as being a Planets Service if they are available to call within Planets (in the Planets Interoperability Framework or PIF) and the tool's invocation details record the details of the entry for the tool in the Planets Services Registry.

4.1.2 Classification

The intention behind the requirement from PC3/4, to 'replace the existing family¹ and type classification' with GDFR classification is open to interpretation. Discussions [MEETING-WIN/08/08/05] concluded that a faceted classification system is required and it should be applicable as broadly as possible to entities in the Technical Registry.

To meet this requirement, it will be necessary to use a subset of the GDFR Facet Value Name pairs combined with additional Facet Name Value pairs to meet the classification requirements of the Technical Registry.

As the list of Facet Names and Facet Values is not definitive, the admin interface will not enforce any rules to control the use of the classification. Therefore, it will be possible to associate any number of Facet Name Value pairs with an entity, despite the fact that the combination (or even the individual values) does not necessarily make sense. The onus will be on administrators to use the Facet Values sensibly.

For example, the following facet values could be assigned to the file format PNG 1.0:

genre:still-image

role:encoding

form:binary

which does not make sense; the role should be file format and not encoding.

If a concrete set of Facet Values could be determined, then it would be possible to build rules into the admin interface to ensure that the added data makes sense.

As for families, they will now be represented by an entity facet value GDFR:role:family. While it makes sense that the aggregate (i.e. family) of certain entities are themselves an entity of that type, this is not the case for all entities. So, a suite of software is still a software package and a collection of bits of hardware is a computer which is still a piece of hardware, but a family of file formats (e.g. TIFF) is not itself a file format and so cannot be modelled as a file format with facet value GDFR:role:family.

In addition, some information that is currently recorded in Pronom 6.2 will be lost if the family and type classification system is removed. For example, in Pronom 6.2 it is possible to associate a File Format Family, WAVE for example, with multiple File Formats. If the family classification system is removed, there is no replacement concept that an entity with a facet value GDFR:role:family can have members that must be File Formats.

However, families are little used in Pronom 6.2; indeed there are only 4 file format families (WAVE, DXF, DWG, RTF) and no families of any other type of entity. Since very little information is recorded about these families, it was concluded that they are not a necessary part of the data model and can be deprecated in PCR version 3 and removed from a future version if no problems are encountered by deprecating entity families.

4.1.3 XCDL/XDEL files

Requirement PC3/D4 – 2.5 implies that both XCDL and XCEL files should be stored in the database.

However, an XCDL xml file is specific to a particular instance of a File Format, and so would have to be stored against a reference file of a particular format. Additionally, the XCDL xml document for a given file may change over time, and so should not be stored and retrieved but recalculated on

¹ A family is a group of related entities. For example the PDF family would include all the entities representing the different versions of the PDF file format.

demand to ensure that the information it contains is accurate. Alternatively, an example file of a given format would have to be uploaded in order for an XCDL to be derived for it.

Therefore, in PRONOM 7, and thus in version 2.1 of the Core Registry, the requirement has been interpreted that XCDL schema files (and XCEL files) should be stored and associated with a file format.

4.1.4 PUIDS

The following entities will have PUIDs:

- all core entities
 - file formats
 - software packages
 - hardware
 - character encoding
 - compression techniques
 - storage media
- technical environments
- pathways
- properties

Note that PCR 2 allows PUIDs to be assigned to agents, documents and IPRs; this is not required and this functionality will be dropped in PCR 3.

PUID assignment will happen on creation of a record that should have a PUID.

The PUID assigned will be of the form PUID_TYPE/NUMBER where PUID TYPE is one of the following:

fmt – for file formats
sfw – for software packages
hdw – for hardware
chr – for character encoding
cmp – for compression techniques
med – for storage media
env – for technical environments
pth – pathways
prp - properties

A framework for the allocation of PUIDs in PCR 3 has been agreed between TNA and Tessella. While the main aim is to ensure (as far as possible) the uniqueness of any allocated PUID, the framework deliberately has been kept as simple as possible. Also, it aims to enable the current way of working without precluding any of the likely future ways that registry will be deployed.

It will be the responsibility of human administrators (via agreements and vigilance) to ensure that PUIDs do not clash between instances, and the responsibility of the software to ensure that each entity within an instance has a unique PUID. While the responsibilities for PUID allocation for file formats, software packages, processes and pathways are clear, there is an overlap of concerns around the edges. For this reason, the idea of allocating ranges of numbers for each type of PUID to each instance of PCR was introduced.

Each installed instance of PCR will, by default, have no allocated ranges of PUIDs for any entities and so will not be able to create any new entities which require PUIDs. There would be human processes and agreements in place for the allocation of PUID ranges to individual instances of PCR; this is outside the scope of the PCR software. There would be a separate allocation (or none) of a range of PUID numbers for each type of entity that is identified by a PUID (i.e. all the core entities, plus pathways, properties and technical environments).

For example:

Instance A could be allocated the following PUID ranges:

- file formats: 1 – 10,000
- software packages: 0
- pathways: 10,001 – 10,300
- etc.

and instance B (used to enter details of software packages and pathways) could be allocated the following PUID ranges:

- file formats: 0
- software packages: 1 – 10,000
- pathways: 1 – 10,000
- etc.

Note that the allocation of ranges of PUIDs is not something that is currently required as Planets will initially only have one, central instance of PCR. So there is no need for any synchronisation between instances. However, it is envisaged that in the future there may be multiple installations of PCR and so the idea of allocating ranges was explored so as not to prevent this future usage.

There will be two ways of allocating a PUID (the number part; the PUID type is determined by the type of the entity) to an entity:

1. A function to allocate the next available PUID (of the correct type) to an entity. This function will only allocate PUIDs from within the allocated ranges and will not be able to create a new PUID if there are no unused PUIDs left in the range (or if the range is empty or set to zero).
2. A function to check whether a user-entered PUID is unique within an instance of Pronom. This function will also check whether the PUID is within the allocated range for that type of entity. It will block the user from saving the entity if the PUID is not unique within that instance of the registry. It will warn the user if the PUID is outside the allocated range for the entity in the registry instance, but allow the user to override the warning and save the entity if the user chooses to do so. This will allow entities that have been assigned PUIDs in one instance of PCR to be manually entered into a different instance of PCR.

Note that the key aim of the software requirements for the allocation of PUIDs is to prevent PUID duplication within an instance of the core registry.

4.1.5 Testbed Results Submission

The submission of Testbed results was discussed in detail at a meeting between people involved in the Testbed and PCR sub-projects (see [TESTBED-AUG-2009]). The set of requirements included in this document was agreed. What was also agreed, was what the ideal implementation that Testbed would like to see was, although this is not binding upon the implementer of these requirements without researching the feasibility in the development time available. This ideal implementation is as follows:

1. Add a CLOB field to the pathway steps table to store aggregated Testbed results.
2. Implement a web service for the submission of Testbed results (as a CLOB probably containing an XML document) together with enough information for an administrator to uniquely identify the pathway step and which Testbed instance supplied the results. The information will be stored in a quarantine area.
3. Create a web page to allow a Testbed person with PCR administration rights to review the Testbed results (probably an XML document) and select the appropriate PCR pathway step and save it to the PCR database with a single button press.
4. The system must implement a web service to retrieve the Testbed results for a pathway step.

4.1.6 InSPECT Significant Properties Data Model

The InSPECT project (<http://www.significantproperties.org.uk>) has defined a data model for significant properties of file formats (see [INSPECT_DICTIONARY]). This data model provides a framework that allows for both the storage of information about such properties and the recording

of measurements of the actual properties of a specific file. As such, only a subset of the data model relating to properties is relevant for PCR. Comparing the two data models we can see how closely the PCR data model for properties maps onto InSPECT's.

InSPECT Semantic Units for Properties	PCR Data Model for Properties
Identifier	PUID
Identifier Type	PUID type
Identifier Value	PUID value
Title	Simple entity name
Description	Simple entity description
Function – class (restricted to content, context, structure, rendering, and behaviour)	This could be mapped onto “simple entity facet name”, if the GDFR facet model is extended appropriately.
Function – sub-class (InSPECT recommend the development of a controlled vocabulary, but one does not currently exist).	This could be mapped onto “simple entity facet value”, if the GDFR facet model is extended appropriately.
Preservation Level	This is not information that should be stored in the PCR, but is something that organisations will want to determine and record in a suitable location.
Specification Registry	This is implied because the data is stored in PCR.
Measurement	<p>This is for recording the value of a specific property of a file before and after transformation; therefore it is not something that should be stored in the PCR.</p> <p>However, it could be useful to store the units that a measurement is made in (e.g. pixels for image width).</p> <p>The allowed variance in measurements for the same instance property before and after migration is already stored for component type instance properties involved in a migration pathway step.</p>
Genre	This performs the same task as ‘function – subclass’ and so could be handled in the same way, i.e. by mapping onto facets.
Genre – Resource Type (e.g. raster image, still image, audio-waveform, moving image, text)	The resource type of a genre is defined as “the category in which a property may be classified”. However, the examples given are all PCR component types. A file format can be associated with one or more component manifestations, which are of a given component type. Therefore, an instance property, whether that is a component type property or file format type property, can be indirectly associated with a component type.

InSPECT Semantic Units for Properties	PCR Data Model for Properties
Genre – Format Type (e.g. MIME type, PUID)	The format type of a genre is defined as “the encoding format in which a property is defined”, but the examples given are file format types, not character encoding types. File format instance properties are associated with a file format and this has a PUID (as do properties themselves) and it may also have a MIME type external identifier.

Therefore, the PCR data model has the capability to store all the data required by the InSPECT property data model.

4.2 Outstanding Issues and Enhancements

This section details issues that have not been fully resolved at the time of writing, as well as ideas for future enhancements which have not been fully specified. Resolutions for the issues are being sought and the document will be updated accordingly in due course.

4.2.1 Invocation Details

As discussed in [MEETING-WIN/08/08/05], software package tools may come in a variety of forms. Some may be part of the Planets Interoperability Framework, some may be available only locally, some may have entries in the Planets Services Registry, and some may be available outside of Planets.

The current data model for software package tools allows details of how to invoke these different types of tools to be stored. Since there are different ways to invoke different types of services, the data model currently stores the invocation details as free text; thus it is up to the administrator to ensure that meaningful information is entered for a tool’s invocation details. In future, when more is known about what types of tools will be stored in the core registry and what types of information need to be retained on how to invoke the tools, the data model can be updated to store more specific information on the details of how to invoke a tool.

4.2.2 Advanced Software Search

Following discussions [MEETING-WIN/08/08/05], the implementation of an advanced software search has been delayed until a later release of the Core Registry.

It was recognised that the fields in Figure 4 were considered “key fields” for an advanced search for Software Packages: Defining the layout and logic for combining the fields into a search was considered to be out of scope of this iteration.

Advanced Search

entity type

select fields to include in search:

name	<input type="text"/>	puid value	<input type="text"/>
version	<input type="text"/>	alias	<input type="text"/>
associated people	<input type="text" value="select person..."/>	identifiers	<input type="text"/>
description	<input type="text"/>	associated processes	<input type="text" value="select process type ..."/>
associated organisations	<input type="text" value="select organisation..."/>	withdrawn date	from <input type="text"/>
release date	from <input type="text"/>	to <input type="text"/>	
	to <input type="text"/>		
support status	<input type="text" value="select support status..."/>		
associated file formats	<input type="text" value="select name and version..."/>		
	<input type="text" value="select file format extension ..."/>		
support period	from <input type="text"/>		
	to <input type="text"/>		

Software Package Search Fields

These fields are only applicable to searches for Software Packages.

service pack level

Figure 4 - Advanced software search fields

4.2.3 Version Control, Rollback and Synchronisation

It is envisaged that in the future there may be multiple instances of PCR and that the data held in them would need to be kept in sync. Therefore, the concepts of version control, rollback and synchronisation, which would build on the existing auditing functionality, have been introduced and discussed at a high level. The details of what is actually required have not been determined yet. However, the basics are included in this document so they are not lost.

More discussion of the user requirements for version control, rollback and synchronisation is expected to be covered in a future vision document, which is forthcoming.

5. Data Requirements

5.1 General Requirements

5.1.1 Main information

Label	Requirement	Necessity
S5.1.1.1	<p>The system must be able to store information on the following core entities:</p> <ul style="list-style-type: none"> • File formats. • File format character encodings. • File format compression techniques. • Software. • Hardware. • Storage Media <p>Source: [P3SD, Current system, P4IM, Clarification 2, TWSRD] Source: [Pronom6 SRD - S4.1.1.1]</p>	M
S5.1.1.2	<p>The system must be able to store information on the following subsidiary entities:</p> <ul style="list-style-type: none"> • File format reference files. • File format external signatures (e.g. file extension). • File format internal signatures. • Software processes that can occur to file formats. • People and organisations. • Documentation. • Intellectual Property Rights. • Pathways (migration, emulation, and extraction). • Software package tools (e.g. for characterisation and migration). • Properties of file formats and deliverable unit components. • Technical environments. • Component manifestation types. • Images. • Reference files. • Classifications. <p>Source: [P3SD, Current system, P4IM, Clarification 2, TWSRD] Source: [Pronom6 SRD - S4.1.1.1]</p>	M
S5.1.1.3	<p>It must be possible to record the relationship between two entities of the same type where the type is one of the following:</p> <ul style="list-style-type: none"> • File formats. • File format character encodings. • File format compression techniques. • Software packages. • Hardware. • Storage media. • Documentation. <p>Source: [P4IM, Clarification 2, TWSRD] Source: [Pronom6 SRD - S4.1.1.2, Pronom7 SRD – S5.1.1.12, Pronom7 SRD – S5.2.3.6]</p>	M

Label	Requirement	Necessity
S5.1.1.4	<p>When recording the relationship between two entities of the same type, the following information must be recorded:</p> <ul style="list-style-type: none"> • The source entity. • The target entity. • The relationship type (linked to the maintained list of relationship types). • The relationship note. <p>Source: [P4IM] Source: [Pronom6 SRD - S4.1.1.3, Pronom7 SRD – S5.1.1.13, Pronom7 SRD – S5.2.3.7]</p>	M

5.1.2 Families

Label	Requirement	Necessity
S5.1.2.1	Future enhancement, so removed from this document.	X

5.1.3 Types

Label	Requirement	Necessity
S5.1.3.1	<p>The system must maintain a list of entity types that can be used to categorise the core entities (e.g. email, database, spreadsheet); although these should be deprecated in favour of using GDFR facets from PCR version 3 onwards.</p> <p>Source: [P4IM] Source: [Pronom6 SRD - S4.1.5.1, Pronom7 SRD – S5.6.16.1]</p>	M
S5.1.3.2	<p>The system must be able to maintain a list of component manifestation types (e.g. document type, image type etc.); although these should be deprecated in favour of using GDFR facets from PCR version 3 onwards.</p> <p>Source: [Tessella] Source: [Pronom7 SRD – S5.6.6.1]</p>	M
S5.1.3.3	<p>The system must maintain a list of file format external signature types.</p> <p>Source: [P4IM] Source: [Pronom6 SRD - S4.1.5.2, Pronom7 SRD – S5.6.18.1]</p>	M
S5.1.3.4	<p>The system must maintain a list of file format orientations (e.g., text, binary); although these should be deprecated in favour of using GDFR facets from PCR version 3 onwards.</p> <p>Source: [P4IM] Source: [Pronom6 SRD - S4.1.6.4]</p>	M
S5.1.3.5	Future enhancement, so removed from this document.	X
S5.1.3.6	Future enhancement, so removed from this document.	X
S5.1.3.7	<p>The system must maintain a list of possible documentation types. These should include:</p> <ul style="list-style-type: none"> • Unknown • Authoritative • Informative • Speculative <p>Source: [P4IM, PA-tool registry - 5.5.1, PA-tool registry - 5.5.3, PA-tool registry - 5.5.7] Source: [Pronom6 SRD - S4.1.5.6, Pronom7 SRD – S5.6.12.1]</p>	M
S5.1.3.8	<p>The system must maintain a list of possible organisation types.</p> <p>Source: [P4IM] Source: [Pronom6 SRD - S4.1.5.7, Pronom7 SRD – S5.6.4.1]</p>	M
S5.1.3.9	<p>The system must maintain a list of agent relationship types.</p> <p>Source: [Tessella] Source: [Pronom7 SRD – S5.6.2.1]</p>	M
S5.1.3.10	<p>The system must maintain a list of possible intellectual property right types.</p> <p>Source: [P4IM, PA-tool registry - 5.4.1] Source: [Pronom6 SRD - S4.1.5.8, Pronom7 SRD – S5.6.24.1]</p>	M

S5.1.3.11	The system must maintain a list of possible pathway roles. These should include: <ul style="list-style-type: none"> • Preservation • Presentation • Extraction Source: [Tessella] Source: [Pronom6 SRD - S4.1.5.9, Pronom7 SRD – S5.6.26.1]	M
S5.1.3.12	The system must maintain a list of External Identifier Types such as URLs, MIME types and ISBNs. Source: [Tessella] Source: [Pronom7 SRD – S5.6.17.1]	M
S5.1.3.13	The system must maintain a list of pathway types such as extraction and migration of objects or technical environments. Source: [Pronom7 SRD – S5.6.28.1]	M
S5.1.3.14	The system must be able to maintain a list of process action types such as actions on objects (i.e. migration) and actions on technical environments (i.e. emulation). Source: [PA-tool registry - 5.2.6] Source: [Pronom7 SRD – S5.6.29.1]	M

5.1.4 Other information

Label	Requirement	Necessity
S5.1.4.1	The system must maintain a list of core entity types. This must use the following entity type codes together with appropriate display names and descriptions: <ul style="list-style-type: none"> • FILE_FORMAT • SOFTWARE_PACKAGE • HARDWARE • COMPRESSION_TECHNIQUE • CHARACTER_ENCODING • STORAGE_MEDIA Source: [Tessella] Source: [Pronom7 SRD – S5.6.1.1, Pronom7 SRD – S5.6.1.2, Pronom7 SRD – S5.6.13.1]	M
S5.1.4.2	The system must maintain a list of countries. Source: [P3SD – 3.2.7.2, Current system] Source: [Pronom6 SRD - S4.1.6.1, Pronom7 SRD – S5.6.9.1]	M
S5.1.4.3	The system must maintain a list of languages. Source: [P3SD – 3.2.7.2, P4IM] Source: [Pronom6 SRD - S4.1.6.2, Pronom7 SRD – S5.6.25.1]	M
S5.1.4.4	The system must maintain a list of intellectual property right jurisdictions. Source: [P4IM] Source: [Pronom6 SRD - S4.1.6.3, Pronom7 SRD – S5.6.23.1]	M
S5.1.4.5	The system must maintain a list of file format byte orders. This should include the following: <ul style="list-style-type: none"> • Big-endian • Little-endian • Big-endian and little-endian • Big-endian or little-endian • Not applicable • Unknown (default) Source: [P4IM, PC/3 – D4 – 2.20] Source: [Pronom6 SRD - S4.1.6.5, Pronom7 SRD – S5.6.20.1]	M
S5.1.4.6	The system must maintain a list of file format disclosure levels (e.g., full, partial, none etc.). Source: [P3SD – 3.2.7.2, P4IM] Source: [Pronom6 SRD - S4.1.6.6, Pronom7 SRD – S5.6.21.1]	M
S5.1.4.7	Future enhancement, so removed from this document.	X
S5.1.4.8	Future enhancement, so removed from this document.	X

Label	Requirement	Necessity
S5.1.4.9	The system must maintain a list of compression lossiness. Source: [P4IM] Source: [Pronom6 SRD - S4.1.6.9, Pronom7 SRD – S5.6.7.1]	M
S5.1.4.10	The system must maintain a list of possible documentation availabilities. Source: [P4IM] Source: [Pronom6 SRD - S4.1.6.10, Pronom7 SRD – S5.6.10.1]	M
S5.1.4.11	The system must maintain a list of byte sequence position types (e.g., “specified offset”, “relative to previous”). Source: [P4IM] Source: [Pronom6 SRD - S4.1.6.11, Pronom7 SRD – S5.6.5.1]	M
S5.1.4.12	The system must maintain a list of software process types (e.g., creates the format, renders the format, identifies the format, validates the format or extracts metadata about files in that format). Source: [P4IM, Clarification 2] Source: [Pronom6 SRD - S4.1.6.12, Pronom7 SRD – S5.6.30.1]	M
S5.1.4.13	The system must maintain a list of content variance types. Source: [P4IM, Clarification 2] Source: [Pronom6 SRD - S4.1.6.13, Pronom7 SRD – S5.6.8.1]	M
S5.1.4.14	The system must maintain a list of relationship types. This must include both a forward and reverse description of the relationship if the type is directional (i.e. ‘supersedes’ and ‘is superseded by’). Source: [P4IM, Clarification 2, Review] Source: [Pronom6 SRD - S4.1.6.14, Pronom7 SRD – S5.6.15.1]	M
S5.1.4.15	The system must maintain a list of agent roles. These should include: <ul style="list-style-type: none"> • Support provider • Developer • Information provider Source: [Tessella] Source: [Pronom7 SRD – S5.6.3.1]	M
S5.1.4.16	The system must maintain a list of possible documentation content types, such as user manuals and installation guides. Source: [PA-tool registry - 5.5.1, PA-tool registry - 5.5.3, PA-tool registry - 5.5.7] Source: [Pronom7 SRD – S5.6.11.1]	M
S5.1.4.17	The system must maintain a list of possible families that can be used to categorise the core entities. Source: [Tessella] Source: [Pronom7 SRD – S5.6.14.1]	M
S5.1.4.18	The system must maintain a list values to use in tri state (yes/no/unknown) fields. Source: [Tessella] Source: [Pronom7 SRD – S5.6.19.1]	M
S5.1.4.19	The system must maintain a list of possible pathway states. These should include: <ul style="list-style-type: none"> • Current • Withdrawn • Unknown Source: [Pronom7 SRD – S5.6.27.1]	M
S5.1.4.20	The system must be able to maintain a list of programming languages. Source: [PA-tool registry – SU1] Source: [Pronom7 SRD – S5.6.31.1]	M

Label	Requirement	Necessity
S5.1.4.21	The system must be able to maintain a list of software package interface types, such as GUI, command line interface and web services. Source: [PA-tool registry – SU1] Source: [Pronom7 SRD – S5.6.33.1]	M
S5.1.4.22	The system must be able to maintain a list of licence types. This list should include the following values: <ul style="list-style-type: none"> • Free • Commercial • Other • Unknown Source: [EMAIL-080728] Source: [Pronom7 SRD – S5.6.34.1]	M
S5.1.4.23	The system must maintain a list of component manifestation types (i.e. instances of particular component types; for example, a Microsoft Word document and an HTML page and its associated files are both instances of the component type 'document'), including the following information: <ul style="list-style-type: none"> • Name • Description • The component type it is an instance of. Source: [Tessella] Source: [Pronom7 SRD – S5.2.8.1, Pronom7 SRD – S5.2.8.2]	M
S5.1.4.24	The system must maintain a list of images including the following information: <ul style="list-style-type: none"> • Display name. • A textual description of the image. • The image itself Source: [Tessella] Source: [Pronom7 SRD – S5.2.8.1, Pronom7 SRD – S5.2.8.2]	M
S5.1.4.25	The system must maintain a list of licences including the following information: <ul style="list-style-type: none"> • Name. • Terms of the licence. • Type of the licence (linked to a list of authority controlled licence types). Source: [PC3/D4 – 2.17, EMAIL-080728, MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.2.15.1, Pronom7 SRD – S5.2.15.2, Pronom7 SRD – S5.2.15.3]	M
S5.1.4.26	The system must maintain a list of XCDL schema files, including the following information: <ul style="list-style-type: none"> • Display name • Description • The file itself. Source: [PC3/D4 – 2.4, EMAIL-080724, MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.2.16.1, Pronom7 SRD – S5.2.16.2]	M
S5.1.4.27	The system must maintain a list of XCEL files, including the following information: <ul style="list-style-type: none"> • Display name • Description • The file itself. Source: [PC3/D4 – 2.4, EMAIL-080724, MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.2.17.1, Pronom7 SRD – S5.2.17.2]	M

5.1.5 PUID Assignment

Label	Requirement	Necessity
-------	-------------	-----------

Label	Requirement	Necessity
S5.1.5.1	<p>The Core Registry must allow PUIDs to be assigned to the following entities:</p> <ul style="list-style-type: none"> • File formats • Software • Hardware • Character encodings • Compression techniques • Storage media • Technical environments • Pathways • Processes <p>Source: [URD: REG-CN 02, MEET 4] Source: [Pronom6 SRD - S4.1.6.16, Pronom7 SRD – S5.1.1.1]</p>	M
S5.1.5.2	<p>The system must be able to maintain a list of PUID Types that will be used in the assigned PUID values.</p> <ul style="list-style-type: none"> • fmt – File Format • sfw - Software • hdw - Hardware • enc – Character Encoding • cpt – Compression Technique • med – Storage Media • ptw - Pathway • prp – Property • env – Technical Environment <p>Source: [Tessella] Source: [Pronom7 SRD – S5.6.32.1]</p>	M
S5.1.5.3	<p>For every record requiring a PUID the system must store the following information:</p> <ul style="list-style-type: none"> • the puid value • the PUID type <p>Source: [Pronom6 SRD - S4.3.1.15, Pronom7 SRD – S5.4.1]</p>	M
S5.1.5.4	<p>For every PUID the system must ensure that the combination of</p> <ul style="list-style-type: none"> • puid type, • puid value <p>is unique within the current installed instance of the registry. Source: [Pronom6 SRD - S4.3.1.15, Pronom7 SRD – S5.4.2] Source: [MEETING-WIN/08/08/05, MEETING-10-08-2009]</p>	M
S5.1.5.5	Future enhancement, so removed from this document.	X
S5.1.5.6	Future enhancement, so removed from this document.	X
S5.1.5.7	<p>It must be possible for a user to enter a PUID for a new entity (of a type that can be assigned a PUID), rather than having the system generate one, and have the system check that the PUID is unique within the current instance of the registry. If the PUID is not unique, then the user should be prevented from saving the entity with that PUID and informed why.</p> <p>Source: [MEETING-10-08-2009]</p>	M
S5.1.5.8	Future enhancement, so removed from this document.	X

5.2 Core Entities

This section covers the information that is stored for the core entities in the system, which are:

- File Formats
- Software Package
- Hardware
- Character Encoding
- Compression technique
- Storage Media

Core entities suffer from obsolescence and their obsolescence affects our ability to render digital information, which is why information about them is stored in the system.

5.2.1 File Format Requirements

5.2.1.1 File format information

Label	Requirement	Necessity
S5.2.1.1.1	<p>The system must maintain a list of file formats including the following information:</p> <ul style="list-style-type: none"> • System ID (not visible to users). • Name. • Version • Description. • Orientation (linked to the maintained list of possible orientations). • Byte order (linked to the maintained list of possible byte orders: big-endian, little-endian or both). • Level of public disclosure (linked to the maintained list of possible disclosures). • Release date. • Whether withdrawn or not • Withdrawal date (if withdrawn). • Notes/comments on the format. <p>Source: [P3SD – 3.2.7.2, P4IM, Meeting 1, PC3/D4 – 2.18] Source: [Pronom6 SRD - S4.3.1.1, Pronom7 SRD – S5.1.0.1, Pronom7 SRD – S5.1.1.3, Pronom7 SRD – S5.1.2.1.1, Pronom7 SRD – S5.1.2.1.2]</p>	M
S5.2.1.1.2	<p>The system must be able to assign multiple external identifiers to each file format.</p> <p>Source: [P4IM] Source: [Pronom6 SRD - S4.3.1.2, Pronom7 SRD – S5.1.1.9]</p>	M
S5.2.1.1.3	<p>The system must be able to record information on multiple file format developers (people or organisations).</p> <p>Source: [P4IM, PA-tool registry - 5.2.3, Pronom6 SRD - S4.3.1.4] Source: [Pronom6 SRD – S4.3.1.3, Pronom7 SRD – S5.1.1.7]</p>	M
S5.2.1.1.4	<p>The system must be able to record information on multiple agents (people or organisations) providing file format support.</p> <p>Source: [P4IM, PA-tool registry - 5.2.3, Pronom6 SRD - S4.3.1.4] Source: [Pronom6 SRD – S4.3.1.4, Pronom7 SRD – S5.1.1.7]</p>	M
S5.2.1.1.5	<p>The system must be able to record information on multiple documentation items associated with the file format.</p> <p>Source: [P4IM, PA-tool registry - 5.2.1, PA-tool registry - 5.2.2, PA-tool registry - 5.2.9, PA-tool registry - 5.9.3, PC3/D4 – 2.11] Source: [Pronom6 SRD – S4.3.1.5, Pronom7 SRD – S5.1.1.4]</p>	M
S5.2.1.1.6	<p>The system must be able to record information on multiple intellectual property rights associated with the file format.</p> <p>Source: [P4IM, PA-tool registry - 5.2.4, PA-tool registry - 5.9.4] Source: [Pronom6 SRD – S4.3.1.6, Pronom7 SRD – S5.1.1.5]</p>	M
S5.2.1.1.7	<p>The system must be able to record multiple file format aliases (holding the name and version for each alias).</p> <p>Source: [P4IM, Review, PA-tool registry - 5.2.10, PA-tool registry - 5.9.6] Source: [Pronom6 SRD – S4.3.1.7, Pronom7 SRD – S5.1.1.6]</p>	M
S5.2.1.1.8	<p>The system must be able to assign the file format to multiple file format types (linked to maintained list of file format types).</p> <p>Source: [P4IM] Source: [Pronom6 SRD – S4.3.1.8, Pronom7 SRD – S5.1.1.14]</p>	M
S5.2.1.1.9	<p>The system must be able to record multiple file format byte orders (each linked to the maintained list of byte orders).</p> <p>Source: [P4IM] Source: [Pronom6 SRD - S4.3.1.9]</p>	M

Label	Requirement	Necessity
S5.2.1.1.10	The system must be able to record multiple file format compression techniques (each linked to the maintained list of compression techniques). Source: [P4IM] Source: [Pronom6 SRD - S4.3.1.10, Pronom7 SRD – S5.1.2.1.3]	M
S5.2.1.1.11	The system must be able to record multiple file format character encodings (each linked to the maintained list of character encodings). Source: [P4IM] Source: [Pronom6 SRD - S4.3.1.11, Pronom7 SRD – S5.1.2.1.4]	M
S5.2.1.1.12	The system must be able to record information on multiple reference files associated with the file format. Source: [P4IM] Source: [Pronom6 SRD - S4.3.1.12, Pronom7 SRD – S5.1.2.5.1]	M
S5.2.1.1.13	The system must be able to assign the file format to multiple file format families. Source: [P4IM] Source: [Pronom6 SRD - S4.3.1.13, Pronom7 SRD – S5.1.1.15]	M
S5.2.1.1.14	The system must be able to assign a priority relationship between two file formats to illustrate that one format replaced the other. Source: [13-Jan-2005] Source: [Pronom6 SRD - S4.3.1.14, Pronom7 SRD – S5.1.2.1.6] N.B. This is used in file format identification by DROID: e.g., if a format is found to be both a RTF1 and RTF2 file but not RTF3, it should be assigned as RTF2.	M
S5.2.1.1.15	It must be possible to associate file formats with one or more Technical Environments Source: [TWSRD] Source: [Pronom6 SRD - S4.14.2, Pronom7 SRD – S5.1.2.1.5]	M
S5.2.1.1.16	The system must be able to associate multiple XCEL files with each file format. Source: [PC3/D4 - 2.4] Source: [Pronom7 SRD – S5.1.2.1.7a, Pronom7 SRD – S5.1.2.7.1]	M
S5.2.1.1.17	The system must be able to associate multiple XCDL schema files with each file format. Source: [PC3/D4 - 2.4] Source: [Pronom7 SRD – S5.1.2.1.7b, Pronom7 SRD – S5.1.2.6.1]	M
S5.2.1.1.18	The system must be able to associate multiple file format assessment records created by other organisations with each file format. The format assessments should consist of the following information: <ul style="list-style-type: none"> a descriptive summary field a link to the detailed assessment on an external website. Source: [PC3/D4 – 2.10] Source: [Pronom7 SRD – S5.1.2.1.8]	M
S5.2.1.1.19	The system must be able to associate each file format with a component manifestation type. Source: [Pronom 6 - Implemented] Source: [Pronom7 SRD – S5.1.2.1.9]	M
S5.2.1.1.20	The system must be able to associate multiple instance properties with each file format. Source: [Pronom6 SRD - S4.13.2, Pronom7 SRD – S5.1.2.2.1]	M
S5.2.1.1.21	The system must be able to associate multiple GDFR facet values with each file format. Source: [PC3/D4 - 2.21, MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.2.1.10, Pronom7 SRD – S5.1.0.3, Pronom7 SRD – S5.6.22.12]	M

5.2.1.2 [External signatures](#)

Label	Requirement	Necessity
S5.2.1.2.1	The system must maintain a list of external signatures including the following information: <ul style="list-style-type: none"> External signature type (linked to a maintained list of values) External signature text (the value). Source: [P4IM] Source: [Pronom6 SRD - S4.3.2.1, Pronom7 SRD – S5.2.12.1]	M
S5.2.1.2.2	The system must be able to assign multiple external signatures to a file format. Source: [P4IM] Source: [Pronom6 SRD - S4.3.2.2, Pronom7 SRD – S5.1.2.3.1]	M
S5.2.1.2.3	The system must be able to assign an external signature to multiple file formats. Source: [Clarification 2] Source: [Pronom6 SRD - S4.3.2.4, Pronom7 SRD – S5.2.12.2]	M
S5.2.1.2.4	Future enhancement, so removed from this document.	X

5.2.1.3 [Internal signatures](#)

Label	Requirement	Necessity
S5.2.1.3.1	The system must maintain a list of internal signatures including the following information: <ul style="list-style-type: none"> Name Note Source: [P4IM, Review] Source: [Pronom6 SRD - S4.3.3.1, Pronom7 SRD – S5.2.13.1]	M
S5.2.1.3.2	It must be possible to assign multiple byte sequences to each internal signature. Source: [P4IM] Source: [Pronom6 SRD - S4.3.3.2, Pronom7 SRD – S5.2.13.5]	M
S5.2.1.3.3	The system must maintain a list of byte sequences including the following information: <ul style="list-style-type: none"> Position type (linked to the maintained list of byte sequence position types). Value. Endianness. Offset. Maximum offset Indirect offset location (the location within the file where the offset for the byte sequence is found). Indirect offset length (the number of bytes specifying the offset). Source: [P4IM, Byte, Review, PRONOM5a] Source: [Pronom6 SRD - S4.3.3.3, Pronom7 SRD – S5.2.13.1.1, Pronom7 SRD – S5.2.13.1.2, Pronom7 SRD – S5.2.13.1.3]	M
S5.2.1.3.4	The system must be able to assign multiple internal signatures to a file format. Source: [P4IM] Source: [Pronom6 SRD - S4.3.3.4, Pronom7 SRD – S5.1.2.4.1]	M
S5.2.1.3.5	The system must be able to assign an internal signature to multiple file formats. Source: [Clarification 2] Source: [Pronom6 SRD - S4.3.3.6, Pronom7 SRD – S5.2.13.4]	M
S5.2.1.3.6	The system must be able to distinguish between internal signatures assigned to a single file format (which should be called a specific internal signature) and those assigned to multiple file formats (which should be called generic internal signatures). Source: [13-Jan-2005] Source: [Pronom6 SRD - S4.3.3.8, Pronom7 SRD – S5.2.13.2]	M

5.2.1.4 Reference Files

Label	Requirement	Necessity
S5.2.1.4.1	The system must maintain a list of file format reference files (e.g., specification documents) including the following information: <ul style="list-style-type: none"> Name. Version. External Identifier (e.g., a URL) Description. Note. Source: [P4IM] Source: [Pronom6 SRD - S4.3.4.1, Pronom7 SRD – S5.2.14.1, Pronom7 SRD – S5.2.14.2, Pronom7 SRD – S5.2.14.3]	M
S5.2.1.4.2	The system must be able to record information on multiple documentation items associated with the reference file. Source: [P4IM] Source: [Pronom6 SRD - S4.3.4.2, Pronom7 SRD – S5.2.14.4]	M
S5.2.1.4.3	The system must be able to record information on multiple intellectual property rights associated with the reference file. Source: [P4IM] Source: [Pronom6 SRD - S4.3.4.3, Pronom7 SRD – S5.2.14.5]	M

5.2.2 Software Package Requirements

Label	Requirement	Necessity
S5.2.2.1	The system must maintain a list of software packages including the following information: <ul style="list-style-type: none"> Software package name Version Description Service pack level. Release date. Whether it is withdrawn or not. Withdrawal date (if withdrawn). Note Default file formats (linked to the maintained list of file formats) i.e. those file formats that the software package can create. Source: [P3SD – 3.2.7.2, P4IM, Clarification 1, PA-tool registry - 5.2.1, PA-tool registry - 5.2.6, PC3/D4 – 2.17, MEETING-WIN/08/08/05] Source: [Pronom6 SRD - S4.4.1, Pronom7 SRD – S5.1.0.1, Pronom7 SRD – S5.1.1.2a, Pronom7 SRD – S5.1.1.2b, Pronom7 SRD – S5.1.1.3, Pronom7 SRD – S5.1.4.1.1, Pronom7 SRD – S5.1.4.1.2a, Pronom7 SRD – S5.1.4.1.2b, Pronom7 SRD – S7.1.2.5.7.5]	M
S5.2.2.2	The system must be able to assign multiple external references to each software package Source: [P4IM] Source: [Pronom6 SRD - S4.4.2, Pronom7 SRD – S5.1.1.9]	M
S5.2.2.3	The system must be able to record for each software package a list of the people and/or organisations that developed the software each linked to a maintained person or organisation in the system. Source: [P4IM, PA-tool registry - 5.2.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.4.3, Pronom7 SRD – S5.1.1.7]	M
S5.2.2.4	The system must be able to record for each software package a list of the people and/or organisations that support the software each linked to a maintained person or organisation in the system. Source: [P4IM, PA-tool registry - 5.2.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.4.4, Pronom7 SRD – S5.1.1.7]	M

Label	Requirement	Necessity
S5.2.2.5	The system must be able to record for each software package a list of the known documentation about the software each linked to a maintained documentation record in the system. Source: [P4IM, PA-tool registry - 5.2.1, PA-tool registry - 5.2.2, PA-tool registry - 5.2.9, PA-tool registry - 5.9.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.4.5, Pronom7 SRD – S5.1.1.4]	M
S5.2.2.6	The system must be able to record for each software package a list of the known intellectual property rights about the software each linked to a maintained intellectual property rights record in the system. Source: [P4IM, PA-tool registry - 5.2.4, PA-tool registry - 5.9.4] Source: [Pronom6 SRD - S4.4.6, Pronom7 SRD – S5.1.1.5]	M
S5.2.2.7	The system must be able to assign multiple aliases to each software package (holding the name and version for each alias). Source: [P4IM , Review, PA-tool registry - 5.2.10, PA-tool registry - 5.9.6] Source: [Pronom6 SRD - S4.4.7, Pronom7 SRD – S5.1.1.6]	M
S5.2.2.8	The system must be able to record for each software package a list of the software packages (e.g., operating systems) on which it depends, each linked to a maintained software package in the system. Source: [P4IM, Review] Source: [Pronom6 SRD - S4.4.8, Pronom7 SRD – S5.1.1.14, Pronom7 SRD – S5.1.4.1.5]	M
S5.2.2.9	The system must be able to record for each software package a list of the hardware (e.g., RAM, processor) on which it depends, each linked to a maintained item of hardware in the system. Source: [P4IM, Review] Source: [Pronom6 SRD - S4.4.9, Pronom7 SRD – S5.1.4.1.6]	M
S5.2.2.10	The system must maintain a series of software images including: <ul style="list-style-type: none"> • The software package (linked to the maintained list of software packages). • The image itself. • A textual description of the image. Source: [P4IM] Source: [Pronom6 SRD - S4.4.10, Pronom7 SRD – S5.1.4.1.7]	M
S5.2.2.11	The system must be able to assign each software package record to multiple software package families. Source: [P4IM] Source: [Pronom6 SRD - S4.4.11, Pronom7 SRD – S5.1.1.15]	M
S5.2.2.12	The system must be able to assign each software package record to multiple software package types (each linked to the maintained list of software package types). Source: [P4IM] Source: [Pronom6 SRD - S4.4.12, Pronom7 SRD – S5.1.1.14]	M
S5.2.2.13	The system must be able to record for each software package a list of the languages in which the software is supplied each linked to a maintained language in the system. Source: [P4IM] Source: [Pronom6 SRD - S4.4.15, Pronom7 SRD – S5.1.4.1.9]	M
S5.2.2.14	Future enhancement, so removed from this document.	X
S5.2.2.15	Until the storage media functionality is added, the system must be able to store information on the storage media on which it is supplied in the same way as in PRONOM 3. Source: [Review, TWSRD] Source: [Pronom6 SRD - S4.4.16, Pronom7 SRD – S5.1.4.1.8]	M

Label	Requirement	Necessity
S5.2.2.16	The system must be able to associate multiple GDFR facet values with each software package. Source: [MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.0.3, Pronom7 SRD – S5.6.22.14]	M
S5.2.2.17	The system must be able to record for each software package the availability of the source code for the package. Source: [PA-tool registry - 5.2.1, PA-tool registry - 5.2.6, PC3/D4 – 2.17, MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.4.1.2c] Note: Any restrictions on what can be done with the source code once it has been obtained should be recorded in the licence details for the software package.	M
S5.2.2.18	The system must be able to record for each software package the availability of the software package itself. Source: [Tessella] Source: [Pronom7 SRD – S5.1.4.1.2d] Note: Any restrictions on the use of the software package should be recorded in its licence details.	M
S5.2.2.19	The system must be able to associate multiple technical environments with each software package. Source: [Pronom7 SRD – S5.1.4.1.11] Note: This association can be used to specify a dependency on a number of items of Hardware and/or Software Packages.	M
S5.2.2.20	The system must be able to associate multiple programming languages with each software package. Source: [PA-tool registry - 5.2.1, PA-tool registry - 5.2.6, PC3/D4 – 2.17] Source: [Pronom7 SRD – S5.1.4.1.13]	M
S5.2.2.21	The system must be able to associate multiple software interface types (such as GUI, command line, etc.) with each software package. Source: [PA-tool registry - 5.2.11] Source: [Pronom7 SRD – S5.1.4.1.14]	M
S5.2.2.22	They system must be able to associate multiple locations where the software package can be obtained or downloaded with each software package. Source: [PA-tool registry - 5.2.1] Source: [Pronom7 SRD – S5.1.4.1.15, Pronom7 SRD – S5.1.4.1.16]	M

5.2.2.1 [Software Components](#)

Label	Requirement	Necessity
S5.2.2.1.1	The system must be able to store the name, version and description of components within software, such as those within JHOVE. Source: [TWSRD] Source: [Pronom6 SRD - S4.4.17, Pronom7 SRD – S5.1.4.1.10]	M
S5.2.2.1.2	Future enhancement, so removed from this document.	X

5.2.2.2 [Software Package Tools](#)

Label	Requirement	Necessity
S5.2.2.2.1	The system must maintain a list of software package tools to allow the name, description and the Java class name of the SDB wrapper for the tool to be stored. Source: [TWSRD] Source: [Pronom6 SRD - S4.12.1, Pronom7 SRD – S5.1.4.2.1, Pronom7 SRD – S5.1.4.2.3]	M

Label	Requirement	Necessity
S5.2.2.2.2	The system must be able to specify that a software package tool can identify, validate, migrate, extract properties or objects, or redact a given file format. Source: [TWSRD] Source: [Pronom6 SRD - S4.12.2]	M
S5.2.2.2.3	The system must be able to associate multiple software package tools (each of which is a wrapper used to invoke the software package to perform a specific task) with each software package. Source: [TWSRD] Source: [Pronom6 SRD - S4.12.3, Pronom7 SRD – S5.1.4.1.12, Pronom7 SRD – S5.1.4.2.2]	M
S5.2.2.2.4	For each software package tool, the system must be able to record whether the tool is capable of batch processing. Source: [PA-tool registry - 5.2.1] Source: [Pronom7 SRD – S5.1.4.1.3]	M
S5.2.2.2.5	For each software package tool, the system must be able to record whether the software package tool is accessible as a Planets service or not. Source: [PA-tool registry - 5.2.1] Source: [Pronom7 SRD – S5.1.4.1.4]	M
S5.2.2.2.6	The system must be able to maintain a list of software package tool invocation details including the following information: <ul style="list-style-type: none"> • an optional parameter string used at invocation time • a description of what the invocation of the tool with the given parameter string does Source: [PA-tool registry - 5.2.9, PA-tool registry - 5.8.1, Tessella] Source: [Pronom7 SRD – S5.1.4.3.1, Pronom7 SRD – S5.1.4.3.3, Pronom7 SRD – S5.1.4.3.4]	M
S5.2.2.2.7	The system must be able to associate multiple sets of invocation details with each software package tool. Source: [PA-tool registry - 5.2.9] Source: [Pronom7 SRD – S5.1.4.3.2]	M
S5.2.2.2.8	For each property extraction software package tool the system must be able to record which properties that tool can extract from a given file format. Source: [Pronom7 implementation]	M
S5.2.2.2.9	Future enhancement, so removed from this document.	X

5.2.2.3 [Software Package Licences](#)

Label	Requirement	Necessity
S5.2.2.3.1	For each software package, the system must be able to record an association with multiple licences. Source: [PC3/D4 – 2.17] Source: [Pronom7 SRD – S5.1.4.4.1]	M
S5.2.2.3.2	For each association between a software package and a licence, the system must be able to record additional licence details as free text. Source: [PC3/D4 – 2.17] Source: [Pronom7 SRD – S5.1.4.4.2]	M

5.2.3 [Hardware Requirements](#)

Label	Requirement	Necessity
-------	-------------	-----------

Label	Requirement	Necessity
S5.2.3.1	<p>The system must maintain a list of hardware including the following information:</p> <ul style="list-style-type: none"> • Hardware name • Version • Description • Release date. • Whether it is withdrawn or not. • Withdrawal date (if withdrawn). • Note <p>Source: [P4IM, PA-tool registry - 5.9.1] Source: [Pronom6 SRD - S4.6.1, Pronom7 SRD – S5.1.0.1, Pronom7 SRD – S5.1.1.2a, Pronom7 SRD – S5.1.1.2b, Pronom7 SRD – S5.1.1.3, Pronom7 SRD – S5.1.3.1, Pronom7 SRD – S5.1.3.4]</p>	M
S5.2.3.2	<p>The system must be able to assign multiple external references to each item of hardware.</p> <p>Source: [P4IM] Source: [Pronom6 SRD - S4.6.2, Pronom7 SRD – S5.1.1.9]</p>	M
S5.2.3.3	<p>The system must be able to record for each item of hardware a list of the people and/or organisations that developed the hardware each linked to a maintained person or organisation in the system.</p> <p>Source: [P4IM, PA-tool registry - 5.2.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.6.3, Pronom7 SRD – S5.1.1.7]</p>	M
S5.2.3.4	<p>The system must be able to record for each item of hardware a list of the people and/or organisations that support the hardware each linked to a maintained person or organisation in the system.</p> <p>Source: [P4IM, PA-tool registry - 5.2.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.6.4, Pronom7 SRD – S5.1.1.7]</p>	M
S5.2.3.5	<p>The system must be able to record for each item of hardware a list of the known documentation about the hardware each linked to a maintained documentation record in the system.</p> <p>Source: [P4IM, PA-tool registry - 5.2.1, PA-tool registry - 5.2.2, PA-tool registry - 5.2.9, PA-tool registry - 5.9.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.6.5, Pronom7 SRD – S5.1.1.4]</p>	M
S5.2.3.6	<p>The system must be able to record for each item of hardware a list of the known intellectual property rights about the hardware each linked to a maintained intellectual property rights record in the system.</p> <p>Source: [P4IM, PA-tool registry - 5.2.4, PA-tool registry - 5.9.4] Source: [Pronom6 SRD - S4.6.6, Pronom7 SRD – S5.1.1.5]</p>	M
S5.2.3.7	<p>The system must be able to assign multiple aliases to each item of hardware (holding the name and version for each alias).</p> <p>Source: [P4IM, Review, PA-tool registry - 5.2.10, PA-tool registry - 5.9.6] Source: [Pronom6 SRD - S4.6.7, Pronom7 SRD – S5.1.1.6]</p>	M
S5.2.3.8	<p>The system must be able to record for each item of hardware a list of the software packages (e.g., operating systems) on which it depends, each linked to a maintained software package in the system.</p> <p>Source: [P4IM] Source: [Pronom6 SRD - S4.6.8, Pronom7 SRD – S5.1.3.2]</p>	M
S5.2.3.9	<p>The system must be able to record for each item of hardware a list of the hardware (e.g., RAM, processor) on which it depends, each linked to a maintained item of hardware in the system.</p> <p>Source: [P4IM] Source: [Pronom6 SRD - S4.6.9, Pronom7 SRD – S5.1.1.14, Pronom7 SRD – S5.1.3.3]</p>	M

Label	Requirement	Necessity
S5.2.3.10	The system must maintain a series of hardware images including: <ul style="list-style-type: none"> The item of hardware (linked to the maintained list of hardware). The image itself. A textual description of the image. Source: [P4IM] Source: [Pronom6 SRD - S4.6.10, Pronom7 SRD – S5.1.3.5]	M
S5.2.3.11	The system must be able to assign each hardware record to multiple hardware families (each linked to the maintained list of hardware families). Source: [P4IM] Source: [Pronom6 SRD - S4.6.11, Pronom7 SRD – S5.1.1.15]	M
S5.2.3.12	The system must be able to assign each hardware record to multiple hardware types (each linked to the maintained list of hardware types). Source: [P4IM] Source: [Pronom6 SRD - S4.6.12, Pronom7 SRD – S5.1.1.14]	M
S5.2.3.13	The system must be able to associate multiple GDFR facet values with each item of hardware. Source: [MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.0.3, Pronom7 SRD – S5.6.22.15]	M

5.2.4 Character Encoding Requirements

Label	Requirement	Necessity
S5.2.4.1	The system must maintain a list of character encodings including the following information: <ul style="list-style-type: none"> System ID (not visible to users). Name. Version Description. Code page. Code unit width. Encoding form width. Release date. Whether it is withdrawn or not. Withdrawal date (if withdrawn). Note. Source: [P4IM] Source: [Pronom6 SRD - S4.8.1, Pronom7 SRD – S5.1.0.1, Pronom7 SRD – S5.1.1.2a, Pronom7 SRD – S5.1.1.2b, Pronom7 SRD – S5.1.1.3, Pronom7 SRD – S5.1.5.1, Pronom7 SRD – S5.1.5.2]	M
S5.2.4.2	The system must be able to assign multiple external identifiers to each character encoding. Source: [P4IM] Source: [Pronom6 SRD - S4.8.2, Pronom7 SRD – S5.1.1.9]	M
S5.2.4.3	The system must be able to record information on multiple character encoding developers (people or organisations). Source: [P4IM, PA-tool registry - 5.2.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.8.3, Pronom7 SRD – S5.1.1.7]	M
S5.2.4.4	The system must be able to record information on multiple agents (people or organisations) providing character encoding support. Source: [P4IM, PA-tool registry - 5.2.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.8.4, Pronom7 SRD – S5.1.1.7]	M

Label	Requirement	Necessity
S5.2.4.5	The system must be able to record information on multiple documentation items associated with the character encoding. Source: [P4IM, PA-tool registry - 5.2.1, PA-tool registry - 5.2.2, PA-tool registry - 5.2.9, PA-tool registry - 5.9.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.8.5, Pronom7 SRD – S5.1.1.4]	M
S5.2.4.6	The system must be able to record information on multiple intellectual property rights associated with the character encoding. Source: [P4IM, PA-tool registry - 5.2.4, PA-tool registry - 5.9.4] Source: [Pronom6 SRD - S4.8.6, Pronom7 SRD – S5.1.1.5]	M
S5.2.4.7	The system must be able to record multiple character encoding aliases (holding the name and version for each alias). Source: [P4IM, Review, PA-tool registry - 5.2.10, PA-tool registry - 5.9.6] Source: [Pronom6 SRD - S4.8.7, Pronom7 SRD – S5.1.1.6]	M
S5.2.4.8	The system must be able to assign the character encoding to multiple character encoding families (linked to maintained list of character encoding families). Source: [P4IM] Source: [Pronom6 SRD - S4.8.8, Pronom7 SRD – S5.1.1.15]	M
S5.2.4.9	The system must be able to associate multiple GDFR facet values with each character encoding. Source: [MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.0.3]	M

5.2.5 Compression Technique Requirements

Label	Requirement	Necessity
S5.2.5.1	The system must maintain a list of compression techniques including the following information: <ul style="list-style-type: none"> • System ID. • Name. • Version • Description. • Lossiness (linked to maintained list of lossiness values). • Four character code. • Release date. • Whether it is withdrawn or not. • Withdrawal date (if withdrawn). • Note. Source: [P4IM] Source: [Pronom6 SRD - S4.9.1, Pronom7 SRD – S5.1.0.1, Pronom7 SRD – S5.1.1.2a, Pronom7 SRD – S5.1.1.2b, Pronom7 SRD – S5.1.1.3, Pronom7 SRD – S5.1.6.1, Pronom7 SRD – S5.1.6.2]	M
S5.2.5.2	The system must be able to assign multiple external identifiers to each compression technique. Source: [P4IM] Source: [Pronom6 SRD - S4.9.2, Pronom7 SRD – S5.1.1.9]	M
S5.2.5.3	The system must be able to record information on multiple compression technique developers (people or organisations). Source: [P4IM, PA-tool registry - 5.2.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.9.3, Pronom7 SRD – S5.1.1.7]	M
S5.2.5.4	The system must be able to record information on multiple agents (people or organisations) providing compression technique support. Source: [P4IM, PA-tool registry - 5.2.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.9.4, Pronom7 SRD – S5.1.1.7]	M

Label	Requirement	Necessity
S5.2.5.5	The system must be able to record information on multiple documentation items associated with the compression technique. Source: [P4IM, PA-tool registry - 5.2.1, PA-tool registry - 5.2.2, PA-tool registry - 5.2.9, PA-tool registry - 5.9.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.9.5, Pronom7 SRD – S5.1.1.4]	M
S5.2.5.6	The system must be able to record information on multiple intellectual property rights associated with the compression technique. Source: [P4IM, PA-tool registry - 5.2.4, PA-tool registry - 5.9.4] Source: [Pronom6 SRD - S4.9.6, Pronom7 SRD – S5.1.1.5]	M
S5.2.5.7	The system must be able to record multiple compression technique aliases (holding the name and version for each alias). Source: [P4IM, Review, PA-tool registry - 5.2.10, PA-tool registry - 5.9.6] Source: [Pronom6 SRD - S4.9.7, Pronom7 SRD – S5.1.1.6]	M
S5.2.5.8	The system must be able to assign the compression technique to multiple compression technique families (linked to maintained list of compression technique families). Source: [P4IM] Source: [Pronom6 SRD - S4.9.8, Pronom7 SRD – S5.1.1.15]	M
S5.2.5.9	The system must be able to associate multiple GDFR facet values with each compression technique. Source: [MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.0.3, Pronom7 SRD – S5.6.22.13]	M

5.2.6 Storage Medium Requirements

Label	Requirement	Necessity
S5.2.6.1	The system must maintain a list of storage media components including the following information: <ul style="list-style-type: none"> • Storage Media Name • Version • Description • Release date. • Whether it is withdrawn or not. • Withdrawal date (if known). • Note. Source: [P4IM] Source: [Pronom6 SRD - S4.7.1, Pronom7 SRD – S5.1.0.1, Pronom7 SRD – S5.1.1.2a, Pronom7 SRD – S5.1.1.2b, Pronom7 SRD – S5.1.1.3, Pronom7 SRD – S5.1.7.1, Pronom7 SRD – S5.1.7.2]	M
S5.2.6.1.1	Future enhancement, so removed from this document.	X
S5.2.6.2	Future enhancement, so removed from this document.	X
S5.2.6.3	The system must be able to record for each storage media record a list of the people and/or organisations that developed the media each linked to a maintained person or organisation in the system. Source: [P4IM, PA-tool registry - 5.2.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.7.3, Pronom7 SRD – S5.1.1.7]	M
S5.2.6.4	The system must be able to record for each storage media record a list of the people and/or organisations that support the media each linked to a maintained person or organisation in the system. Source: [P4IM, PA-tool registry - 5.2.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.7.4, Pronom7 SRD – S5.1.1.7]	M
S5.2.6.5	The system must be able to record for each storage media record a list of the known documentation about the media each linked to a maintained documentation record in the system. Source: [P4IM, PA-tool registry - 5.2.1, PA-tool registry - 5.2.2, PA-tool registry - 5.2.9, PA-tool registry - 5.9.3, PC3/D4 – 2.11] Source: [Pronom6 SRD - S4.7.5, Pronom7 SRD – S5.1.1.4]	M

Label	Requirement	Necessity
S5.2.6.6	Future enhancement, so removed from this document.	X
S5.2.6.7	Future enhancement, so removed from this document.	X
S5.2.6.8	The system must be able to record for each storage media record a list of the known intellectual property rights about the media each linked to a maintained intellectual property rights record in the system. Source: [P4IM, PA-tool registry - 5.2.4, PA-tool registry - 5.9.4] Source: [Pronom6 SRD - S4.7.8, Pronom7 SRD – S5.1.1.5]	M
S5.2.6.9	The system must be able to assign multiple aliases for each storage media (holding the name and version for each alias). Source: [P4IM, Review, PA-tool registry - 5.2.10, PA-tool registry - 5.9.6] Source: [Pronom6 SRD - S4.7.8, Pronom7 SRD – S5.1.1.6]	M
S5.2.6.10	Future enhancement, so removed from this document.	X
S5.2.6.11	Future enhancement, so removed from this document.	X
S5.2.6.12	The system must maintain a series of storage media images including: <ul style="list-style-type: none"> • The storage medium (linked to the maintained list of storage media). • The image itself. • A textual description of the image. Source: [P4IM] Source: [Pronom6 SRD - S4.7.11, Pronom7 SRD – S5.1.7.5]	M
S5.2.6.13	The system must be able to assign each storage media record to multiple storage media families (each linked to the maintained list of storage media families). Source: [P4IM] Source: [Pronom6 SRD - S4.7.12, Pronom7 SRD – S5.1.1.15]	M
S5.2.6.14	The system must be able to assign each storage media record to multiple storage media types (each linked to the maintained list of storage media types). Source: [P4IM] Source: [Pronom6 SRD - S4.7.13, Pronom7 SRD – S5.1.1.14]	M
S5.2.6.15	The system must be able to associate multiple GDFR facet values with each storage medium. Source: [MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.0.3 Pronom7 SRD – S5.6.22.16]	M

5.3 Subsidiary Entity Requirements

The system needs to store a variety of information about people, organisations, documentation, intellectual property rights etc. The need to record information on these entities is to support the core entities that the system is really about. This section covers the information that needs to be held on each of these subsidiary entities.

5.3.1 People and Organisations (Agents)

Label	Requirement	Necessity
-------	-------------	-----------

Label	Requirement	Necessity
S5.3.1.1	<p>The system must be able to maintain a list of relevant people including the following information:</p> <ul style="list-style-type: none"> • System ID • Name. • Job title. • Organisation to which they belong. • Address (single field) (optional) • Personal Telephone (optional). • Personal E-mail address (optional). • Personal Web site URL (optional) • History (optional). • Description (optional notes). <p>Source: [P4IM, TWSRD, PA-tool registry - 5.1.1] Source: [Pronom6 SRD - S4.2.1.1, Pronom7 SRD – S5.1.0.1, Pronom7 SRD – S5.2.1.2, Pronom7 SRD – S5.2.1.9]</p>	M
S5.3.1.2	<p>The system must be able to maintain a list of relevant organisations including the following information:</p> <ul style="list-style-type: none"> • System ID • Organisation Name. • Address (single field) (optional). • Telephone number (optional). • Contact e-mail address (optional). • Web site URL (optional) • History (optional) • Description (optional notes). <p>Source: [P3SD – 3.2.7.2, P4IM, Clarification 2, TWSRD, PA-tool registry - 5.1.1, PA-tool registry - 5.1.2] Source: [Pronom6 SRD - S4.2.1.2, Pronom7 SRD – S5.1.0.1, Pronom7 SRD – S5.2.1.2, Pronom7 SRD – S5.2.1.10]</p>	M
S5.3.1.3	<p>The system must be able to record the role a person or organisation plays in any relationship it has with a core or subsidiary entity.</p> <p>Source: [Pronom7 SRD – S5.1.1.8]</p>	M
S5.3.1.4	<p>The system must be able to associate multiple GDFR facet values with each person or organisation.</p> <p>Source: [MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.0.3]</p>	M
S5.3.1.5	<p>The system must be able to record the country where a person's or organisation's main residence is located.</p> <p>Source: [Pronom6 - Implemented] Source: [Pronom7 SRD – S5.2.1.3]</p>	M
S5.3.1.6	<p>The system must be able to record whether an agent is an individual or an organisation and what the type of that organisation is.</p> <p>Source: [Pronom6 – Implemented, PA-tool registry - 5.1.2] Source: [Pronom7 SRD – S5.2.1.4 Pronom7 SRD – S5.2.1.10]</p>	M
S5.3.1.7	<p>The system must be able to associate multiple web sites (each with a URL and a description) with each agent (person or organisation).</p> <p>Source: [Pronom6 – Implemented, PA-tool registry - 5.1.1, PA-tool registry - 5.1.2] Source: [Pronom7 SRD – S5.2.1.5, Pronom7 SRD – S5.2.1.6]</p>	M
S5.3.1.8	<p>The system must be able to associate multiple events (each with a date and a description) with each agent (person or organisation).</p> <p>Source: [Pronom6 - Implemented] Source: [Pronom7 SRD – S5.2.1.7, Pronom7 SRD – S5.2.1.8]</p>	M
S5.3.1.9	<p>The system must be able to associate each person with multiple organisations.</p> <p>Source: [PA-tool registry - 5.1.4] Source: [Pronom7 SRD – S5.2.1.11]</p>	M

5.3.2 Documentation

Label	Requirement	Necessity
S5.3.2.1	<p>The system must be able to hold the following information about documentation:</p> <ul style="list-style-type: none"> • System ID (not displayed to users) • Title • Full bibliographical reference for citation display purposes. • Documentation publication date. • Documentation type (linked to maintained list of documentation types). • Documentation availability (linked to maintained list of documentation availability types, such as 'public', 'restricted', 'not available'). • Documentation availability notes. • Documentation Content Type (linked to maintained list of documentation content types such as 'user manual', and 'installation guide'). <p>Source: [P4IM, PA-tool registry - 5.5.1, PA-tool registry - 5.5.3, PA-tool registry - 5.5.7] Source: [Pronom6 SRD - S4.2.2.1, Pronom7 SRD – S5.1.0.1, Pronom7 SRD – S5.2.3.1]</p>	M
S5.3.2.2	<p>The system must be able to assign multiple external references to each documentation item.</p> <p>Source: [P4IM, PA-tool registry - 5.5.6] Source: [Pronom6 SRD - S4.2.2.2, Pronom7 SRD – S5.2.3.2]</p>	M
S5.3.2.3	<p>The system must be able to assign multiple author (a person or organisation) references to each documentation item.</p> <p>Source: [P4IM, PA-tool registry - 5.5.2] Source: [Pronom6 SRD - S4.2.2.3, Pronom7 SRD – S5.2.3.3]</p>	M
S5.3.2.4	<p>The system must be able to assign multiple publisher (a person or organisation) references to each documentation item.</p> <p>Source: [P4IM] Source: [Pronom6 SRD - S4.2.2.4, Pronom7 SRD – S5.2.3.4]</p>	M
S5.3.2.5	<p>The system must be able to assign multiple intellectual property rights to each documentation item.</p> <p>Source: [P4IM, PA-tool registry - 5.5.4] Source: [Pronom6 SRD - S4.2.2.5, Pronom7 SRD – S5.2.3.5]</p>	M
S5.3.2.6	<p>The system must be able to associate multiple GDFR facet values with each item of documentation.</p> <p>Source: [MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.0.3]</p>	M

5.3.3 Intellectual Property Rights

Label	Requirement	Necessity
S5.3.3.1	<p>The system must be able to hold the following information about intellectual property rights:</p> <ul style="list-style-type: none"> • System ID (not displayed to users) • Intellectual property right type (linked to maintained list of intellectual property right types). • Date. • Jurisdiction (linked to maintained list of jurisdictions) • IPR details <p>Source: [P4IM, PA-tool registry - 5.4.1] Source: [Pronom6 SRD - S4.2.3.1, Pronom7 SRD – S5.1.0.1, Pronom7 SRD – S5.2.2.1]</p>	M
S5.3.3.2	<p>The system must be able to assign multiple external identifies to each intellectual property right.</p> <p>Source: [P4IM] Source: [Pronom6 SRD - S4.2.3.2, Pronom7 SRD – S5.2.2.2]</p>	M

Label	Requirement	Necessity
S5.3.3.3	The system must be able to assign multiple owners to each intellectual property right. Source: [P4IM, PA-tool registry - 5.4.2] Source: [Pronom6 SRD - S4.2.3.3, Pronom7 SRD – S5.2.2.3]	M
S5.3.3.4	The system must be able to associate multiple GDFR facet values with each intellectual property right. Source: [MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.0.3]	M

5.3.4 Technical Environments

Label	Requirement	Necessity
S5.3.4.1	It must be possible to describe a standard technical environment, including the following information: <ul style="list-style-type: none"> Name Description Source: [TWSRD] Source: [Pronom6 SRD - S4.14.1, Pronom7 SRD – S5.2.4.1a, Pronom7 SRD – S5.1.0.1] N.B. This is essentially a brief description of a typical combination of software or hardware (e.g. “a browser with a PDF-compatible viewer add-in” or “An appropriate version of Microsoft Office on Windows on a PC”).	M
S5.3.4.2	For each technical environment, the system must be able to record a PUID. Source: [MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.2.4.1b]	M
S5.3.4.3	The system must be able to associate multiple items of hardware with each technical environment. Source: [PA-tool registry - 5.3.3] Source: [Pronom7 SRD – S5.2.4.2]	M
S5.3.4.4	The system must be able to associate multiple software packages with each technical environment. Source: [PA-tool registry - 5.3.4] Source: [Pronom7 SRD – S5.2.4.3]	M
S5.3.4.5	The system must be able to associate multiple documents with each technical environment. Source: [PA-tool registry - 5.3.5] Source: [Pronom7 SRD – S5.2.4.4]	M
S5.3.4.6	The system must be able to associate multiple GDFR facet values with each technical environment. Source: [MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.0.3]	M

5.3.5 Processes

A (software) process is an action that can be performed by a software package. Processes can be classified as either ‘processes for objects’ or ‘processes for technical environments’, where ‘processes for objects’ includes both ‘processes for components’ and ‘processes for file formats’.

Label	Requirement	Necessity
-------	-------------	-----------

Label	Requirement	Necessity
S5.3.5.1	<p>The system must maintain a list of the processes by which software packages interact with file formats including the following information:</p> <ul style="list-style-type: none"> • The file format used. • The software package used. • Software process type (linked to a list of authority controlled software process types). • A descriptive note • Priority of the process with respect to other processes of the same process type. • A description of known limitations of the specific software process. <p>Source: [P4IM, Clarification 2, Review, PC3/D4 – 2.12] Source: [Pronom6 SRD - S4.5.1, Pronom7 SRD – S5.2.5.1.1, Pronom7 SRD – S5.2.5.1.2, Pronom7 SRD – S5.2.5.1.3, Pronom7 SRD – S5.2.5.1.4]</p>	M
S5.3.5.2	<p>The system must be able to record for each software process a list of the known documentation about the process each linked to a maintained documentation record in the system.</p> <p>Source: [P4IM] Source: [Pronom6 SRD - S4.5.3, Pronom7 SRD – S5.2.5.1.6]</p>	M
S5.3.5.3	<p>The system must be able to associate each software process with details of how the software package tool can be invoked.</p> <p>Source: [Tessella] Source: [Pronom7 SRD – S5.2.5.1.5]</p>	M
S5.3.5.4	<p>The system must record that the following software process types (which are grouped as 'processes for objects') are associated with the 'action on file format' process action type:</p> <ul style="list-style-type: none"> • Format identification • Format validation • Characterisation property extraction • Embedded object extraction • Redaction • Migration <p>Source: [Pronom7 SRD – S5.6.30.2]</p>	M
S5.3.5.5	<p>The system must record that the following software process types (which are grouped as 'processes for technical environments') are associated with the 'action on technical environment' process action type:</p> <ul style="list-style-type: none"> • Emulation <p>Source: [Pronom7 SRD – S5.6.30.2]</p>	M
S5.3.5.6	<p>When recording a render software process it must be possible to record a note on the content invariance using a maintained list of allowed invariance descriptions.</p> <p>Source: [P4IM, Clarification 1] Source: [Pronom6 SRD - S4.5.2, Pronom7 SRD – S5.2.5.2.3]</p>	M
S5.3.5.7	Future enhancement, so removed from this document.	X
S5.3.5.8	<p>When recording an extract metadata process, it must be possible to describe multiple metadata elements that can be extracted.</p> <p>Source: [P4IM, Clarification 2] Source: [Pronom6 SRD - S4.5.5, Pronom7 SRD – S5.2.5.2.4]</p>	M
S5.3.5.9	<p>When recording a component manifestation identification process, it must be possible to associate the process with the component manifestation type it can identify.</p> <p>Source: [Tessella] Source: [Pronom7 SRD – S5.2.5.2.5]</p>	M

Label	Requirement	Necessity
S5.3.5.10	When recording a component manifestation measurement process, it must be possible to associate the process with the component manifestation type it can measure. Source: [Tessella] Source: [Pronom7 SRD – S5.2.5.2.6]	M
S5.3.5.11	When recording an emulation process, it must be possible to associate the process with the technical environment being emulated. Source: [Tessella] Source: [Pronom7 SRD – S5.2.5.3.2]	M
S5.3.5.12	The system must be able to record the priority of a software process with respect to other processes that can perform the same task. Source: [Tessella] Source: [Pronom7 SRD – S5.3.3.1.1]	M

5.3.6 Pathways

This section covers the information held about Pathways, which applies to both File Format Pathways and Technical Environment Pathways.

Label	Requirement	Necessity
S5.3.6.1	The system must maintain a list of pathways (a defined sequence of steps to achieve a desired outcome, where each step may use a software package tool to carry out a process) including the following information: <ul style="list-style-type: none"> • System ID (not displayed to users) • Name • Description • Type of pathway action (file format migration, file format object extraction, technical environment emulation) Source: [PA-tool registry – 5.7.1] Source: [Pronom6 SRD - S4.11.1, Pronom7 SRD – S5.1.0.1, Pronom7 SRD – S5.2.6.1.4] Note: this currently does not allow for a pathway to be a mixture of actions on environments and actions on objects.	M
S5.3.6.2	The system must be able to associate multiple GDFR facet values with each pathway. Source: [MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.0.3]	M
S5.3.6.3	For each pathway, the system must be able to assign the role that the pathway can be used for (e.g. extraction, presentation, preservation). Source: [Pronom6 – current implementation] Source: [Pronom7 SRD – S5.2.6.2.1]	M
S5.3.6.4	For each role that a pathway can be used for, the system must be able to record the status of the Pathway (unknown/current/withdrawn). Source: [Pronom6 – current implementation] Source: [Pronom7 SRD – S5.2.6.2.2]	M
S5.3.6.5	The system must allow a pathway to be approved for a particular role (e.g. extraction, presentation, preservation). Source: [URD: REG-CN 03, MEET 4] Source: [Pronom6 SRD - S4.11.2, Pronom7 SRD – S5.2.6.2.3]	M
S5.3.6.6	For each approval of a pathway for a particular role, the system must store the following information: <ul style="list-style-type: none"> • Date that the pathway was approved for the particular role • Agent who approved the pathway for the particular role Source: [Pronom6 – current implementation] Source: [Pronom7 SRD – S5.3.2.1.3, Pronom7 SRD – S5.3.2.1.4]	M

Label	Requirement	Necessity
S5.3.6.7	<p>The system must be able to associate each pathway with multiple steps.</p> <p>Source: [PC3/D4 – 5.7.2, PA-tool registry - 5.8.1] Source: [Pronom7 SRD – S5.2.6.1.3]</p>	M
S5.3.6.8	<p>For each step in a pathway, the system must be able to record the following information:</p> <ul style="list-style-type: none"> • a sequence number (to indicate the position of the step in the ordered sequence of steps that makes up the pathway) • the software package tool used in the step (optional) • the invocation details of the software package tool used in the step (optional) <p>Source: [TWSRD, PA-tool registry - 5.8.1] Source: [Pronom6 SRD - S4.12.4, Pronom7 SRD – S5.2.6.3.1, Pronom7 SRD – S5.2.6.3.2, Pronom7 SRD – S5.2.6.3.3, Pronom7 SRD – S5.2.6.4.1] Note: Each step can participate in one pathway. If the same conceptual step is present in >1 pathway, it will need to be entered as >1 physical step to ensure that the sequence numbers make sense.</p>	M
S5.3.6.9	<p>Where a step is for a technical environment emulation pathway, the system must be able to record the following information:</p> <ul style="list-style-type: none"> • PUID of the technical environment the emulation process runs in (source technical environment) • PUID of the technical environment being emulated (target technical environment) <p>Source: [PC3/D4 – 5.7.5] Source: [Pronom7 SRD – S5.2.6.6.1]</p>	M
S5.3.6.10	<p>Where a step is for a file format pathway, the system must be able to record the following information:</p> <ul style="list-style-type: none"> • Maximum size of input file • PUID of the input file <p>Source: [PC3/D4 – 5.7.1, PC3/D4 – 5.7.4] Source: [Pronom7 SRD – S5.2.6.5.1.1, Pronom7 SRD – S5.2.6.5.1.2]</p>	M
S5.3.6.11	<p>Where a step is for a file format <i>migration</i> pathway, the system must be able to record the following information:</p> <ul style="list-style-type: none"> • PUID of the output file <p>Source: [PC3/D4 – 5.7.5] Source: [Pronom7 SRD – S5.2.6.5.2.1]</p>	M
S5.3.6.12	<p>The system must be able to associate multiple migration pathway object properties with each step for a file format <i>migration</i> pathway, and store the following information for each property:</p> <ul style="list-style-type: none"> • Associated instance property of the source component • Associated instance property of the target component • Variance of the property (a number) • Notes <p>Source: [Tessella] Source: [Pronom7 SRD – S5.2.6.5.2.2, Pronom7 SRD – S5.2.6.5.2.3, Pronom7 SRD – S5.2.6.5.2.4]</p>	M
S5.3.6.13	<p>The system must allow a component property (migration pathway object property) to be approved for the migration pathway role.</p> <p>Source: [Tessella] Source: [Pronom7 SRD – S5.3.2.2.1, Pronom7 SRD – S5.3.2.2.2, S5.3.2.2.3]</p>	M

Label	Requirement	Necessity
S5.3.6.14	For each approval of a component property (migration pathway object property), the system must store the following information: <ul style="list-style-type: none"> • Date that the component property (migration pathway object property) was approved • Agent who approved the component property (migration pathway object property) Source: [Pronom6 – current implementation] Source: [Pronom7 SRD – S5.3.2.2.4, Pronom7 SRD – S5.3.2.2.5]	M
S5.3.6.15	Future enhancement, so removed from this document.	X

5.3.7 Properties for File Formats and Components

The properties of file formats and components can be associated with risk scores which are indicative of how difficult it is to preserve a file or component with a particular property value (e.g. it is harder to preserve a file whose format's compression property has a value of lossy, than one whose format's compression property has a value of lossless). These risk scores can then be combined for all inherent properties of a particular file format to provide an overall risk score for that format.

Label	Requirement	Necessity
S5.3.7.1	The system must maintain a list of properties including the following information: <ul style="list-style-type: none"> • Name • Description • PUID Source: [PA-tool registry – 5.7.1] Source: [Pronom7 SRD – S5.1.0.1, Pronom7 SRD – S5.2.7.1.1a, Pronom7 SRD – S5.2.7.1.1b, Pronom7 SRD – S5.2.7.1.2]	M
S5.3.7.2	The system must be able to hold information about <i>inherent</i> file format properties, including the risk associated with this property. The risk may be dependent on the value of the property. Source: [URD: REG-CN 03, MEET4] Source: [Pronom6 SRD - S4.13.1, Pronom7 SRD – S5.2.7.1.3, Pronom7 SRD – S5.3.1.1.1, Pronom7 SRD – S5.3.1.1.2]	M
S5.3.7.3	The system must be able to hold information about file format <i>instance</i> properties, including the risk associated with this property. Source: [URD: REG-CN 03, MEET4] Source: [Pronom6 SRD - S4.13.2, Pronom7 SRD – S5.1.2.2.1, Pronom7 SRD – S5.2.7.1.3, Pronom7 SRD – S5.3.1.1.1, Pronom7 SRD – S5.3.1.1.2]	M
S5.3.7.4	The system must be able to hold information about component instance properties including maximum and minimum thresholds. Source: [URD: REG-CN 03, MEET4] Source: [Pronom6 SRD - S4.13.3]	M
S5.3.7.5	The system must be able to associate each component type with multiple component instance properties. Source: [Tessella] Source: [Pronom7 SRD – S5.6.6.1]	M
S5.3.7.6	The system must be able to associate multiple GDFR facet values with each property. Source: [MEETING-WIN/08/08/05] Source: [Pronom7 SRD – S5.1.0.3]	M
S5.3.7.7	The system must be able to associate multiple options (the discrete values of a property) with each property. For example, the inherent compression property has the discrete values (which would be its property options): lossy compression, lossless compression, no compression. Source: [Pronom7 SRD – S5.2.7.2.2]	M

Label	Requirement	Necessity
S5.3.7.8	For each property option, the system must be able to record the following information: <ul style="list-style-type: none"> • Display Name • Description. Source: [Pronom7 SRD – S5.2.7.2.1]	M
S5.3.7.9	The system must be able to associate a file format with a particular value of an inherent property (i.e. with a particular inherent property option) and flag whether this inherent property value is considered to be high risk for the file format. Source: [Tessella] Source: [Pronom7 SRD – S5.3.1.3.1, Pronom7 SRD – S5.3.1.3.2]	M
S5.3.7.10	The system must be able to associate a risk score with each value of an inherent property (i.e. with a particular inherent property option). Source: [Pronom7 SRD – S5.3.1.2.1, Pronom7 SRD – S5.3.1.2.2]	M
S5.3.7.11	The system must be able to hold information about the risks associated with file format instance properties, whether these properties are: <ul style="list-style-type: none"> • continuous, and as such are described by ranges of numerical values, or <ul style="list-style-type: none"> • discrete, and as such are described by a set of text values or enumerations. Source: [Tessella] Source: [Pronom7 SRD - S5.3.1.4.1, Pronom7 SRD - S5.3.1.4.3]	M
S5.3.7.12	For each risk score associated with a value of a file format instance property, the system must be able to record the following: <ul style="list-style-type: none"> • the value of the risk score • a flag to indicate whether the value of the instance property is considered high risk. Source: [Pronom7 SRD - S5.3.1.4.2, Pronom7 SRD - S5.3.1.4.4]	M
S5.3.7.13	For each risk score associated with a file format instance property that is continuous, the system must be able to record the following: <ul style="list-style-type: none"> • the minimum value of the instance property associated with the particular risk score • the maximum value of the instance property associated with the particular risk score Source: [Pronom7 SRD - S5.3.1.4.2]	M
S5.3.7.14	For a risk score associated with a file format instance property that is discrete, the system must be able to associate it with a value of the instance property (i.e. with a particular instance property option). Source: [Pronom7 SRD - S5.3.1.4.4]	M

5.3.8 Provenance information

Note that provenance information is distinct from the audit trail. Provenance records contain information about the provider of the information to the Core Registry and can be used to judge its quality. Audit trail entries record information about the actions that have been carried out by an administrator to update the Core Registry.

Label	Requirement	Necessity
-------	-------------	-----------

Label	Requirement	Necessity
S5.3.8.1	<p>The system must be able to hold a single provenance entry for each of the following pieces of information held in the system:</p> <ul style="list-style-type: none"> • File formats. • Character encodings. • Compression techniques. • Reference files. • Internal signatures. • Software packages. • Hardware. • Storage media. • Software processes that can occur to file formats. • People. • Organisations. • Documentation. • External Identifiers. • Intellectual Property Rights. <p>Source: [P4IM, Clarification 1, PC3/D4 – 2.18] Source: [Pronom6 SRD - S4.1.2.1, Pronom7 SRD – S5.2.9.3, Pronom7 SRD – S5.1.1.16]</p>	M
S5.3.8.2	<p>For each provenance entry, the system must be able to record:</p> <ul style="list-style-type: none"> • The individual or organisation who provided the information (linked to maintained data on person or organisation agents). • Date on which the information was provided. • Notes on the provenance information. • Date on which the information was last updated. <p>Source: [P3SD – 3.2.7.2, P4IM] Source: [Pronom6 SRD - S4.1.2.2, Pronom7 SRD – S5.2.9.1, Pronom7 SRD – S5.2.9.2]</p>	M
S5.3.8.3	<p>The default values for the provenance entries should be:</p> <ul style="list-style-type: none"> • The date the information was provided: current system date. • Notes on the provenance information: Empty. • Date the information was updated: current system date. <p>Source: [Clarification 2] Source: [Pronom6 SRD - S4.1.2.3, Pronom7 SRD – S5.2.9.4]</p>	M
S5.3.8.4	<p>The default value for the organisation that provided the provenance information should be configurable for each instance of the Core Registry.</p> <p>Source: [Tessella] Source: [Pronom7 SRD – S5.2.9.4]</p>	M

5.3.9 External identifiers

Label	Requirement	Necessity
S5.3.9.1	<p>The system must be able to hold multiple external identifiers for each of the following information held in the system:</p> <ul style="list-style-type: none"> • File formats. • File format character encodings. • File format compression techniques. • Software packages. • Hardware. • Storage media. • Documentation. • Intellectual Property Rights. • Reference Files (only one external identifier is required for a reference file). <p>Source: [P4IM, Clarification 2, Review] Source: [Pronom6 SRD - S4.1.3.1, Pronom7 SRD – S5.2.10.1]</p>	M

Label	Requirement	Necessity
S5.3.9.2	<p>The system must maintain the following information about each external identifier:</p> <ul style="list-style-type: none"> Identifier type (linked to a maintained list of identifier types). Identifier value. Notes on the identifier. <p>Source: [P4IM, Clarification 2] Source: [Pronom6 SRD - S4.1.3.2, Pronom7 SRD – S5.2.10.2, Pronom7 SRD – S5.2.10.3, Pronom7 SRD – S5.1.1.10]</p>	M
S5.3.9.3	Future enhancement, so removed from this document.	X

5.3.10 Classification Facet Values

Label	Requirement	Necessity
S5.3.10.1	<p>The system must be able to record a list of GDFR facets from the GDFR system of faceted classification including the following information:</p> <ul style="list-style-type: none"> Facet name Description <p>Source: [PC3/D4 - 2.21] Source: [Pronom7 SRD – S5.6.22.1]</p>	M
S5.3.10.2	<p>The system must be able to record a list of GDFR facet value records from the GDFR system of faceted classification including the following information:</p> <ul style="list-style-type: none"> a link to a Facet a Facet Value Description <p>Source: [PC3/D4 - 2.21] Source: [Pronom7 SRD – S5.6.22.2]</p>	M
S5.3.10.3	<p>The system must include the following facet names in the list of facets:</p> <ul style="list-style-type: none"> genre (mandatory) role (mandatory) subsidiary-genre composition form constraint basis domain transform <p>Source: [PC3/D4 - 2.21] Source: [Pronom7 SRD – S5.6.22.3]</p>	M
S5.3.10.4	<p>The system must include the following in the list of facet values associated with the “genre” facet name:</p> <ul style="list-style-type: none"> aggregate any database dataset executable model moving-image other presentation sound spreadsheet still-image text <p>Source: [PC3/D4 - 2.21] Source: [Pronom7 SRD – S5.6.22.4]</p>	M

Label	Requirement	Necessity
S5.3.10.5	The system must include the following in the list of facet values associated with the “role” facet name: <ul style="list-style-type: none"> • family • file-format • encoding • serialization Source: [PC3/D4 - 2.21] Source: [Pronom7 SRD – S5.6.22.5]	M
S5.3.10.6	The system must include the following in the list of facet values associated with the “composition” facet name. <ul style="list-style-type: none"> • unitary • container-bundle • container-wrapper Source: [PC3/D4 - 2.21] Source: [Pronom7 SRD – S5.6.22.6]	M
S5.3.10.7	The system must include the following in the list of facet values associated with the “form” facet name. <ul style="list-style-type: none"> • binary • text Source: [PC3/D4 - 2.21] Source: [Pronom7 SRD – S5.6.22.7]	M
S5.3.10.8	The system must include the following in the list of facet values associated with the “constraint” facet name. <ul style="list-style-type: none"> • structured • unstructured Source: [PC3/D4 - 2.21] Source: [Pronom7 SRD – S5.6.22.8]	M
S5.3.10.9	The system must include the following in the list of facet values associated with the “basis” facet name. <ul style="list-style-type: none"> • sampled • symbolic Source: [PC3/D4 - 2.21] Source: [Pronom7 SRD – S5.6.22.9]	M
S5.3.10.10	The system must include the following in the list of facet values associated with the “domain” facet name. <ul style="list-style-type: none"> • astronomy • cad-cam • gis • web-archive Source: [PC3/D4 - 2.21] Source: [Pronom7 SRD – S5.6.22.10]	M
S5.3.10.11	The system must include the following in the list of facet values associated with the “transform” facet name. <ul style="list-style-type: none"> • compression • encryption • message-digest Source: [PC3/D4 - 2.21] Source: [Pronom7 SRD – S5.6.22.11]	M

5.4 System Information

This section contains details of the information recorded by the system in order to support some of the functional requirements

5.4.1 Audit Entries

This section contains details of the information recorded by the system to support the auditing of changes to the data held in the database.

Label	Requirement	Necessity
S5.4.1.1	The system must create an audit entry whenever data is created, updated, or deleted from the Core Registry Source: [PC3/D4 – 2.13] Source: [Pronom7 SRD – S6.3.1]	M
S5.4.1.2	For every database table for which changes are to be audited, the system must be able to uniquely identify the record within the database by the table name and a single column primary key column. Source: [Tessella] Source: [Pronom7 SRD – S5.5.1.1] Note: Consistency across all auditable tables will allow a generic approach to auditing of all tables to be taken.	M
S5.4.1.3	For each audit entry, the system must be able to record the following information: <ul style="list-style-type: none"> • Table affected by the change • Primary key of the record that changed • Modification Time • Modification Type (add/delete/modify). • The administrator who made the modification. Source: [Tessella] Source: [Pronom7 SRD – S5.5.1.2, Pronom7 SRD – S5.5.1.3, Pronom7 SRD – S5.5.1.4, Pronom7 SRD – S5.5.1.5]	M
S5.4.1.4	Future enhancement, so removed from this document.	X
S5.4.1.5	Future enhancement, so removed from this document.	X
S5.4.1.6	Future enhancement, so removed from this document.	X
S5.4.1.7	An audit entry should be added automatically whenever one of the following entities is added, modified or deleted: <ul style="list-style-type: none"> • agent • pathway, • hardware component, • documentation, • software product, • technical environment and • IPR. Source: [PA-tool registry – 5.7.7] Source: [Pronom7 SRD – S5.5.1.6] Note: Changes to all database tables will be audited.	M
S5.4.1.8	Future enhancement, so removed from this document.	X
S5.4.1.9	Future enhancement, so removed from this document.	X
S5.4.1.10	Future enhancement, so removed from this document.	X

5.4.2 Version Control, Rollback and Synchronisation

In the future it would be desirable to extend the auditing functionality in the system to include version control, rollback and synchronisation.

5.4.3 Technology Watch Alerts

PLATO (the preservation planning Planets sub-project) has requested a technology watch function in PCR to identify changes to the data which will potentially impact preservation planning and send out alerts to subscribers when these changes occur.

Label	Requirement	Necessity
S5.4.3.1	The system must generate alerts when changes (i.e. creation, modification, deletion) are made to specified information stored in the system. Source: [PC3/D6 – 2.2]	M

Label	Requirement	Necessity
S5.4.3.2	<p>The specified information to generate alerts for is any changes (creation/modification/deletion) to the following.</p> <p>For all core entities:</p> <ul style="list-style-type: none"> • description • name • version • release date • note <p>In addition:</p> <p>For file formats:</p> <ul style="list-style-type: none"> • binary flag • disclosure level • byte order. <p>For software packages:</p> <ul style="list-style-type: none"> • service pack details • package availability • source availability. <p>For compression techniques:</p> <ul style="list-style-type: none"> • lossiness. <p>For character encodings:</p> <ul style="list-style-type: none"> • code page • code unit width • encoding form width. <p>Source: [PC3/D6 – 2.2, TESTBED-AUG-2009]</p>	M
S5.4.3.3	Future enhancement, so removed from this document.	X
S5.4.3.4	Future enhancement, so removed from this document.	X
S5.4.3.5	<p>The system must publish each alert to all subscribers to that alert.</p> <p>Source: [PC3/D6 – 2.3, PC3/D6 – 2.4]</p>	M
S5.4.3.6	Future enhancement, so removed from this document.	X

5.4.4 Quarantine Area for Testbed Results Storage

In the future, Testbed (the testing and research environment Planets sub-project) will be storing aggregated results in PCR. These results will be submitted to a PCR web service and stored in a quarantine area until an administrator can review them.

Label	Requirement	Necessity
S5.4.4.1	Future enhancement, so removed from this document.	X

6. Legacy search interface Requirements

The legacy search interface allow users to search the repository, view the results of these searches and see detailed reports about the core entities held within the repository. This section describes these requirements.

6.1 General Requirements

The legacy search web interface is designed to be part of the National Archives' web site and therefore must conform to the National Archives' site style. Similarly, the functional parts of the Core Registry are wrapped in descriptive pages that describe its context and purpose. These pages need to be configurable by the National Archives and thus are not part of the deliverables.

Label	Requirement	Necessity
S6.1.1	The user interface must be web browser-based. Source: [P4URD – 1] Source: [Pronom6 SRD – S5.1.1, Pronom7 SRD – S7.1.7.1]	M
S6.1.2	The user interface must be intuitive and easy to use without training. Source: [P4URD – 2] Source: [Pronom6 SRD – S5.1.2, Pronom7 SRD – S7.1.7.2] N.B. Marked as desirable because impossible to objectively validate.	D
S6.1.3	The Core Registry user interface should contain the home page provided by TNA. Source: [Current system, Clarification 2] Source: [Pronom6 SRD – S5.1.3, Pronom7 SRD – S7.1.7.4]	M
S6.1.4	The home page must be configurable by appropriately qualified IT personnel. Source: [Current system] Source: [Pronom6 SRD – S5.1.4, Pronom7 SRD – S7.1.7.8]	M
S6.1.5	The home page should allow users to enter a search page. Source: [Current system] Source: [Pronom6 SRD – S5.1.5, Pronom7 SRD – S7.1.1.1]	M
S6.1.6	The home page should allow users to enter a help system. Source: [Current system] Source: [Pronom6 SRD – S5.1.6, Pronom7 SRD – S7.1.1.2]	M
S6.1.7	The home page should allow users to enter an on-line submission form. Source: [Current system] Source: [Pronom6 SRD – S5.1.7, Pronom7 SRD – S7.1.1.3]	M
S6.1.8	The source code for the search pages must be delivered so that it can be configured by appropriately qualified IT personnel. Source: [Clarification 2] Source: [Pronom6 SRD – S5.1.8, Pronom7 SRD – S7.1.7.5]	M
S6.1.9	The source code for the help page must be delivered so that it can be configured by appropriately qualified IT personnel. Source: [Clarification 2] Source: [Pronom6 SRD – S5.1.9, Pronom7 SRD – S7.1.7.6]	M
S6.1.10	The source code for the on-line submission form must be delivered so that it can be configured by appropriately qualified IT personnel. Source: [Clarification 2] Source: [Pronom6 SRD – S5.1.10, Pronom7 SRD – S7.1.7.7]	M
S6.1.11	The Core Registry must be able to resolve PUID URIs to human-readable XHTML. Source: [TWSRD] Source: [Pronom6 SRD – S5.1.11, Pronom7 SRD – S7.3.1.4]	M
S6.1.12	Future enhancement, so removed from this document.	X

6.2 Search and Reporting Requirements

6.2.1 Search layout

Label	Requirement	Necessity
S6.2.1.1	The search must be laid out on a number of tabs: <ul style="list-style-type: none"> File format tab. Persistent Unique Identifier tab. Software tab. Vendor tab. Lifecycles tab. Simple search Migration pathways Source: [Current system, P4URD – 19, P4URD – 22, P4URD – 24, TWSRD, PA-tool registry – STU4, PA-tool registry – STU5, PA-tool registry – STU6, PA-tool registry – STU7, PA-tool registry – STU8] Source: [Pronom6 SRD – S6.1.0.1, Pronom7 SRD – S7.1.2.1.1]	M
S6.2.1.2	The file format tab must allow: <ul style="list-style-type: none"> Searches for formats (see section 6.2.2). Searches for software packages capable of processing (creating/rendering/identifying/validating/extracting metadata from) file formats (see section 6.2.3). Source: [Current system, P4URD – 17, P4URD – 18] Source: [Pronom6 SRD – S6.1.0.2, Pronom7 SRD – S7.1.2.4.1.1]	M
S6.2.1.3	The Persistent Unique Identifier (PUID) tab must allow: <ul style="list-style-type: none"> Search by PUID (see section 6.2.4). Source: [P4URD – 19] Source: [Pronom6 SRD – S6.1.0.3, Pronom7 SRD – S7.1.2.1.2]	M
S6.2.1.4	The software tab must allow: <ul style="list-style-type: none"> Searches for software (see section 6.2.5) Source: [Current system] Source: [Pronom6 SRD – S6.1.0.4]	M
S6.2.1.5	The vendor tab (see section 0) must allow: <ul style="list-style-type: none"> Search by vendor name. Search by developer of core entities. Source: [Current system, P4URD - 23, 11-Feb-2005] Source: [Pronom6 SRD – S6.1.0.5]	M
S6.2.1.6	The lifecycles tab must allow: <ul style="list-style-type: none"> Searches by support period (see section 6.2.7) Searches by release dates (see section 6.2.8). Source: [Current system, P4URD- 24, P4URD- 25, P4URD- 26, P4URD- 27] Source: [Pronom6 SRD – S6.1.0.6, Pronom7 SRD – S7.1.2.8.2.1]	M
S6.2.1.7	An additional search tab for simple searches that searches across formats, software, character encodings and compression techniques. Source: [TWSRD] Source: [Pronom6 SRD – S6.1.0.7, Pronom7 SRD – S7.1.2.10.1]	M
S6.2.1.8	Future enhancement, so removed from this document.	X
S6.2.1.9	Future enhancement, so removed from this document.	X
S6.2.1.10	The Migration pathways search tab must allow: <ul style="list-style-type: none"> Searches by start and end format (see 6.1.18) Source: [TWSRD] Source: [Pronom6 SRD – S6.1.0.9]	M

6.2.2 Searches for formats by extension, name or risk (format tab)

Label	Requirement	Necessity
S6.2.2.1	The user interface must allow users to search for formats using a given extension. Source: [Current system, P3SD - S2.3.1.1, SD – S2.3.1.9] Source: [Pronom6 SRD – S6.1.1.1, Pronom7 SRD – S7.1.2.4.4.1]	M

Label	Requirement	Necessity
S6.2.2.2	The user interface must allow users to search for formats by providing the format name (or partial name) or alias (or part thereof). Source: [P4URD - 17, DemoProgress] Source: [Pronom6 SRD – S6.1.1.2, Pronom7 SRD – S7.1.2.4.5.1]	M
S6.2.2.3	When listing the file formats, the search results must show: <ul style="list-style-type: none"> List of possible extensions (concatenated together in alphabetical order). Persistent unique identifier. Format name. Format version. Source: [Current system, P4URD – 29, Clarification 2, Review]. Source: [Pronom6 SRD – S6.1.1.3]	M
S6.2.2.4	All items in this search result listing must be a hyperlink to the format detailed report associated with the format in that row. Source: [P4URD - 29] Source: [Pronom6 SRD – S6.1.1.4, Pronom7 SRD – S7.1.2.4.1.3]	M
S6.2.2.5	The Core Registry must be able to search on format risks by specifying a format type (optional) and a risk threshold (above or below). The resulting report should show a list of formats. Source: [TWSRD] Source: [Pronom6 SRD – S6.1.1.5, Pronom7 SRD – S7.1.2.4.6.1]	M
S6.2.2.6	The user interface must allow users to search for formats by file format instance property (e.g. image width). Source: [PC3/D4 – 2.3, PC3/D4 – 2.14] Source: [Pronom7 SRD – S7.1.2.4.3.1]	M

6.2.3 Searches for software packages capable of processing formats (format tab)

Label	Requirement	Necessity
S6.2.3.1	The user interface must allow users to search for all software packages capable of processing (creating/rendering/identifying/validating/extracting metadata from) all formats with a given extension. Source: [Current system, P3SD - S2.3.1.1, SD - S2.3.1.9] Source: [Pronom6 SRD – S6.1.2.1, Pronom7 SRD – S7.1.2.5.2.1]	M
S6.2.3.2	The user interface must allow users to search for all software packages capable of processing (creating/rendering/identifying/validating/extracting metadata from) all formats with a given name (or part of name). Source: [P4URD - 18] Source: [Pronom6 SRD – S6.1.2.2, Pronom7 SRD – S7.1.2.5.2.2]	M
S6.2.3.3	When searching for software packages capable of processing a given file format, users must be able to restrict the search by: <ul style="list-style-type: none"> Software that can create the format Software that can render the format Software that can identify the format Software that can validate the format Software that can extract metadata from the format Software that can carry out any of the above processes (the default). Source: [10-Feb-2005] Source: [Pronom6 SRD – S6.1.2.2.1, Pronom7 SRD – S7.1.2.5.2.3]	M
S6.2.3.4	When searching for software packages capable of processing (creating/rendering/identifying/validating/extracting metadata from) a given file format, users must be able to restrict the search by: <ul style="list-style-type: none"> Unsupported software packages. Supported software packages. All software packages (the default). Source: [P3SD - S2.3.1.1, SD - S2.3.1.6] Source: [Pronom6 SRD – S6.1.2.3, Pronom7 SRD – S7.1.2.5.2.4]	M

Label	Requirement	Necessity
S6.2.3.5	Future enhancement, so removed from this document.	X
S6.2.3.6	Future enhancement, so removed from this document.	X
S6.2.3.7	When searching for software packages capable of processing a given file format, users must be able to restrict the search by: <ul style="list-style-type: none"> The operating system the software runs under The hardware the software runs on Source: [PC3/D4 – 2.7] Source: [Pronom7 SRD – S7.1.2.5.2.7]	M
S6.2.3.8	The results of the search must list all file formats together with information on the software packages that can create/render/identify/validate/extract metadata from that format. Source: [Current system, Clarification 1] Source: [Pronom6 SRD – S6.1.2.4, Pronom7 SRD – S7.1.2.5.3.1]	M
S6.2.3.9	When listing the formats the results must show PUID, format name and format version. Source: [Clarification 1] Source: [Pronom6 SRD – S6.1.2.4.01, Pronom7 SRD – S7.1.2.5.3.2]	M
S6.2.3.10	All format listing items must be a hyperlink to the format detailed report associated with the format in that row (see S6.1.12.1 below). Source: [Clarification 2] Source: [Pronom6 SRD – S6.1.2.4.2, Pronom7 SRD – S7.1.2.2.1]	M
S6.2.3.11	When listing the software packages that can process each format, the search results must show: <ul style="list-style-type: none"> Software package name. Software package version. Release date Vendor name. The types of processes – render (read), create (write), identify, validate and/or extract metadata - that the software package can perform on the format, as long as the search included those process types. The invariance description, if the format can be rendered by the software package. Source: [Current system, 10-Feb-2005, PA-tool registry – STU1] Source: [Pronom6 SRD – S6.1.2.5, Pronom7 SRD – S7.1.2.5.3.3, Pronom7 SRD – S7.1.2.5.3.4a]	M
S6.2.3.12	The software package name, software package version, release date and vendor name columns are always included in the results. The process column only includes those processes which are specified in the search criteria. The invariance column is only displayed when the render process is included in the search criteria. Source: [10-Feb-2005, DemoProgress] Source: [Pronom6 SRD – S6.1.2.5.1, Pronom7 SRD – S7.1.2.5.3.5]	M
S6.2.3.13	If no software package can process (create/render/identify/validate/extract metadata from) a format in the list, the search results should inform the user that “No software packages are compatible with this file format”. Source: [Current system] Source: [Pronom6 SRD – S6.1.2.6]	M
S6.2.3.14	The software package name column in the list of search results must be a link to a detailed description of the software package. Source: [Current system] Source: [Pronom6 SRD – S6.1.2.7, Pronom7 SRD – S7.1.2.2.1]	M
S6.2.3.15	The vendor name column in the list of search results must be a link to a detailed description of the vendor. Source: [Current system] Source: [Pronom6 SRD – S6.1.2.8, Pronom7 SRD – S7.1.2.2.1]	M

6.2.4 Search for format by PUID (PUID tab)

Label	Requirement	Necessity
S6.2.3.1	The user interface must allow users to search for a format by Persistent unique identifier. Source: [P4URD – 19, Clarification 1] Source: [Pronom6 SRD – S6.1.3.1, Pronom7 SRD – S7.1.2.1.2, Pronom7 SRD – S7.1.2.4.2.1, Pronom7 SRD – S7.1.2.10.2]	M
S6.2.4.2	For any file format PUID searched for, the system must display the detailed report associated with that file format. Source: [Pronom6 Implementation] Source: [Pronom7 SRD – S7.1.2.4.2.2, Pronom7 SRD – S7.1.2.10.3]	M

6.2.5 Search for software (software tab)

Label	Requirement	Necessity
S6.2.5.1	The user interface must allow users to search for all software by name or part of name. Source: [P3SD - S2.3.1.1, P3SD - S2.3.1.4] Source: [Pronom6 SRD – S6.1.4.1, Pronom7 SRD – S7.1.2.5.1.1]	M
S6.2.5.2	The user interface must allow users to search for all software by vendor name or part of vendor name. Source: [P3SD - S2.3.1.1, P3SD - S2.3.1.4] Source: [Pronom6 SRD – S6.1.4.2, Pronom7 SRD – S7.1.2.5.1.2]	M
S6.2.5.3	When listing the software packages, the search results must show: <ul style="list-style-type: none"> • Software package name. • Software package version. • Vendor name. Source: [Current system] Source: [Pronom6 SRD – S6.1.4.3, Pronom7 SRD – S7.1.2.5.1.3]	M
S6.2.5.4	The software name column in the list of search results must be a link to a detailed description of the software package. Source: [Current system] Source: [Pronom6 SRD – S6.1.4.4, Pronom7 SRD – S7.1.2.2.1]	M
S6.2.5.5	The vendor name column in the list of search results must be a link to a detailed description of the vendor. Source: [Current system] Source: [Pronom6 SRD – S6.1.4.5, Pronom7 SRD – S7.1.2.2.1]	M
S6.2.5.6	Future enhancement, so removed from this document.	X
S6.2.5.7	Future enhancement, so removed from this document.	X
S6.2.5.8	Future enhancement, so removed from this document.	X

6.2.6 Search by vendor name (vendor tab)

Label	Requirement	Necessity
S6.2.5.1	The user interface must allow users to search for all vendors by vendor name or part of vendor name. Source: [P3SD - S2.3.1.1, P3SD - S2.3.1.3] Source: [Pronom6 SRD – S6.1.5.1, Pronom7 SRD – S7.1.2.7.2.1]	M
S6.2.6.2	The user must be able to optionally choose from the subset of core entity categories ('file format', 'software package', 'character encoding' or 'compression technique') and restrict the search results to those vendors that develop that core entity. Source: [P4URD – 23, Clarification 1, 11-Feb-2005] Source: [Pronom6 SRD – S6.1.5.1.1, Pronom7 SRD – S7.1.2.7.2.2]	M
S6.2.6.3	The list of search results for vendor must show: <ul style="list-style-type: none"> • Vendor name. • Vendor URL. Source: [Current system, Clarification 2] Source: [Pronom6 SRD – S6.1.5.2, Pronom7 SRD – S7.1.2.7.1.1]	M

Label	Requirement	Necessity
S6.2.6.4	The vendor name column in the list of search results must be a link to a detailed description of the vendor. Source: [Current system] Source: [Pronom6 SRD – S6.1.5.3, Pronom7 SRD – S7.1.2.2.1, Pronom7 SRD – S7.1.2.7.1.2]	M
S6.2.6.5	The vendor URL must be a hyperlink which when clicked should open a new browser page at the specified URL. Source: [Current system] Source: [Pronom6 SRD – S6.1.5.4, Pronom7 SRD – S7.1.2.2.2]	M

6.2.7 Search by support period (lifecycles tab)

Label	Requirement	Necessity
S6.2.7.1	The user interface must allow users to search for all file formats or software packages, supported (or not supported) on a given date. Source: [P3SD - S2.3.1.1, P3SD - 2.3.1.5] Source: [Pronom6 SRD – S6.1.7.1, Pronom7 SRD – S7.1.2.8.2.9]	M
S6.2.7.2	The user must be able to specify whether they wish to look at supported or unsupported core entities through choosing 'supported' or 'unsupported' from a list box. Source: [P4URD – 25] Source: [Pronom6 SRD – S6.1.7.2, Pronom7 SRD – S7.1.2.8.2.10]	M
S6.2.7.3	The user must be able to choose a core entity category (limited to 'file format' and 'software component') and restrict the search results to those core entities within that category. Source: [P4URD – 27, Clarification 1, Review] Source: [Pronom6 SRD – S6.1.7.2.1, Pronom7 SRD – S7.1.2.8.2.11]	M
S6.2.7.4	Future enhancement, so removed from this document.	X
S6.2.7.5	When listing the core entities, the search results must show: <ul style="list-style-type: none"> • Core entity name. • Core entity version. • Vendor name. • Supported until date (showing 'No date set' if not withdrawn or 'Withdrawn (date unknown)' if marked as withdrawn but with no withdrawn date). Source: [Current system, Clarification 2] Source: [Pronom6 SRD – S6.1.7.3, Pronom7 SRD – S7.1.2.8.1.1]	M
S6.2.7.6	The core entity name column in the list of search results must be a link to a detailed description of the core entity. Source: [Current system] Source: [Pronom6 SRD – S6.1.7.4, Pronom7 SRD – S7.1.2.2.1]	M
S6.2.7.7	The vendor name column in the list of search results must be a link to a detailed description of the vendor. Source: [Current system] Source: [Pronom6 SRD – S6.1.7.5, Pronom7 SRD – S7.1.2.2.1]	M
S6.2.7.8	When searching for either 'supported' or 'unsupported' core entities (limited to file formats and software packages) on a given date, the system must show all core entities that have been withdrawn but for which no withdrawal date has been set. The only exception to this is that when there is a query for 'supported' core entities on a date in the future, the system must not show any core entities that have been withdrawn, but for which no withdrawal date has been set. Source: [Clarification 2, 14-Apr-2005] Source: [Pronom6 SRD – S6.1.7.6, Pronom7 SRD – S7.1.2.8.2.13]	M

Label	Requirement	Necessity																																																																	
S6.2.7.9	<p>When searching for either 'supported' or 'unsupported' core entities on a given date, if the release date and/or the withdrawn date of a core entity is unknown, then the logic for whether or not that core entity should be displayed is set out in the following tables.</p> <table border="1"> <thead> <tr> <th rowspan="2">Release date</th> <th rowspan="2">Withdrawn date</th> <th rowspan="2">Withdrawn flag</th> <th colspan="2">Core entity is displayed</th> </tr> <tr> <th>Supported queries</th> <th>Unsupported queries</th> </tr> </thead> <tbody> <tr> <td rowspan="4">NULL</td> <td rowspan="2">NULL</td> <td>Y</td> <td>Y</td> <td>Y</td> </tr> <tr> <td>N</td> <td>Y</td> <td>N</td> </tr> <tr> <td>> query date</td> <td>Y</td> <td>Y</td> <td>N</td> </tr> <tr> <td>= query date</td> <td>Y</td> <td>Y</td> <td>Y</td> </tr> <tr> <td rowspan="4">> query date</td> <td rowspan="4">NULL</td> <td>< query date</td> <td>Y</td> <td>N</td> </tr> <tr> <td>Y</td> <td>N</td> <td>Y</td> </tr> <tr> <td>N</td> <td>N</td> <td>Y</td> </tr> <tr> <td>Y</td> <td>Y</td> <td>Y</td> </tr> <tr> <td rowspan="2">= query date</td> <td rowspan="2">NULL</td> <td>N</td> <td>Y</td> <td>N</td> </tr> <tr> <td>Y</td> <td>Y</td> <td>Y</td> </tr> <tr> <td rowspan="2">< query date</td> <td rowspan="2">NULL</td> <td>Y</td> <td>Y</td> <td>Y</td> </tr> <tr> <td>N</td> <td>Y</td> <td>N</td> </tr> </tbody> </table> <p>N.B. Withdrawn flag = N corresponds to the withdrawn flag being set to false, or being null. <i>If the query date after the date on which it was recorded that the entity has been withdrawn, then for the case where the withdrawn flag is set and the withdrawn date is unknown, the core entity should be considered unsupported; i.e. it should not be shown for queries for supported core entities. This is as shown in the following table.</i></p> <table border="1"> <thead> <tr> <th rowspan="2">Query date</th> <th rowspan="2">Withdrawn date</th> <th rowspan="2">Withdrawn flag</th> <th colspan="2">Core entity is displayed</th> </tr> <tr> <th>Supported queries</th> <th>Unsupported queries</th> </tr> </thead> <tbody> <tr> <td>> withdraw recorded date</td> <td>NULL</td> <td>Y</td> <td>N</td> <td>Y</td> </tr> </tbody> </table> <p>Source: [23-Feb-2005, 14-Apr-2005] Source: [Pronom6 SRD – S6.1.7.7, Pronom7 SRD – S7.1.2.8.2.14]</p>	Release date	Withdrawn date	Withdrawn flag	Core entity is displayed		Supported queries	Unsupported queries	NULL	NULL	Y	Y	Y	N	Y	N	> query date	Y	Y	N	= query date	Y	Y	Y	> query date	NULL	< query date	Y	N	Y	N	Y	N	N	Y	Y	Y	Y	= query date	NULL	N	Y	N	Y	Y	Y	< query date	NULL	Y	Y	Y	N	Y	N	Query date	Withdrawn date	Withdrawn flag	Core entity is displayed		Supported queries	Unsupported queries	> withdraw recorded date	NULL	Y	N	Y	M
Release date	Withdrawn date				Withdrawn flag	Core entity is displayed																																																													
		Supported queries	Unsupported queries																																																																
NULL	NULL	Y	Y	Y																																																															
		N	Y	N																																																															
	> query date	Y	Y	N																																																															
	= query date	Y	Y	Y																																																															
> query date	NULL	< query date	Y	N																																																															
		Y	N	Y																																																															
		N	N	Y																																																															
		Y	Y	Y																																																															
= query date	NULL	N	Y	N																																																															
		Y	Y	Y																																																															
< query date	NULL	Y	Y	Y																																																															
		N	Y	N																																																															
Query date	Withdrawn date	Withdrawn flag	Core entity is displayed																																																																
			Supported queries	Unsupported queries																																																															
> withdraw recorded date	NULL	Y	N	Y																																																															

6.2.8 Search by release dates (lifecycles tab)

Label	Requirement	Necessity
S6.2.8.1	<p>The user interface must allow users to search for all core entities (limited to 'file formats' or 'software packages') released before a given date. Source: [P3SD - S2.3.1.1, P3SD - 2.3.1.7] Source: [Pronom6 SRD – S6.1.8.1, Pronom7 SRD – S7.1.2.8.2.2]</p>	M
S6.2.8.2	<p>The user interface must allow users to search for all core entities (limited to 'file formats' or 'software packages') released after a given date. Source: [P3SD - S2.3.1.1, P3SD - 2.3.1.7, Clarification 2] Source: [Pronom6 SRD – S6.1.8.1.1, Pronom7 SRD – S7.1.2.8.2.3]</p>	M
S6.2.8.3	<p>The user interface must allow users to search for all core entities (limited to 'file formats' or 'software packages') released within a given date range. Source: [P3SD - S2.3.1.1, P3SD - 2.3.1.7, Clarification 2] Source: [Pronom6 SRD – S6.1.8.2, Pronom7 SRD – S7.1.2.8.2.4]</p>	M

Label	Requirement	Necessity
S6.2.8.4	The user must be able to choose these three options by selecting from a list box ('before', 'after', 'between') with the first two options showing one date box and the latter option displaying two. Source: [P4URD – 26, Clarification 2] Source: [Pronom6 SRD – S6.1.8.3, Pronom7 SRD – S7.1.2.8.2.5]	M
S6.2.8.5	The user must be able to choose a core entity category (limited to 'file format' or 'software package') and restrict the search results to those core entities within that category. Source: [P4URD – 27, Clarification 1, Review] Source: [Pronom6 SRD – S6.1.8.3.1, Pronom7 SRD – S7.1.2.8.2.6]	M
S6.2.8.6	Future enhancement, so removed from this document.	X
S6.2.8.7	When listing the core entities, the search results must show: <ul style="list-style-type: none"> • Core entity name. • Core entity version. • Vendor name. • Release date. Source: [Current system] Source: [Pronom6 SRD – S6.1.8.4, Pronom7 SRD – S7.1.2.8.1.1]	M
S6.2.8.8	The core entity name column in the list of search results must be a link to a detailed description of the core entity. Source: [Current system] Source: [Pronom6 SRD – S6.1.8.5, Pronom7 SRD – S7.1.2.8.1.2]	M
S6.2.8.9	The vendor name column in the list of search results must be a link to a detailed description of the vendor. Source: [Current system] Source: [Pronom6 SRD – S6.1.8.6, Pronom7 SRD – S7.1.2.8.1.3]	M
S6.2.8.10	When searching for core entities by release date, those core entities with a null release date should not be included in the results. Source: [23-Feb-2005] Source: [Pronom6 SRD – S6.1.8.7, Pronom7 SRD – S7.1.2.8.2.8]	M

6.2.9 Storage media search (storage media tab)

In the future the legacy search interface should allow users to search by storage media type.

6.2.10 Simple search (simple search tab)

Label	Requirement	Necessity
S6.2.10.1	The system must provide a simple search facility allowing keyword searching across multiple registry fields from a single search box. The tables to be used in the search are: <ul style="list-style-type: none"> • Format • Software • Character encoding • Compression technique The fields to search are: <ul style="list-style-type: none"> • Name • Description • Version • Aliases • Identifiers Source: [URD: REG-UI 01, MEET 4, Review] Source: [Pronom6 SRD – S6.1.11.1, Pronom7 SRD – S7.1.2.10.1]	M
S6.2.10.2	The name and version fields should be searchable as a combined field. E.g. It should be possible to search for Excel 2000 and have Excel 2000 returned in the results, and preferably on the first page of results. Source: [Tessella] Source: [Pronom7 SRD – S7.1.2.10.4]	D

6.2.11 Detailed format report

Label	Requirement	Necessity
S6.2.11.1	<p>A detailed report on a file format must display all the information held on that format in a tabular format including the following sub-divisions:</p> <ul style="list-style-type: none"> • Summary. • Resources (i.e. documentation and additional resources). • Signatures. • Compression. • Character Encoding. • Rights (i.e. IPR). • Related Files (i.e. Reference Files, XCEL files, XCDL schema files). • Risk Assessments. <p>Source: [P4URD – 30, Clarification 2, 27-Jan-2005, TWSRD, PC3/D4 – 2.11, MEETING-WIN/08/08/05] Source: [Pronom6 SRD – S6.1.12.1, Pronom7 SRD – S7.1.2.4.7.1] Note: Documentation has been renamed Resources so that it is more generic.</p>	M
S6.2.11.2	<p>Any external reference held about a file format that is a URL must be a hyperlink which when clicked should open a new browser page at the specified URL.</p> <p>Source: [Clarification 2] Source: [Pronom6 SRD – S6.1.12.2, Pronom7 SRD – S7.1.2.2.2]</p>	M
S6.2.11.3	<p>Any external reference held about a file format character encoding that is a URL must be a hyperlink which when clicked should open a new browser page at the specified URL.</p> <p>Source: [Clarification 2] Source: [Pronom6 SRD – S6.1.12.3, Pronom7 SRD – S7.1.2.2.2]</p>	M
S6.2.11.4	<p>Any external reference held about a file format compression technique that is a URL must be a hyperlink which when clicked should open a new browser page at the specified URL.</p> <p>Source: [Clarification 2] Source: [Pronom6 SRD – S6.1.12.4, Pronom7 SRD – S7.1.2.2.2]</p>	M
S6.2.11.5	<p>Any external reference held about file format documentation that is a URL must be a hyperlink which when clicked should open a new browser page at the specified URL.</p> <p>Source: [Clarification 2] Source: [Pronom6 SRD – S6.1.12.5, Pronom7 SRD – S7.1.2.2.2]</p>	M
S6.2.11.6	<p>Any external reference held about a file format intellectual property rights that is a URL must be a hyperlink which when clicked should open a new browser page at the specified URL.</p> <p>Source: [Clarification 2] Source: [Pronom6 SRD – S6.1.12.6, Pronom7 SRD – S7.1.2.2.2]</p>	M
S6.2.11.7	<p>Any external reference held about a file format reference file that is a URL must be a hyperlink which when clicked should open a new browser page at the specified URL.</p> <p>Source: [Clarification 2] Source: [Pronom6 SRD – S6.1.12.7, Pronom7 SRD – S7.1.2.2.2]</p>	M
S6.2.11.8	<p>The detailed format report should include a list of all the other formats that are related to the format that is being reported on, together with the nature of the relationship (e.g., “is superseded by”) and a hyperlink to the detailed report on the other format.</p> <p>Source: [27-Jan-2005, PC3/D4 – 2.6, MEETING-WIN/08/08/05] Source: [Pronom6 SRD – S6.1.12.8, Pronom7 SRD – S7.1.2.4.7.4] N.B. Relationships recording priority information should not be displayed since the ‘priority’ relationship is used purely within the DROID format identification algorithm, and has no more general significance. Displaying this information is potentially confusing to users.</p>	M

Label	Requirement	Necessity
S6.2.11.9	<p>The detailed format report must list each internal signature including the name, description of the signature and the position type, offset, endian-ness and value of each byte sequence.</p> <p>Source: [PRONOM5a] Source: [Pronom6 SRD – S6.1.12.9, Pronom7 SRD – S7.1.2.4.7.12]</p>	M
S6.2.11.10	<p>The detailed format report should show information on the property risks (as a separate section)</p> <p>Source: [TWSRD, PC3/D4 – 2.19] Source: [Pronom6 SRD – S6.1.12.10, Pronom7 SRD – S7.1.2.4.7.22, Pronom7 SRD – S7.1.2.4.7.23]</p> <p>Note: in the case of inherent properties this will be a risk score associated with the value selected for the file format. In the case of instance properties, this will be a risk score associated with a particular value or range for the instance property for the given file format e.g. the instance property “image width” for PNG has a risk of 1000 when the value of the instance property is greater than 6000 pixels.</p>	M
S6.2.11.11	<p>The detailed format report should show information on the format properties (inherent & instance) held.</p> <p>Source: [TWSRD, PC3/D4 – 2.19] Source: [Pronom6 SRD – S6.1.12.11, Pronom7 SRD – S7.1.2.4.7.21]</p>	M
S6.2.11.12	<p>The detailed format report should display the facet name and value for each facet associated with the file format. This information should be displayed in the following order of importance:</p> <ul style="list-style-type: none"> • genre (mandatory) • subsidiary-genre • role (mandatory) • subsidiary-genre • composition • form • constraint • basis • domain • transform <p>Source: [PC3/D4 – 2.21] Source: [Pronom7 SRD – S7.1.2.4.7.5, Pronom7 SRD – S7.1.2.4.7.6]</p>	M
S6.2.11.13	<p>The detailed format report should display information about external assessments of a file format, including both a descriptive summary of the assessment and a link to the detailed assessment on an external website.</p> <p>Source: [PC3/D4 – 2.10] Source: [Pronom7 SRD – S7.1.2.4.7.8]</p>	M
S6.2.11.14	<p>The detailed format report should display (on the resources tab) any additional information stored about the file format in the Core Registry in the form of a link to either an external URL where the information can be found, or to a detailed agent report for a relevant organisation, and a description of the nature of the link.</p> <p>Source: [PC3/D4 – 2.11] Source: [Pronom7 SRD – S7.1.2.4.7.9, Pronom7 SRD – S7.1.2.4.7.10, Pronom7 SRD – S7.1.2.4.7.11]</p>	M
S6.2.11.15	<p>The detailed format report should display (on the related files tab) a list of XCEL files for the file format, and allow users to download them.</p> <p>Source: [PC3/D4 – 2.5] Source: [Pronom7 SRD – S7.1.2.4.7.17, Pronom7 SRD – S7.1.2.4.7.18]</p>	M

Label	Requirement	Necessity
S6.2.11.16	The detailed format report should display (on the related files tab) a list of XCDL schema files for the file format, and allow users to download them. Source: [PC3/D4 – 2.5] Source: [Pronom7 SRD – S7.1.2.4.7.19, Pronom7 SRD – S7.1.2.4.7.20]	M

6.2.12 Common requirements for detailed core entity reports

Label	Requirement	Necessity
S6.2.12.1	A detailed report on a core entity must display all the information held on that core entity in a tabular format. Source: [Pronom6 SRD – S6.1.13.1, Pronom7 SRD – S7.1.2.3.5]	M
S6.2.12.2	Future enhancement, so removed from this document.	X
S6.2.12.3	Future enhancement, so removed from this document.	X
S6.2.12.4	Future enhancement, so removed from this document.	X
S6.2.12.5	Future enhancement, so removed from this document.	X
S6.2.12.6	Within a detailed core entity report, the vendor and support vendor names must be a hyperlink to detailed description of the vendor. Source: [Current system] Source: [Pronom6 SRD – S6.1.13.2]	M
S6.2.12.7	Within a detailed core entity report, the file formats rendered (read), created (written), identified, validated and metadata extracted fields must be a hyperlink to a detailed description of the file format. Source: [P4URD – 31, 10-Feb-2005] Source: [Pronom6 SRD – S6.1.13.3, Pronom7 SRD – S7.1.2.2.1]	M
S6.2.12.8	Future enhancement, so removed from this document.	X
S6.2.12.9	Any external reference held about a core entity's documentation that is a URL must be a hyperlink which when clicked should open a new browser page at the specified URL. Source: [Clarification 2] Source: [Pronom6 SRD – S6.1.13.6, Pronom7 SRD – S7.1.2.2.2]	M
S6.2.12.10	Any external reference held about a core entity's intellectual property rights that is a URL must be a hyperlink which when clicked should open a new browser page at the specified URL. Source: [Clarification 2] Source: [Pronom6 SRD – S6.1.13.7, Pronom7 SRD – S7.1.2.2.2]	M

6.2.13 Detailed software report

Label	Requirement	Necessity
S6.2.13.1	A detailed report on a software package must display all the information held on that format in a tabular format including the following sub-divisions: <ul style="list-style-type: none"> • Summary. • File formats processed. • Images. • Documentation • IPR. Source: [Current system, 27-Jan-2005, PA-tool registry – STU9, PA-tool registry – STU9] Source: [Pronom6 SRD – S6.1.13.1.1, Pronom7 SRD – S7.1.2.5.7.2, Pronom7 SRD – S7.1.2.5.7.7]	M
S6.2.13.2	Any external reference held about a software package that is a URL must be a hyperlink which when clicked should open a new browser page at the specified URL. Source: [Clarification 2] Source: [Pronom6 SRD – S6.1.13.4, Pronom7 SRD – S7.1.2.2.2]	M

Label	Requirement	Necessity
S6.2.13.3	The detailed software report should include a list of all the other software packages that are related to the software that is being reported on, together with the nature of the relationship (e.g., "is superseded by") and a hyperlink to the detailed report on the other software. Source: [27-Jan-2005] Source: [Pronom6 SRD – S6.1.13.8, Pronom7 SRD – S7.1.2.5.7.10]	M
S6.2.13.4	The detailed software report must include the following information: <ul style="list-style-type: none"> • Associated programming language • Availability of source code • Availability of software package • Associated licence details Source: [PC3/D4 – 2.17] Source: [Pronom7 SRD – S7.1.2.5.7.4a, Pronom7 SRD – S7.1.2.5.7.4b]	M
S6.2.13.5	Future enhancement, so removed from this document.	X
S6.2.13.6	Future enhancement, so removed from this document.	X
S6.2.13.7	Future enhancement, so removed from this document.	X
S6.2.13.8	Future enhancement, so removed from this document.	X

6.2.14 Detailed agent report

Label	Requirement	Necessity
S6.2.14.1	A detailed description of an agent must display all the information held on that agent in a tabular format. Source: [Current system] Source: [Pronom6 SRD – S6.1.14.1, Pronom7 SRD – S7.1.2.7.3.1]	M
S6.2.14.2	Within a detailed agent report, the company website entry must be a hyperlink which when clicked should open a new browser page at the specified URL. Source: [Current system] Source: [Pronom6 SRD – S6.1.14.2, Pronom7 SRD – S7.1.2.2.2]	M
S6.2.14.3	Within a detailed agent report, the support URL must be a hyperlink which when clicked should open a new browser page at the specified URL. Source: [Current system] Source: [Pronom6 SRD – S6.1.14.3, Pronom7 SRD – S7.1.2.2.2]	M
S6.2.14.4	Within a detailed agent report, the contact e-mail must be a hyperlink which when clicked should open the default e-mail client with a new message addressed to that e-mail address. Source: [Current system] Source: [Pronom6 SRD – S6.1.14.4, Pronom7 SRD – S7.1.2.7.3.2]	M

6.2.15 General search results

Label	Requirement	Necessity
S6.2.15.1	If a search results in no data being returned, the user must be suitably informed. Source: [Current system] Source: [Pronom6 SRD – S6.1.16.1, Pronom7 SRD – S7.1.2.2.3]	M
S6.2.15.2	Each search results page that displays a list of results must indicate the nature of the search that was performed. Source: [Current system] Source: [Pronom6 SRD – S6.1.16.2, Pronom7 SRD – S7.1.2.2.4]	M

Label	Requirement	Necessity
S6.2.15.3	<p>The results of a query must be sortable by the following columns (when they are present):</p> <ul style="list-style-type: none"> • Extension list. • PUID • Name. • Release date. • Vendor name. • Read format invariance. • The date the format is supported until. <p>Source: [Current system, P3SD - S2.3.1.1, Clarification 1, Clarification 2, Review] Source: [Pronom6 SRD – S6.1.16.3, Pronom7 SRD – S7.1.2.2.5] N.B. The release date is currently not sortable in the search by software packages capable of reading/writing a format search results.</p>	M
S6.2.15.4	<p>When the sort criteria does not distinguish the order of two or more rows in a query result, the rows must be further ordered by release date and then by system ID.</p> <p>Source: [Review] Source: [Pronom6 SRD – S6.1.16.3.1, Pronom7 SRD – S7.1.2.2.6]</p>	M
S6.2.15.5	<p>For a search on file formats by extension, or name, the results should always include secondary ordering by format name and then by format version.</p> <p>Source: [DemoProgress] Source: [Pronom6 SRD – S6.1.16.3.2, Pronom7 SRD – S7.1.2.4.1.4]</p>	M
S6.2.15.6	<p>For a search on software, the results should include secondary ordering by software name, then version.</p> <p>Source: [DemoProgress] Source: [Pronom6 SRD – S6.1.16.3.3, Pronom7 SRD – S7.1.2.5.1.4]</p>	M
S6.2.15.7	<p>Search results must be shown in a paged format, with an indicator of how many pages there are in total.</p> <p>Source: [P3SD - S2.3.1.2] Source: [Pronom6 SRD – S6.1.16.4, Pronom7 SRD – S7.1.2.2.7]</p>	M
S6.2.15.8	<p>A facility must be provided to save each set of search results as XML.</p> <p>Source: [P3SD - S2.3.1.8, P4URD - 19, P4URD – 32, P4URD - 35] Source: [Pronom6 SRD – S6.1.16.5, Pronom7 SRD – S7.1.2.2.8]</p>	M
S6.2.15.9	<p>A facility must be provided to save each set of search results as CSV.</p> <p>Source: [P3SD - S2.3.1.8, P4URD – 19, P4URD – 32, P4URD – 35] Source: [Pronom6 SRD – S6.1.16.6, Pronom7 SRD – S7.1.2.2.9]</p>	M
S6.2.15.10	<p>A facility must be provided to view each set of search results in a 'printer friendly' format.</p> <p>Source: [P3SD - S2.3.1.8, P4URD - 19, P4URD - 32, P4URD - 35] Source: [Pronom6 SRD – S6.1.16.7, Pronom7 SRD – S7.1.2.2.10]</p>	M

6.2.16 General detailed reports

Label	Requirement	Necessity
S6.2.16.1	<p>It must be possible to provide a hyperlink in any of the detailed reports to an accompanying web page (e.g., containing supporting RTF or PDF documentation) with any valid URL.</p> <p>Source: [P4URD – 59, Clarification 2] Source: [Pronom6 SRD – S6.1.17.1, Pronom7 SRD – S7.1.2.2.2]</p>	M
S6.2.16.2	<p>A facility must be provided to save each report as XML.</p> <p>Source: [P3SD - S2.4.3.1, P3SD - S2.3.1.8, P4URD – 13, P4URD - 19, P4URD – 32, P4URD – 35, Clarification 1] Source: [Pronom6 SRD – S6.1.17.2, Pronom7 SRD – S7.1.2.3.1]</p>	M

Label	Requirement	Necessity
S6.2.16.3	A facility must be provided to save each report as CSV. Source: [P3SD - S2.3.1.8, P4URD – 13, P4URD – 19, P4URD – 32, P4URD – 35] Source: [Pronom6 SRD – S6.1.17.3, Pronom7 SRD – S7.1.2.3.2]	M
S6.2.16.4	A facility must be provided to view each report in a 'printer friendly' format. Source: [P3SD - S2.3.1.8, P4URD - 19, P4URD - 32, P4URD - 35] Source: [Pronom6 SRD – S6.1.17.4, Pronom7 SRD – S7.1.2.3.3]	M
S6.2.16.5	Every report type must have an associated XML schema. Source: [P3SD – S2.4.3.1, Clarification 2] Source: [Pronom6 SRD – S6.1.17.5, Pronom7 SRD – S7.1.2.3.4]	M
S6.2.16.6	Any provenance entry in a detailed report must contain a link to the detailed report for the source (agent), together with the source date and description. Source: [Review, 18-Feb-2005] Source: [Pronom6 SRD – S6.1.17.6, Pronom7 SRD – S7.1.2.2.1]	M

6.2.17 Migration Pathways Search

Label	Requirement	Necessity
S6.2.17.1	The user interface must allow users to search for migration pathways by specifying an optional input (start) format type and an optional output (end) format type. The resulting report should show a list of migration pathways. Source: [TWSRD, PA-tool registry – STU5, PA-tool registry – STU6, PA-tool registry – STU7] Source: [Pronom6 SRD – S6.1.18.1, Pronom7 SRD – S7.1.2.6.1.1, Pronom7 SRD – S7.1.2.6.1.2]	M
S6.2.17.2	Future enhancement, so removed from this document.	X
S6.2.17.3	When listing the migration pathways, the search results must show the input file format and output file format, with each result in the list being a link to the detailed report for that migration pathway. Source: [PA-tool registry – STU5, PA-tool registry – STU6, PA-tool registry – STU7] Source: [Pronom7 SRD – S7.1.2.6.1.5, Pronom7 SRD – S7.1.2.6.1.6]	M
S6.2.17.4	Future enhancement, so removed from this document.	X
S6.2.17.5	Future enhancement, so removed from this document.	X
S6.2.17.6	The detailed migration pathway report must include the following information: <ul style="list-style-type: none"> • Status of pathway • Name • Type of action • Type of tools in the pathway • Steps in the pathway • Input file format • Maximum file size for the input file • Output file format (for file format migrations) Source: [PA-tool registry – STU8] Source: [Pronom7 SRD – S7.1.2.6.3.1, Pronom7 SRD – S7.1.2.6.3.2, Pronom7 SRD – S7.1.2.6.3.4]	M
S6.2.17.7	Future enhancement, so removed from this document.	X
S6.2.17.8	Future enhancement, so removed from this document.	X

Label	Requirement	Necessity
S6.2.17.9	The detailed migration pathway report must include the following information about each step in the pathway: <ul style="list-style-type: none"> • Software package name • Software package tool name • Software package tool parameters • Migration pathway step number Source: [PA-tool registry – STU4, PA-tool registry – STU8] Source: [Pronom7 SRD – S7.1.2.6.3.5]	M
S6.2.17.10	Future enhancement, so removed from this document.	X

6.3 File Format Comparison

In the future it should be possible to compare two file formats.

6.4 Help system Requirements

Label	Requirement	Necessity
S6.4.1	The Help system must be a set of HTML pages, one for each search and report screen. Source: [P4URD - 36, 8-Apr-2005] Source: [Pronom6 SRD – S7.2.1.1, Pronom7 SRD – S7.1.6.1]	M
S6.4.2	The Help system must contain context sensitive information on each search and report screen. Source: [P4URD - 37] Source: [Pronom6 SRD – S7.2.1.2, Pronom7 SRD – S7.1.6.2]	M
S6.4.3	The help system text must be updated to cover all new functionality. Source: [Clarification 2, Review] Source: [Pronom6 SRD – S7.2.1.3, Pronom7 SRD – S7.1.6.3]	M
S6.4.4	The help system must include on the File Format summary web page an explanatory note as to how the Format Risk field should be interpreted. Source: [PC3/D4 – 2.8] Source: [Pronom7 SRD – S7.1.6.4]	M

6.5 Non-functional Requirements

Label	Requirement	Necessity
S6.5.1	The system should respond to 95% of queries within 1 second when there are 100 concurrent users or fewer. Source: [TWSRD] Source: [Pronom6 SRD – S6.2.9, Pronom7 SRD – S7.1.7.3]	D

7. Administration Requirements

This section details the requirements for the web based administration interface for the system. In addition to providing functionality to maintain the data in the repository, the administration interface also provides functionality to search and browse the data for users with the REPOSITORY_READER role.

7.1 General Requirements

Label	Requirement	Necessity
S7.1.1	The administration user interface must be web-browser-based. Source: [PC3/4] Source: [Pronom7 SRD – S7.2.5.4.1]	M
S7.1.2	The master system should be able to support 5 concurrent users. Source: [Clarification 1, Clarification 2] Source: [Pronom6 SRD – S7.1.2.1, Pronom7 SRD – S7.2.5.4.2]. N.B. This requirement is a statement of intent of the level of use the administration application will be designed and implemented to support. The requirement has been marked as desirable as we do not intend to test it due to the cost involved.	D

7.2 Security

This section details the security requirements for the Administration Interface.

The user requirements, on which these system requirements are based, do not have any explicit security requirements stated. In order to support the Audit functionality, it is necessary that users of the system authenticate themselves before making any changes to the data in the system. Without authentication there is no way to provide the details of the creator/editor of a record, unless the users manually enter that information when they edit information.

As such, the following requirements are included to provide a basic level of security and also to support the generation of audit entries when the contents of the database are modified.

Label	Requirement	Necessity
S7.2.1	The system must require a user to authenticate in order to edit data through the administration interface. Source: [Tessella] Source: [Pronom7 SRD – S7.2.1.1]	M
S7.2.2	The system must allow the user to login. Source: [Tessella] Source: [Pronom7 SRD – S7.2.1.2]	M
S7.2.3	The system must allow the user to logout. Source: [Tessella] Source: [Pronom7 SRD – S7.2.1.3]	M
S7.2.4	Future enhancement, so removed from this document.	X
S7.2.5	Future enhancement, so removed from this document.	X
S7.2.6	Future enhancement, so removed from this document.	X
S7.2.7	Users of the administration interface who have not authenticated themselves will be granted the REPOSITORY_READER level of authorisation. Source: [Tessella, MEETING-26-08-2009] Source: [Pronom7 SRD – S7.2.1.11]	M
S7.2.8	Authenticated users of the administration interface will be granted the REPOSITORY_SYSTEM_ADMINISTRATOR level of authorisation. Source: [Tessella] Source: [Pronom7 SRD – S7.2.1.7, Pronom7 SRD – S7.2.1.8]	M
S7.2.9	Future enhancement, so removed from this document.	X

7.3 Repository Reader Functionality

This section covers the functionality provided to a user with the role REPOSITORY_READER.

7.3.1 Core Entities

This section covers the functionality available to a user with role REPOSITORY_READER which relate to Core Entities.

7.3.1.1 Common

Label	Requirement	Necessity
	Common – List	
S7.3.1.1.1	The system must allow users to list all entries for each of the core entities. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.1]	M
S7.3.1.1.2	When listing the core entities, the system must show: <ul style="list-style-type: none"> • PUID • name • version text • has been withdrawn flag Source: [Tessella, Review] Source: [Pronom6 SRD - S7.1.1.12, Pronom7 SRD – S7.2.2.1.2, Pronom7 SRD – S7.2.2.1.3]	M
	Common – View	
S7.3.1.1.3	The system must allow users to view an individual core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.7]	M
S7.3.1.1.4	The detailed view of an individual core entity must include the following information: <ul style="list-style-type: none"> • name • version text • release date • has been withdrawn flag • withdrawal date Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.8]	M
S7.3.1.1.5	The detailed view of an individual core entity must include the following provenance information: <ul style="list-style-type: none"> • source name (i.e. name of agent) • source date • source information last updated date • provenance note Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.9, Pronom7 SRD – S7.2.2.2.10.2]	M
S7.3.1.1.6	The detailed view of an individual core entity must include the following information for each GDFR value associated with the core entity: <ul style="list-style-type: none"> • GDFR facet name • GDFR facet value Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.10, Pronom7 SRD – S7.2.2.1.11, Pronom7 SRD – S7.2.2.1.12]	M
S7.3.1.1.7	The detailed view of an individual core entity must include the following information for each external identifier associated with the core entity: <ul style="list-style-type: none"> • Identifier type • Identifier text Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.13]	M

Label	Requirement	Necessity
S7.3.1.1.8	The detailed view of an individual core entity must include the following information for each alias associated with the core entity: <ul style="list-style-type: none"> Alias name Alias version Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.14]	M
S7.3.1.1.9	The detailed view of an individual core entity must include a list of documents associated with the core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.15]	M
S7.3.1.1.10	The detailed view of an individual core entity must include a list of intellectual property rights associated with the core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.16]	M
S7.3.1.1.11	The detailed view of an individual core entity must include a list of agents associated with the core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.17]	M
S7.3.1.1.12	The detailed view of an individual core entity must include a list of external identifiers associated with the core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.18]	M
S7.3.1.1.13	The detailed view of an individual core entity must include a list of all other core entities that are related to the core entity being displayed giving the name and version of the related core entity together with the nature of the relationship. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.19, Pronom7 SRD – S7.2.2.1.20]	M
S7.3.1.1.14	Future enhancement, so removed from this document.	X

7.3.1.2 [File Formats](#)

Label	Requirement	Necessity
S7.3.1.2.1	The system must allow users to list file formats in such a way as to meet all the common requirements for listing core entities. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.2.1]	M
S7.3.1.2.2	The system must allow users to view an individual file format in such a way as to meet all the common requirements for viewing an individual core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.2.4]	M
S7.3.1.2.3	The detailed view of an individual file format must include the following information with the file format: <ul style="list-style-type: none"> file format orientation (binary/text/unknown) byte order (big-endian / little-endian / either / unknown) public disclosure level (full, partial, none, unknown). Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.2.5]	M
S7.3.1.2.4	The detailed view of an individual file format must include the following information for each external signature associated with the file format: <ul style="list-style-type: none"> External signature type External signature text Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.2.6]	M

Label	Requirement	Necessity
S7.3.1.2.5	The detailed view of an individual file format must include the following information for each reference file associated with the file format: <ul style="list-style-type: none"> • name • version Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.2.7]	M
S7.3.1.2.6	The detailed view of an individual file format must include the following information for each technical environment associated with the file format: <ul style="list-style-type: none"> • PUID • name Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.2.8, Pronom7 SRD – S7.2.2.1.2.9]	M
S7.3.1.2.7	The detailed view of an individual file format must include the following information for each XCEL file associated with the file format: <ul style="list-style-type: none"> • XCEL file name Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.2.10]	M
S7.3.1.2.8	The detailed view of an individual file format must include the following information for each assessment associated with the file format: <ul style="list-style-type: none"> • descriptive summary field • external website link text Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.2.11]	M
S7.3.1.2.9	The detailed view of an individual file format must include the following information for each component manifestation type associated with the file format: <ul style="list-style-type: none"> • component manifestation type name • manifestation type • description of the component manifestation type Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.2.12]	M
S7.3.1.2.10	The detailed view of an individual file format must include the following information for each property associated with the file format: <ul style="list-style-type: none"> • name • description (instance properties only) • is high risk flag • risk score Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.2.13, Pronom7 SRD – S7.2.2.1.2.14]	M
S7.3.1.2.11	The detailed view of an individual file format must include the following information for each internal signature associated with the file format: <ul style="list-style-type: none"> • Name • Flag to indicate if the internal signature is specific to only one file format, or generic (i.e. relates to more than one file format). Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.2.15]	M
S7.3.1.2.12	The detailed view of an individual file format must include the following information for each process associated with the file format: <ul style="list-style-type: none"> • Process type • Software tool name • Software tool PUID Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.2.16]	M
S7.3.1.2.13	Future enhancement, so removed from this document.	X

7.3.1.3 [Software Packages](#)

Label	Requirement	Necessity
-------	-------------	-----------

Label	Requirement	Necessity
S7.3.1.3.1	The system must allow users to list all software packages in such a way as to meet all the common requirements for listing core entities. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.4.1]	M
S7.3.1.3.2	The system must allow users to view a software package in such a way as to meet all the common requirements for viewing an individual core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.4.4]	M
S7.3.1.3.3	The detailed view of an individual core entity must include the following information for the software package: <ul style="list-style-type: none"> • service pack level • default file format (name and version text) should be indicated in the list of target relationships • availability of source Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.4.5]	M
S7.3.1.3.4	The detailed view of an individual software package must include the name of each language associated with the software package. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.4.6]	M
S7.3.1.3.5	The detailed view of an individual software package must include the display name of each image associated with the software package. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.4.7]	M
S7.3.1.3.6	The detailed view of an individual software package must include the name of each programming language associated with the software package. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.4.8]	M
S7.3.1.3.7	The detailed view of an individual software package must include the following information for each software package component associated with the software package: <ul style="list-style-type: none"> • name • version • description Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.4.9]	M
S7.3.1.3.8	The detailed view of an individual software package must include the PUID of each technical environment associated with the software package. The detailed view of an individual software package must include the following information for each technical environment associated with the software package: <ul style="list-style-type: none"> • PUID • name Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.4.10]	M
S7.3.1.3.9	The detailed view of an individual software package must include the following information for each software tool associated with the software package: <ul style="list-style-type: none"> • name • software tool Service Type Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.4.12]	M

7.3.1.4 [Hardware](#)

Label	Requirement	Necessity
-------	-------------	-----------

Label	Requirement	Necessity
S7.3.1.4.1	The system must allow users to list all hardware in such a way as to meet all the common requirements for listing core entities. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.3.1]	M
S7.3.1.4.2	The system must allow users to view an individual item of hardware in such a way as to meet all the common requirements for viewing an individual core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.3.2]	M
S7.3.1.4.3	The detailed view of an individual item of hardware must include the following information for each hardware component associated with the item of hardware: <ul style="list-style-type: none"> • name • version • description Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.3.3]	M
S7.3.1.4.4	The detailed view of an individual item of hardware must include the display name of any image associated with the item of hardware. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.3.4]	M

7.3.1.5 [Character Encoding](#)

Label	Requirement	Necessity
S7.3.1.5.1	The system must allow users to list all character encodings in such a way as to meet all the common requirements for listing core entities. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.5.1]	M
S7.3.1.5.2	The system must allow users to view an individual character encoding in such a way as to meet all the common requirements for viewing an individual core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.5.3]	M
S7.3.1.5.3	The detailed view of an individual character encoding must include the following information: <ul style="list-style-type: none"> • code unit width • encoding form width Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.5.4]	M

7.3.1.6 [Compression Technique](#)

Label	Requirement	Necessity
S7.3.1.6.1	The system must allow users to list all compression techniques in such a way as to meet all the common requirements for listing core entities. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.6.1]	M
S7.3.1.6.2	The system must allow users to view an individual compression technique in such a way as to meet all the common requirements for viewing an individual core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.6.3]	M
S7.3.1.6.3	The detailed view of an individual compression technique must include the lossiness of that compression technique. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.6.4]	M

7.3.1.7 [Storage Media](#)

Note: the storage of detailed storage media data is an enhancement.

Label	Requirement	Necessity
S7.3.1.7.1	The system must allow users to list all storage media in such a way as to meet all the common requirements for listing core entities. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.7.1]	M
S7.3.1.7.2	The system must allow users to view an individual storage medium in such a way as to meet all the common requirements for viewing an individual core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.1.7.2]	M

7.3.2 Subsidiary Entities

This section covers the functionality available to a user with role REPOSITORY_READER which relate to subsidiary entities.

7.3.2.1 Common

Label	Requirement	Necessity
	Common – List	
S7.3.2.1.1	The system must allow users to list all entries for each of the subsidiary entities. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.1.1]	M
S7.3.2.1.2	When listing subsidiary entities, the system must show each entity's PUID, if the subsidiary entity is allocated a PUID. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.1.2] N.B. Applies to agents, technical environments, pathways and properties.	M
	Common – View	
S7.3.2.1.3	The system must allow users to view an individual subsidiary entity in detail, including its PUID if it has one. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.1.4, Pronom7 SRD – S7.2.2.2.1.5]	M
S7.3.2.1.4	The detailed view of an individual subsidiary entity must include the following provenance information: <ul style="list-style-type: none"> • source name (i.e. name of agent) • source date • source information last updated date • provenance note Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.1.6, Pronom7 SRD – S7.2.2.2.10.2]	M
S7.3.2.1.5	The detailed view of an individual subsidiary entity must include the agent's role for each agent associated with the subsidiary entity: Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.1.7] N.B. This applies to IPR, documents, and agents.	M

7.3.2.2 Agents

Label	Requirement	Necessity
S7.3.2.2.1	When listing agents, the system must display the name of the agent. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.2.1]	M

Label	Requirement	Necessity
S7.3.2.2.2	The detailed view of an agent must include the following information: <ul style="list-style-type: none"> • Agent type (organisation/individual/institute/NGO etc) • Name • Address • Telephone • Email • Country Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.2.3]	M
S7.3.2.2.3	The detailed view of an agent must include the URL of each web site associated with the agent: Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.2.4]	M
S7.3.2.2.4	For agents who are individuals (rather than organisations), the detailed view of an agent must include the following information: <ul style="list-style-type: none"> • job title • organisation name Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.2.5]	M
S7.3.2.2.5	The detailed view of an agent must list all other agents associated with the agent being viewed. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.2.6]	M

7.3.2.3 [Intellectual Property Rights](#)

Label	Requirement	Necessity
S7.3.2.3.1	The detailed view of an individual IPR entry must include the following information: <ul style="list-style-type: none"> • Intellectual property right type • Intellectual property right jurisdiction type • licence details • description Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.3.2]	M
S7.3.2.3.2	The detailed view of an individual IPR entry must list all external identifiers associated with that IPR entry. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.1.8]	M

7.3.2.4 [Documents](#)

Label	Requirement	Necessity
S7.3.2.4.1	The detailed view of an individual document must include the following information: <ul style="list-style-type: none"> • title • bibliographical reference for citation • publication date • document type • document availability type • document content type (user guide, installation guide etc) Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.4.3]	M
S7.3.2.4.2	The detailed view of an individual document must list all IPR entries associated with that document. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.4.4]	M

Label	Requirement	Necessity
S7.3.2.4.3	The detailed view of an individual document must list all other documents related to that document, including a description of the relationship. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.4.5]	M

7.3.2.5 [Technical Environments](#)

Label	Requirement	Necessity
S7.3.2.5.1	When listing technical environments, the system must display the following information: <ul style="list-style-type: none"> • name • description Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.5.1]	M
S7.3.2.5.2	The detailed view of an individual technical environment must include the following information: <ul style="list-style-type: none"> • name • description Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.5.2]	M
S7.3.2.5.3	The detailed view of an individual technical environment must list all hardware associated with that technical environment. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.5.3]	M
S7.3.2.5.4	The detailed view of an individual technical environment must list all software packages associated with that technical environment. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.5.4]	M
S7.3.2.5.5	The detailed view of an individual technical environment must list all documents associated with that technical environment. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.1.9]	M

7.3.2.6 [Processes](#)

Label	Requirement	Necessity
S7.3.2.6.1	When listing processes, the system must display the following information: <ul style="list-style-type: none"> • type • object Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.6.1]	M
S7.3.2.6.2	The detailed view of an individual process must include the following information: <ul style="list-style-type: none"> • Process type (by name) • Software package name • Software package tool name (if there are invocation details for the process) • known limitations • note text • file format name (only processes for objects) • technical environment being emulated (only for emulation processes) • target file format (only for migration processes) • property name (only for property extraction processes) Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.6.2, Pronom7 SRD – S7.2.2.2.6.3, Pronom7 SRD – S7.2.2.2.6.5, Pronom7 SRD – S7.2.2.2.6.6, Pronom7 SRD – S7.2.2.2.6.7]	M

7.3.2.7 Pathways

Label	Requirement	Necessity
S7.3.2.7.1	When listing pathways, the system must display the following information: <ul style="list-style-type: none"> • name • description Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.7.1]	M
S7.3.2.7.2	The detailed view of an individual pathway must include the following information: <ul style="list-style-type: none"> • name • description • pathway type (file format migration, technical environment emulation, file format object extraction) • list of steps in the pathway Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.7.2, Pronom7 SRD – S7.2.2.2.7.3]	M
S7.3.2.7.3	The detailed view of an individual pathway must include the following information for each role associated with the pathway: <ul style="list-style-type: none"> • pathway role (Preservation, Presentation, Extraction) • pathway role status (current, withdrawn, unknown) Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.7.4]	M
S7.3.2.7.4	When listing pathway steps, the system must display the following information: <ul style="list-style-type: none"> • name • description • sequence number Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.7.5]	M
S7.3.2.7.5	The detailed view of an individual pathway step must include the following information: <ul style="list-style-type: none"> • name • description • sequence number Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.7.6]	M
S7.3.2.7.6	The detailed view of an individual pathway step that is an OBJECT_MIGRATION must include the following information: <ul style="list-style-type: none"> • source file format PUID • source file format name • source file format version • target file format PUID • target file format name • target file format version • maximum file size Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.7.8]	M

Label	Requirement	Necessity
S7.3.2.7.7	<p>The detailed view of an individual pathway step that is an OBJECT_MIGRATION must include the following information for each property associated with the migration pathway step:</p> <ul style="list-style-type: none"> • source instance property PUID • source instance property Name • target instance property PUID • target instance property Name • variance • notes <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.7.9]</p>	M

7.3.2.8 [Properties](#)

Label	Requirement	Necessity
S7.3.2.8.1	<p>When listing properties, the system must display the following information:</p> <ul style="list-style-type: none"> • name • description <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.8.1]</p>	M
S7.3.2.8.2	<p>The detailed view of an individual property must include the following information:</p> <ul style="list-style-type: none"> • name • description • property type (instance/inherent) <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.8.2]</p>	M

7.3.2.9 [Component Manifestation Types](#)

Label	Requirement	Necessity
S7.3.2.9.1	<p>When listing component manifestation types, the system must display the following information:</p> <ul style="list-style-type: none"> • name • description • a manifestation type (image/audio/document etc) <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.9.1]</p>	M
S7.3.2.9.2	<p>The detailed view of a component manifestation type must include the following information:</p> <ul style="list-style-type: none"> • name • description • a manifestation type (image/audio/document etc) <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.9.2]</p>	M

7.3.2.10 [External Identifiers](#)

Label	Requirement	Necessity
S7.3.2.10.1	<p>When listing external identifiers, the system must display the following information:</p> <ul style="list-style-type: none"> • external identifier text (such as a url) • external identifier type name (e.g. mime-type, url, ISBN) <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.11.1]</p>	M

Label	Requirement	Necessity
S7.3.2.10.2	<p>The detailed view of an external identifier must include the following information:</p> <ul style="list-style-type: none"> external identifier text (such as a url) external identifier type name (e.g. mime-type, url, ISBN) external identifier note <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.11.2]</p>	M

7.3.2.11 [Images](#)

Label	Requirement	Necessity
S7.3.2.11.1	<p>When listing images, the system must display the following information:</p> <ul style="list-style-type: none"> name description <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.12.1]</p>	M
S7.3.2.11.2	<p>The detailed view of an image must include the following information:</p> <ul style="list-style-type: none"> name description some way to view the image binary <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.12.2, Pronom7 SRD – S7.2.2.2.12.3]</p>	M

7.3.2.12 [External Signatures](#)

Label	Requirement	Necessity
S7.3.2.12.1	<p>When listing external signatures, the system must display the following information:</p> <ul style="list-style-type: none"> External signature type (File extension, Mac OS data fork etc) External signature value <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.13.1]</p>	M
S7.3.2.12.2	<p>The detailed view of an individual external signature must include the following information:</p> <ul style="list-style-type: none"> External signature type (File extension, Mac OS data fork etc) External signature value External signature note <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.13.2]</p>	M

7.3.2.13 [Internal Signatures](#)

Label	Requirement	Necessity
S7.3.2.13.1	<p>When listing internal signatures, the system must display the following information:</p> <ul style="list-style-type: none"> Signature name Whether the signature is generic or specific <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.14.1]</p>	M
S7.3.2.13.2	<p>The detailed view of an individual internal signature must include the following information:</p> <ul style="list-style-type: none"> Signature name Whether the signature is generic or specific Note <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.14.2]</p>	M

Label	Requirement	Necessity
S7.3.2.13.3	<p>The detailed view of an individual internal signature must include the following information (in the order given below) for each byte sequence associated with the internal signature:</p> <ul style="list-style-type: none"> • byte sequence position name (Absolute from BOF, Absolute from EOF, Variable, Indirect From BOF, Indirect From EOF, Unknown) • offset • max offset • indirect offset location • indirect offset length • value (i.e. the byte sequence itself) • byte order (Little-endian, Big-endian, Big-endian or little-endian) <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.14.3]</p>	M

7.3.2.14 Reference Files

Label	Requirement	Necessity
S7.3.2.14.1	<p>When listing reference files, the system must display the following information:</p> <ul style="list-style-type: none"> • file name • file version <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.15.1]</p>	M
S7.3.2.14.2	<p>The detailed view of an individual reference file must include the following information:</p> <ul style="list-style-type: none"> • file name • file version • file description • file notes <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.15.2]</p>	M
S7.3.2.14.3	<p>The detailed view of an individual reference file must list all IPR entries associated with the reference file.</p> <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.2.15.3]</p>	M

7.3.3 Policy Information

7.3.3.1 Pathway Policy Information

Label	Requirement	Necessity
S7.3.3.1.1	<p>The detailed view of an individual pathway must include the following information for the approval of each role associated with the pathway:</p> <ul style="list-style-type: none"> • approval date • the name of the agent who approved the pathway for the particular role <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.3.1.1]</p>	M
S7.3.3.1.2	<p>The detailed view of an individual pathway step must include the following information for the approval of each migration property associated with the pathway step:</p> <ul style="list-style-type: none"> • approval date • the name of the Agent who approved the migration property for the particular pathway step <p>Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.3.1.2]</p>	M

7.3.3.2 Process Policy Information

Label	Requirement	Necessity
-------	-------------	-----------

Label	Requirement	Necessity
S7.3.3.2.1	The detailed view of an individual process must include any process priority values associated with that process. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.3.2.1]	M

7.3.3.3 File Format Risk Policy Information

Label	Requirement	Necessity
S7.3.3.3.1	The detailed view of an individual file format must include the following information for any inherent property risks associated with the file format: <ul style="list-style-type: none"> • Inherent property selected values <ul style="list-style-type: none"> ○ value ○ risk score ○ is high risk flag • Total inherent risk score for file format Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.3.3.1]	M
S7.3.3.3.2	The detailed view of an individual file format must include the following information for any instance property risks associated with the file format: <ul style="list-style-type: none"> • Risk score • Whether it's flagged as being high risk and for instance properties which have a set of discrete values: <ul style="list-style-type: none"> • Instance property value and for instance properties which have continuous values divided into a set of ranges: <ul style="list-style-type: none"> • Instance property range min • Instance property range max Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.3.3.2, Pronom7 SRD – S7.2.2.3.3.3]	M

7.3.4 Search Functionality

In addition to the main search functionality in the .NET legacy, the Java administration and general public interface also includes some search functionality, which is detailed in the following requirements.

Label	Requirement	Necessity
S7.3.4.1	It must be possible for a user to search for an entity by its PUID. Source: [PCR 2 implementation]	M
S7.3.4.2	It must be possible for a user to search for a core entity or a subsidiary entity by keyword when that keyword is in the entity's name or description. Source: [PCR 2 implementation]	M
S7.3.4.3	It must be possible for a user to restrict the keyword search to a specific type of core entity (e.g. file formats). Source: [PCR 2 implementation]	M
S7.3.4.4	It must be possible for a user to search for a pathway by specifying the PUIDs of the input (start) and output (end) file formats. This search should return all pathways that start from the specified file format and end at the specified file format. Source: [PCR 2 implementation]	M
S7.3.4.5	It must be possible for a user to search for pathways by specifying the PUID of the input (start) file format. This search should return all pathways that start from the given file format. Source: [PCR 2 implementation]	M
S7.3.4.6	It must be possible for a user to search for pathways by specifying the PUID of the output (end) file format. This search should return all pathways that result in the given file format. Source: [PCR 2 implementation]	M

Label	Requirement	Necessity
S7.3.4.7	It must be possible for a user to search for pathways by specifying the PUID of the target (end) technical environment. This search should return all pathways that emulate the given technical environment. Source: [PCR 2 implementation]	M

7.4 Repository Administrator Functionality

This section details the requirements for a user with the role REPOSITORY_ADMINISTRATOR (described as 'an administrator' in the requirements in this section).

Users with the REPOSITORY_ADMINISTRATOR role can carry out all functionality available to a REPOSITORY_READER.

7.4.1 Core Entities

Label	Requirement	Necessity
S7.4.1.1	It must be possible for an administrator to create a new core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.1.1]	M
S7.4.1.2	It must be possible for an administrator to edit an existing core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.1.5, Pronom7 SRD – S7.2.3.1.9, Pronom7 SRD – S7.2.4.2.1]	M
S7.4.1.3	When viewing a list of core entities, it must be possible for an administrator to edit an existing core entity or create a new one. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.1.2, Pronom7 SRD – S7.2.3.1.3]	M
S7.4.1.4	When viewing or editing an individual core entity, it must be possible for an administrator to delete that core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.1.6, Pronom7 SRD – S7.2.3.1.8]	M
S7.4.1.5	When editing an individual core entity, the system must indicate that it is in edit mode. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.1.7]	M
S7.4.1.6	When editing an existing core entity, it must be possible for an administrator to add or remove items to/from the displayed lists of associated entities. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.1.10, Pronom7 SRD – S7.2.4.2.2, Pronom7 SRD – S7.2.3.1.11, Pronom7 SRD – S7.2.4.2.3, Pronom7 SRD – S7.2.3.1.12, Pronom7 SRD – S7.2.4.2.4]	M
S7.4.1.7	When editing an existing core entity, the system must provide an interface to edit associated items in the displayed lists when those associated items do not have an independent existence outside of the core entity being edited. Source: [Tessella]	M
S7.4.1.8	When editing an existing core entity, it must be possible to save all changes to the individual core entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.1.13]	M
S7.4.1.9	When editing an existing core entity, it must be possible to abandon all changes to the individual core entity (e.g. by navigating away from the page without saving the changes). Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.1.14]	M

7.4.2 Subsidiary Entities

Label	Requirement	Necessity
-------	-------------	-----------

Label	Requirement	Necessity
S7.4.2.1	It must be possible for an administrator to create a new subsidiary entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.2.1]	M
S7.4.2.2	It must be possible for an administrator to edit an existing subsidiary entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.2.2, Pronom7 SRD – S7.2.3.3.4, Pronom7 SRD – S7.2.3.3.8, Pronom7 SRD – S7.2.4.3.1]	M
S7.4.2.3	When viewing or editing an individual subsidiary entity, it must be possible for an administrator to delete that subsidiary entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.3.5, Pronom7 SRD – S7.2.3.3.7]	M
S7.4.2.4	When editing an individual subsidiary entity, the system must indicate that it is in edit mode. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.3.6]	M
S7.4.2.5	When editing an existing subsidiary entity, it must be possible for an administrator to add or remove items to/from the displayed lists of associated entities. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.3.9, Pronom7 SRD – S7.2.4.3.2, Pronom7 SRD – S7.2.3.3.10, Pronom7 SRD – S7.2.4.3.3]	M
S7.4.2.6	When editing an individual subsidiary entity, the system must provide an interface to edit associated items in the displayed lists when those associated items do not have an independent existence outside of the subsidiary entity being edited. Source: [Tessella, Pronom7 SRD – S7.2.3.3.11]	M
S7.4.2.7	When editing an existing subsidiary entity, it must be possible to save all changes to the individual subsidiary entity. Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.3.12]	M
S7.4.2.8	When editing an existing subsidiary entity, it must be possible to abandon all changes to the individual subsidiary entity (e.g. by navigating away from the page without saving the changes). Source: [Tessella] Source: [Pronom7 SRD – S7.2.3.3.13]	M

7.4.3 Audit Trail

This section details the requirements for the display of audit information.

Label	Requirement	Necessity
S7.4.3.1	The system must be able to list all audit entries for any audited entity in the system. Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.4.1]	M
S7.4.3.2	When listing audit entries, the system must display the following information: <ul style="list-style-type: none"> • modification time • modification type • name/identifier of the person who modified the record Source: [Tessella] Source: [Pronom7 SRD – S7.2.2.4.2]	M
S7.4.3.3	Future enhancement, so removed from this document.	X
S7.4.3.4	Future enhancement, so removed from this document.	X
S7.4.3.5	Future enhancement, so removed from this document.	X
S7.4.3.6	Future enhancement, so removed from this document.	X
S7.4.3.7	Future enhancement, so removed from this document.	X
S7.4.3.8	Future enhancement, so removed from this document.	X

7.4.4 Lists of Values

This section details the requirements for maintenance of the lists of values used to restrict selections for particular fields in the system.

Within the system, the following are considered to be lists of values (LOV):

Agent Relationship Types	External Identifier Types
Agent Roles	External Signature Types
Agent Types	Flag States
Byte Sequence Positions	Format Byte Orders
Component Types	Format Disclosure Levels
Compression Lossiness Types	Intellectual Property Right Jurisdictions
Content Variance Types	Intellectual Property Right Types
Countries	Languages
Document Content Types	Pathway Roles
Document Types	Pathway States
Documentation Availabilities	Pathway Types
Entity Class Ext Id Types	Process Action Types
Entity Classes	Process Types
Entity Families	Programming Languages
Entity Relationship Types	PUID Types
Entity Types	Software Package Interface Types

They are characterised by the following:

- They are populated with values as part of the system installation.
- They are used to populate drop down lists of values.
- They are infrequently modified.
- They may be used to determine behaviour of the system in workflows.

Label	Requirement	Necessity
S7.4.4.1	The look-up tables in admin screen drop-down lists must appear in alphabetical order. Source: [TWSRD] Source: [Pronom6 SRD - S7.1.1.6, Pronom7 SRD – S7.2.2.5.1]	M
S7.4.4.2	Future enhancement, so removed from this document.	X
S7.4.4.3	Future enhancement, so removed from this document.	X
S7.4.4.4	Future enhancement, so removed from this document.	X
S7.4.4.5	Future enhancement, so removed from this document.	X
S7.4.4.6	Future enhancement, so removed from this document.	X
S7.4.4.7	Future enhancement, so removed from this document.	X
S7.4.4.8	Future enhancement, so removed from this document.	X
S7.4.4.9	Future enhancement, so removed from this document.	X
S7.4.4.10	Future enhancement, so removed from this document.	X
S7.4.4.11	Future enhancement, so removed from this document.	X
S7.4.4.12	Future enhancement, so removed from this document.	X

7.4.5 DROID Signature File Generation

This section covers the requirements relating to the generation of signature files for the file format identification tool, DROID, from the information about internal signatures stored in the Core Registry.

Label	Requirement	Necessity
S7.4.5.1	Administrators must be able to initiate an automatic process to create an XML signature file containing: <ul style="list-style-type: none"> • The date the file was created. 	M

Label	Requirement	Necessity
	<ul style="list-style-type: none"> The list of all known file formats with name, version, PUID. All external signatures. All internal signatures including all byte sequences. Source: [P4URD-54, Review] Source: [DROID3 – S3.6.1.2]	
S7.4.5.2	The XML signature file must conform to a TNA-approved XML schema. Source: [Review] Source: [DROID3 – S3.6.1.2.1]	M
S7.4.5.3	Each time this XML file is created, the file must be given a unique version number. Source: [Tessella] Source: [DROID3 – S3.6.1.3]	M
S7.4.5.4	It must be possible to configure the location where the signature file will be saved to. Source: [Feedback1] Source: [DROID3 – S3.6.1.4]	M
S7.4.5.5	When a new XML signature file is generated, the Core Registry record of which file formats can be identified by DROID should be updated to add or remove those entries relating to those file formats which have been added or removed or whose internal signature has been added or removed. This is so that the Core Registry/DROID combination will remain self-consistent. Source: [TWSRD] Source: [Pronom6 SRD – S7.1.1.10, Pronom7 SRD – S6.2.2] Note: This information is recorded in the Core Registry as a set of processes linked to the DROID Pronom 6.2 (SDB) software package tool.	M

7.4.6 Technology Watch Alerts

Label	Requirement	Necessity
S7.4.6.1	It must be possible for administrators to subscribe, either themselves or others, to receive the technology watch alerts. Source: [PC3/D6 – 2.3] Note: These are described in section 5.4.3.	M

7.4.7 Review Testbed Results

In the future it should be possible for an administrator to review submitted Testbed results and either accept or reject them.

7.4.8 PUID Range Assignment

This section covers requirements relating to PUID assignment.

Label	Requirement	Necessity
S7.4.8.1	Future enhancement, so removed from this document.	X

8. Interfaces for External Systems

This section covers the interface requirements that are not browser based. This covers web services based on SOAP or REST.

8.1 REST Interface

This section details the Representational State Transfer (REST) interface that the system must provide.

Label	Requirement	Necessity
S8.1.1	Future enhancement, so removed from this document.	X
S8.1.2	The Core Registry must be able to resolve PUID URIs to machine-readable XML. Source: [URD: REG-SI 02, MEET 4] Source: [Pronom6 SRD – S9.1.4, Pronom7 SRD – S7.3.1.2]	M

8.2 SOAP Web Services

This section details the SOAP web services that the system must provide.

8.2.1 Characterisation

This section details the SOAP web services that the system must provide as part of the system's support for characterisation.

Label	Requirement	Necessity
S8.2.1.1	For backward compatibility with Pronom 6.2, the system should implement the following web services <ul style="list-style-type: none"> • getIdentificationTools • getValidationTools • getPropertyExtractionTools • getToolProperties • getObjectExtractionTools • getSchema • getDTD in accordance with the wsdl found at http://www.nationalarchives.gov.uk/PRONOM/Services/Contract/PRONOMcharacterisation.wsdl . Source: [Pronom 6 implementation] Source: [Pronom6 SRD – S9.1.1, Pronom7 SRD – S7.4.1.2]	M
S8.2.1.2	The system must implement the following web service method <ul style="list-style-type: none"> • getIdentificationTools so that it lists all Software Tools that are associated with the IDENTIFY Software Tool Capability for one or more File Formats. Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.1.3]	M
S8.2.1.3	The system must implement the following web service method <ul style="list-style-type: none"> • getValidationTools so that it lists all Software Tools that are associated with the VALIDATE Software Tool Capability for a specific File Format. Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.1.4]	M

Label	Requirement	Necessity
S8.2.1.4	The system must implement the following web service method <ul style="list-style-type: none"> getPropertyExtractionTools so that it lists all Software Tools that are associated with the PROPERTY_EXTRACT Software Tool Capability for a specific File Format. Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.1.5]	M
S8.2.1.5	The system must implement the following web service method <ul style="list-style-type: none"> getToolProperties so that it lists all Software Tools that are associated with the PROPERTY_EXTRACT Software Tool Capability and the Properties that the Software Tool can extract. Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.1.6]	M
S8.2.1.6	The system must implement the following web service method <ul style="list-style-type: none"> getObjectExtractionTools so that it lists all Software Tools that are associated with the OBJECT_EXTRACT Software Tool. Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.1.7]	M
S8.2.1.7	The system must implement the following web service method <ul style="list-style-type: none"> getSchema so that it returns an xml schema for a specific Pronom related xml namespace and/or location. Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.1.8]	M
S8.2.1.8	The system must implement the following web service method <ul style="list-style-type: none"> getDTD so that it returns an xml DTD for a specific Pronom related xml DTD public ID or system ID. Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.1.9]	M

8.2.2 Preservation Planning

This section details the SOAP web services that the system must provide as part of the system's support for preservation planning.

Label	Requirement	Necessity
-------	-------------	-----------

Label	Requirement	Necessity
S8.2.2.1	<p>The system must, for backward compatibility with Pronom 6.2, implement the web services below in accordance with the wsdl with the namespace of http://pp.pronom.nationalarchives.gov.uk:</p> <ul style="list-style-type: none"> • getFormatsAtRisk • getFormatsByMigrationType • getFormatRisk • getFormatPropertyRisk • getMigrationPathways • getComponentIdentificationTools • getComponentMeasurementTools • getComponentProperties • setPathwayApproval • setPathwayCurrent <p>Source: [Pronom 6 implementation] Source: [Pronom6 SRD – S9.1.1, Pronom7 SRD – S7.4.2.2]</p>	M
S8.2.2.2	<p>The system must implement the following method</p> <ul style="list-style-type: none"> • getFormatsAtRisk <p>so that it returns a list of file formats with risk higher than given threshold.</p> <p>Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.2.3]</p>	M
S8.2.2.3	<p>The system must implement the following method</p> <ul style="list-style-type: none"> • getFormatsByMigrationType <p>so that it returns a list of file formats that can be migrated by a migration tool with a specified pathway type (PRESERVATION or EXTRACTION are expected)</p> <p>Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.2.4]</p>	M
S8.2.2.4	<p>The system must implement the following method</p> <ul style="list-style-type: none"> • getFormatRisk <p>so that it returns the format risk for a given file format</p> <p>Source: [Pronom 6 implementation, PC3/D4 – 2.9] Source: [Pronom7 SRD – S7.4.2.5, Pronom7 SRD – S7.4.2.1.6]</p>	M
S8.2.2.5	<p>The system must implement the following method</p> <ul style="list-style-type: none"> • getFormatPropertyRisk <p>so that it returns the file format inherent property risk for a given File Format Inherent Property value.</p> <p>Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.2.6]</p>	M
S8.2.2.6	<p>The system must implement the following method</p> <ul style="list-style-type: none"> • getComponentIdentificationTools <p>so that it returns a list of software tools that can be used to identify groups of files as individual components.</p> <p>Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.2.7]</p>	M
S8.2.2.7	<p>The system must implement the following method</p> <ul style="list-style-type: none"> • getComponentMeasurementTools <p>so that it returns a list of software tools that can be used to measure groups of files as individual components.</p> <p>Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.2.8]</p>	M

Label	Requirement	Necessity
S8.2.2.8	The system must implement the following method <ul style="list-style-type: none"> getComponentProperties so that it returns a list of properties that should be measured for a particular type of component. Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.2.9]	M
S8.2.2.9	The system must implement the following method <ul style="list-style-type: none"> setPathwayApproval so that sets the approval status for a specific migration pathway to approved. Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.2.10]	M
S8.2.2.10	The system must implement the following method <ul style="list-style-type: none"> setPathwayCurrent so that it sets a specific migration pathway as the default pathway for its source file format. Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.2.11]	M

8.2.2.1 [Preservation Planning – File Formats](#)

This section details the SOAP web services that the system must provide related to preservation planning and file formats.

Label	Requirement	Necessity
S8.2.2.1.1	The system must provide an interface to retrieve a list of properties for a specific file format. Source: [PC3/D4 – 2.1] Source: [Pronom7 SRD – S7.4.2.1.1] Note: This has been requested by Planets sub-project PP/4.	M
S8.2.2.1.2	Future enhancement, so removed from this document.	X
S8.2.2.1.3	Future enhancement, so removed from this document.	X
S8.2.2.1.4	The system must provide an interface to retrieve a list of file formats that are associated with a given file format property. Source: [PC3/D4 – 2.3, PC3/D4 – 2.14] Source: [Pronom7 SRD – S7.4.2.1.3, Pronom7 SRD – S7.4.2.1.7] Note: an example file format property would be image width.	M
S8.2.2.1.5	Future enhancement, so removed from this document.	X
S8.2.2.1.6	The system must provide an interface to retrieve an XCEL file for a specific file format. Source: [PC3/D4 – 2.5] Source: [Pronom7 SRD – S7.4.2.1.4]	M
S8.2.2.1.7	The system should provide an interface to retrieve an XCDL schema file for a specific file format. Source: [PC3/D4 – 2.5] Source: [Pronom7 SRD – S7.4.2.1.5]	D

8.2.2.2 [Preservation Planning – Software](#)

This section details the SOAP web services that the system must provide related to preservation planning and software packages.

Label	Requirement	Necessity
-------	-------------	-----------

Label	Requirement	Necessity
S8.2.2.2.1	<p>The system must provide an interface to retrieve a list of software package tools:</p> <ul style="list-style-type: none"> • getSoftwarePackages <p>Source: [PA-tool-registry – SU1] Source: [Pronom7 SRD – S7.4.2.2.1]</p>	M
S8.2.2.2.2	<p>The system must, for the interface to retrieve a list of software package tools associated with a given software package, allow the optional specification of the following:</p> <ul style="list-style-type: none"> • a software package ID (PUID) <p>Source: [PA-tool-registry – SU11] Source: [Pronom7 SRD – S7.4.2.2.2]</p>	M
S8.2.2.2.3	<p>The system must, for the interface to retrieve a list of software packages, allow the optional specification of the following if a software package ID has been specified:</p> <ul style="list-style-type: none"> • return all details flag <p>Source: [PA-tool-registry – SU11] Source: [Pronom7 SRD – S7.4.2.2.3]</p>	M
S8.2.2.2.4	<p>The system must, for the interface to retrieve a list of software package tools, allow the specification of one of the following:</p> <ul style="list-style-type: none"> • a source file format • a target technical environment <p>Source: [Pronom 6 implementation, PA-tool-registry – SU1, PA-tool-registry – SU3, PA-tool-registry – SU5] Source: [Pronom7 SRD – S7.4.2.2.4]</p>	M
S8.2.2.2.5	<p>The system must, for the interface to retrieve a list of software package tools, allow the following optional parameters:</p> <ul style="list-style-type: none"> • Type of package tool (such as Planets service or non-Planets service or application) <p>Source: [PA-tool-registry – SU1, PA-tool-registry – SU2, PA-tool-registry – SU3] Source: [Pronom7 SRD – S7.4.2.2.5]</p>	M
S8.2.2.2.6	<p>The system must, for the interface to retrieve a list of software package tools for a specific file format, allow the following optional parameters:</p> <ul style="list-style-type: none"> • Type of action (action on object, action on environment) <p>Source: [PA-tool-registry – SU2, PA-tool-registry – SU3] Source: [Pronom7 SRD – S7.4.2.2.6]</p>	M
S8.2.2.2.7	<p>The system must, in response to a request for software package tools for a specific file format, return:</p> <ul style="list-style-type: none"> • a list of software package descriptions for software packages that can • read the specific file format or • recreate a software package to read the specific file format. <p>Source: [PA-tool-registry – SU1] Source: [Pronom7 SRD – S7.4.2.2.7]</p>	M

Label	Requirement	Necessity
S8.2.2.2.8	<p>The system must, in response to a request for software package tools, for each software package tool description return the following information:</p> <ul style="list-style-type: none"> • ID of software package tool • Name of software package tool • Version of software package tool • Creator of software package • Description of software package tool • Type of software package tool (Planets service or non-Planets service) • Software and hardware dependencies for the software package. <p>Source: [PA-tool-registry – SU1, PA-tool-registry – SU5] Source: [Pronom7 SRD – S7.4.2.2.1]</p> <p>Note: the software and hardware dependencies might be specified as</p> <ul style="list-style-type: none"> • a reference to a specific defined technical environment • as a link to multiple software packages and hardware • as a combination of both a software package and software and hardware dependencies. 	M
S8.2.2.2.9	<p>The system must, in response to a request for software package tools, return, for each software package tool, the following additional information if a detailed response has been requested:</p> <ul style="list-style-type: none"> • list of associated documentation • list of publishers • release date • still supported flag • support withdrawn date (if known) • a list of file formats that can be read by the software package • a list of file formats that can be written by the software package (for actions on objects) • a list of target environments that the software package can emulate (for actions on environments). <p>Source: [PA-tool-registry – SU11] Source: [Pronom7 SRD – S7.4.2.2.2]</p> <p>Note: It may be more sensible to use the existing terminology for software package capabilities instead of read/write. e.g.: IDENTIFY VALIDATE PROPERTY_EXTRACT REDACT MIGRATE OBJECT_EXTRACT and EMULATE</p>	M
S8.2.2.2.10	<p>The system must, in response to a request for software packages, return, for each software package, the following additional information if a detailed response has been requested and the software package is a Planets service:</p> <ul style="list-style-type: none"> • invocation details <p>Source: [PA-tool-registry – SU11] Source: [Pronom7 SRD – S7.4.2.2.3]</p>	M

Label	Requirement	Necessity
S8.2.2.2.11	<p>The system must, in response to a request for software packages, return, for each software package, the following additional information if a detailed response has been requested and the software package is not a Planets service:</p> <ul style="list-style-type: none"> location details <p>Source: [PA-tool-registry – SU11] Source: [Pronom7 SRD – S7.4.2.2.4]</p>	M

8.2.2.3 [Preservation Planning – Migration Pathways](#)

This section details the SOAP web services that the system must provide related to preservation planning and Migration Pathways.

Label	Requirement	Necessity
S8.2.2.3.1	<p>The system must implement the following method</p> <ul style="list-style-type: none"> getMigrationPathways <p>so that it returns a list of migration pathways.</p> <p>Source: [Pronom 6 implementation, PA-tool registry – SU6] Source: [Pronom7 SRD – S7.4.2.3.1]</p>	M
S8.2.2.3.2	<p>The system must, for backward compatibility with Pronom 6.2, implement the web services below in accordance with the wsdl with the namespace of http://pp.pronom.nationalarchives.gov.uk:</p> <ul style="list-style-type: none"> getMigrationPathways <p>Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.2.3.2]</p>	M
S8.2.2.3.3	<p>The system must, for the interface to retrieve a list of migration pathways, allow the optional specification of the following:</p> <ul style="list-style-type: none"> a migration pathway ID <p>Source: [PA-tool registry – SU9, PA-tool registry – SU10] Source: [Pronom7 SRD – S7.4.2.3.3]</p>	M
S8.2.2.3.4	<p>The system must, for the interface to retrieve a list of migration pathways, allow the optional specification of the following if a migration pathway ID has been specified:</p> <ul style="list-style-type: none"> return all details flag <p>Source: [PA-tool registry – SU10] Source: [Pronom7 SRD – S7.4.2.3.4]</p>	M
S8.2.2.3.5	<p>The system must, for the interface to retrieve a list of migration pathways, allow the optional specification of one of the following:</p> <ul style="list-style-type: none"> a source file format a target technical environment <p>Source: [Pronom 6 implementation, PA-tool registry – SU7, PA-tool registry – SU6] Source: [Pronom7 SRD – S7.4.2.3.5]</p>	M
S8.2.2.3.6	<p>The system must, for the interface to retrieve a list of migration pathways, allow the optional specification of the following, if a target technical environment has not been specified:</p> <ul style="list-style-type: none"> a target file format <p>Source: [Pronom 6 implementation, PA-tool registry – SU8] Source: [Pronom7 SRD – S7.4.2.3.6]</p>	M

Label	Requirement	Necessity
S8.2.2.3.7	<p>The system must, for the interface to retrieve a list of migration pathways, allow the following optional parameters:</p> <ul style="list-style-type: none"> Type of tool (Planets service or non-Planets service or Application) <p>Source: [PA-tool registry – SU6] Source: [Pronom7 SRD – S7.4.2.3.7]</p>	M
S8.2.2.3.8	<p>The system must, for the interface to retrieve a list of migration pathways for a specific file format, allow the following optional parameters:</p> <ul style="list-style-type: none"> Type of tool (Planets service or non-Planets service or application) <p>Source: [PA-tool registry – SU6] Source: [Pronom7 SRD – S7.4.2.3.8]</p>	M
S8.2.2.3.9	<p>The system must, for the interface to retrieve a list of migration pathways, allow the following optional parameters:</p> <ul style="list-style-type: none"> Type of action (action on object, action on environment) <p>Source: [PA-tool registry – SU6, PA-tool registry – SU7, PA-tool registry – SU8] Source: [Pronom7 SRD – S7.4.2.3.9]</p>	M
S8.2.2.3.10	<p>The system must, in response to a request for migration pathways, return, for each migration pathway, the following information:</p> <ul style="list-style-type: none"> ID of migration pathway (PUID) status of migration pathway Type of action (action on object/action on environment) Details of each migration pathway step <p>Source: [PA-tool registry – SU6] Source: [Pronom7 SRD – S7.4.2.3.10]</p>	M
S8.2.2.3.11	<p>The system must, in response to a request for migration pathways, return, for each migration pathway, the following additional information if a detailed response has been requested:</p> <ul style="list-style-type: none"> maximum size of input file <p>Source: [PA-tool registry – SU10] Source: [Pronom7 SRD – S7.4.2.3.11]</p>	M
S8.2.2.3.12	<p>The system must, for each migration pathway step in a migration pathway, return the following information:</p> <ul style="list-style-type: none"> Migration pathway step number Software tool ID Software tool name Software tool parameters Type of tools (Planets service/ non-Planets service) <p>Source: [PA-tool registry – SU6] Source: [Pronom7 SRD – S7.4.2.3.12]</p>	M
S8.2.2.3.13	<p>The system must, for each migration pathway step in a file format migration pathway, return the following additional information:</p> <ul style="list-style-type: none"> input file format output file format <p>Source: [PA-tool registry – SU6] Source: [Pronom7 SRD – S7.4.2.3.13]</p>	M

Label	Requirement	Necessity
S8.2.2.3.14	The system must, for each migration pathway step in an emulation pathway, return the following additional information: <ul style="list-style-type: none"> target technical environment Source: [PA-tool registry – SU6] Source: [Pronom7 SRD – S7.4.2.3.14]	M

8.2.3 Redaction

This section details the SOAP web services that the system must provide as part of the system's support for redaction.

Label	Requirement	Necessity
S8.2.3.1	The system must provide the following operations as web services <ul style="list-style-type: none"> getRedactionTools Source: [Pronom 6 implementation] Source: [Pronom6 SRD – S9.1.1, Pronom7 SRD – S7.4.3.1]	M
S8.2.3.2	For backward compatibility with Pronom 6.2, the system must implement the following web service in accordance with the wsdl with the namespace of http://red.pronom.nationalarchives.gov.uk , which is given in Appendix A: <ul style="list-style-type: none"> getRedactionTools Source: [Pronom 6 implementation] Source: [Pronom6 SRD – S9.1.1, Pronom7 SRD – S7.4.3.2]	M

8.2.4 Digital Repository System Access Control

This section details the SOAP web services that the system must provide as part of the system's support for TNA's digital repository system.

Label	Requirement	Necessity
S8.2.4.1	Future enhancement, so removed from this document.	X

8.2.5 DROID

This section details the SOAP web services that the system must provide as part of the system's support for the DROID file format identification tool.

Label	Requirement	Necessity
S8.2.5.1	For backward compatibility with Pronom 6.2, the system must implement the web services below in accordance with the wsdl published at http://www.nationalarchives.gov.uk/Pronom/Services/Contract/PRONOM.wsdl <ul style="list-style-type: none"> getSignatureFileVersionV1 getSignatureFileV1 Source: [Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.5.2]	M
S8.2.5.2	The system must implement the following method <ul style="list-style-type: none"> getSignatureFileVersionV1 so that it returns the version of the current DROID signature file. Source: [DROID3 – S3.6.2.1, Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.5.3]	M

Label	Requirement	Necessity
S8.2.5.3	The system must implement the following method <ul style="list-style-type: none">• getSignatureFileV1 so that it returns the current signature file for DROID. Source: [DROID3 – S3.6.3.1, Pronom 6 implementation] Source: [Pronom7 SRD – S7.4.5.4]	M

9. System Wide Requirements

This section covers the requirements of the system needed to ensure that the product can actually be delivered and used and to ensure long-term quality.

9.1 Platform Requirements

The system will use two different web servers:

- The legacy search interface (see section 6), SOAP web services (see section 8.2) and REST interface (see section 8.1) will use Microsoft's IIS web server.
- The administration interface (see section 7), and SOAP web services (see section 8.2), will use a web server which supports version 6 of Java Virtual Machines (JVM).

Label	Requirement	Necessity
S9.1.1	The System will run on Windows 2003 Server. Source: [P4URD – 7, P3SD – S.2.7.2.1] Source: [Pronom6 SRD – S8.1.1, Pronom7 SRD – S8.1.1.1]	M
S9.1.2	The System will run on SQL Server 2000 SP3. Source: [P4URD – 7, P3SD – S.2.7.2.1] Source: [Pronom6 SRD – S8.1.2, Pronom7 SRD – S8.1.1.2]	M
S9.1.3	The System will use IIS 6 as the web server for the public web interface. Source: [P3SD – S.2.7.2.1] Source: [Pronom6 SRD – S8.1.3, Pronom7 SRD – S8.1.2.1]	M
S9.1.4	The system will use IIS 6 to provide the REST interface. Source: [Pronom6 SRD - S8.1.3, Pronom7 SRD – S8.1.2.2]	M
S9.1.5	The system will use Tomcat 6 to provide the Administration Web Interface. Source: [Tessella] Source: [Pronom7 SRD – S8.1.2.3]	M
S9.1.6	The system will use Tomcat 6 to provide the Web Service Interface. Source: [Tessella] Source: [Pronom7 SRD – S8.1.2.4]	M
S9.1.7	Future enhancement, so removed from this document.	X
S9.1.8	The client-side elements of the system must be compatible with: <ul style="list-style-type: none"> • IE 7.0 on Windows XP (SP3). • Firefox 3.0 on Windows XP (SP3). Source: [NPD/4678/CL/CSC/2008JUL02/10:04:03 - PC representative on the PLANETS Technical Coordination Committee] Source: [Pronom7 SRD – S8.1.3.6, which supersedes Pronom6 SRD – S8.1.4]	M

9.2 Documentation Requirements

Label	Requirement	Necessity
S9.2.1	The database schema must be documented. Source: [P4URD – 12] Source: [Pronom6 SRD – S8.2.2, Pronom7 SRD – S8.2.2]	M

Label	Requirement	Necessity
S9.2.2	The following documents must be delivered: <ul style="list-style-type: none"> • Software Requirements Document. • System and Acceptance Test Plan. • Architectural Design Document. • System Test Script. • System Maintenance / Installation Guide. Source: [PD] Source: [Pronom6 SRD – S8.2.3, Pronom7 SRD – S8.2.3]	M
S9.2.3	Future enhancement, so removed from this document.	X

9.3 Installation Requirements

Label	Requirement	Necessity
S9.3.1	The System will be supplied as a set of installers, and database scripts. Source: [P4URD – 11, P3SD - S2.7.5.1] Source: [Pronom6 SRD – S8.3.1, Pronom7 SRD – S8.3.1]	M
S9.3.2	All existing data must be ported to the new database schema. Source: [Tessella] Source: [Pronom6 SRD – S8.3.2, Pronom7 SRD – S8.3.2]	M
S9.3.3	Comprehensive, written installation instructions must be provided. Source: [P4URD – 11, Clarification 2] Source: [Pronom6 SRD – S8.3.3, Pronom7 SRD – S8.3.3]	M

9.4 Security Requirements

See individual functional requirements for appropriate security requirements.

9.5 Quality, Reliability and Maintainability Requirements

Label	Requirement	Necessity
S9.5.1	The system should make use of XSL transformations in order to avoid hard-coding report structures. Source: [P3SD - S2.7.7.2] Source: [Pronom6 SRD – S8.5.2, Pronom7 SRD – S8.5.2]	D
S9.5.2	The system will be Year 2000 compliant. Dates will be stored in a compliant format. Dates will be represented to users as DD-Mmm-YYYY. Source: [P4URD – 9, P3SD - S2.7.7.3] Source: [Pronom6 SRD – S8.5.3, Pronom7 SRD – S8.5.3]	M
S9.5.3	The system must conform to e-GIF version 6. Source: [P4URD – 10] Source: [Pronom6 SRD – S8.5.5, Pronom7 SRD – S8.5.5]	M
S9.5.4	Future enhancement, so removed from this document.	X
S9.5.5	The system must validate input to safeguard against incorrect or malformed requests which may jeopardise system security, in particular embedded SQL or HTML within query strings. Source: [P3SD – S2.3.3.3] Source: [Pronom6 SRD – S8.5.10, Pronom7 SRD – S8.5.10]	M

Appendix A Redaction Web Service WSDL

The following is the WSDL for the redaction web service.

```
<?xml version="1.0" encoding="utf-8"?>
<wsdl:definitions name="RedactionService"
xmlns:http="http://schemas.xmlsoap.org/wsdl/http/"
```

```

xmlns:soap="http://schemas.xmlsoap.org/wsdl/soap/"
xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:tns="http://red.pronom.nationalarchives.gov.uk"
targetNamespace="http://red.pronom.nationalarchives.gov.uk"
xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
xmlns:mime="http://schemas.xmlsoap.org/wsdl/mime/"
xmlns:wsdl="http://schemas.xmlsoap.org/wsdl/"
xmlns:xip="http://www.nationalarchives.gov.uk/XIP">
  <wsdl:types>
    <xs:schema elementFormDefault="qualified"
xmlns:tns="http://red.pronom.nationalarchives.gov.uk"
targetNamespace="http://red.pronom.nationalarchives.gov.uk"
xmlns:xip="http://www.nationalarchives.gov.uk/XIP">

      <!-- Get Messages Schemas -->

      <xs:element name="PUID">
        <xs:complexType >
          <xs:sequence>
            <xs:element name="Value" type="xs:string"
maxOccurs="1" minOccurs="0"/>
          </xs:sequence>
        </xs:complexType>
      </xs:element>

      <!-- Response message schemas -->

      <xs:element name="ToolList">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="Tool" type="tns:Tool"
minOccurs="0" maxOccurs="unbounded"/>
          </xs:sequence>
        </xs:complexType>
      </xs:element>

      <!-- Types -->

      <xs:complexType name="Tool">
        <xs:sequence>
          <xs:element name="className" type="xs:string"/>
          <xs:element name="toolName" type="xs:string"/>
          <xs:element name="toolTypeID" type="xs:int"/>
          <xs:element name="toolID" type="xs:int"/>
          <xs:element name="toolPriority" type="xs:int"/>
        </xs:sequence>
      </xs:complexType>

    </xs:schema>
  </wsdl:types>
  <wsdl:message name="getRedactionToolsIn">
    <wsdl:part name="messagePart" element="tns:PUID"/>
  </wsdl:message>
  <wsdl:message name="getRedactionToolsOut">
    <wsdl:part name="messagePart" element="tns:ToolList"/>
  </wsdl:message>
  <wsdl:portType name="RedactionService">
    <wsdl:operation name="getRedactionTools">
      <wsdl:documentation
xmlns:wsdl="http://schemas.xmlsoap.org/wsdl/"> Gets a list of redaction
tools for a format or for all formats. The input parameter is a string
representing the format's PUID (e.g., fmt/44) otherwise it is blank to
return the tools for all formats.</wsdl:documentation>

```

```
        <wsdl:input message="tns:getRedactionToolsIn"/>
        <wsdl:output message="tns:getRedactionToolsOut"/>
    </wsdl:operation>
</wsdl:portType>
<wsdl:binding name="RedactionServiceSoap" type="tns:RedactionService">
    <soap:binding transport="http://schemas.xmlsoap.org/soap/http"
style="document"/>
    <wsdl:operation name="getRedactionTools">
        <soap:operation
soapAction="http://red.pronom.nationalarchives.gov.uk/getRedactionToolsIn"
style="document"/>
        <wsdl:input>
            <soap:body use="literal"/>
        </wsdl:input>
        <wsdl:output>
            <soap:body use="literal"/>
        </wsdl:output>
    </wsdl:operation>
</wsdl:binding>

    <wsdl:service name="RedactionService">
        <wsdl:port name="RedactionServiceSoap"
binding="tns:RedactionServiceSoap">
            <soap:address location="RedactionService?wsdl"/>
        </wsdl:port>
    </wsdl:service>
</wsdl:definitions>
```