

Fedora and the Preservation of University Records Project

2.1 Ingest Guide

Version
1.0

Date
September 2006

**Digital Collections and Archives, Tufts University
Manuscripts & Archives, Yale University**

An Electronic Record Research Grant funded by the
National Historical Publications and Records Commission

Digital Collections and Archives
Tisch Library Building
Tufts University
Medford, Massachusetts 02155
<http://dca.tufts.edu>

Manuscripts and Archives
Yale University Library
Yale University
P.O. Box 208240
New Haven, Connecticut 06520-8240
<http://www.library.yale.edu/mssa/>

© 2006 Tufts University and Yale University

Co-Principle Investigators
Kevin Glick, Yale University
Eliot Wilczek, Tufts University

Project Analyst
Robert Dockins, Tufts University

This document is available online at
http://dl.tufts.edu/view_pdf.jsp?urn=tufts:central:dca:UA069:UA069.004.001.00006
(September 2006)
and
<http://dca.tufts.edu/features/nhprc/reports/ingest/index.html>
(September 2006)

Fedora and the Preservation of University Records Project Website at
<http://dca.tufts.edu/features/nhprc/index.html>

Funded by the
National Historical Publications and Records Commission
Grant Number 2004-083

Fedora and the Preservation of University Records Project

PART ONE: INTRODUCTION

- 1.1 Project Overview
- 1.2 System Model
- 1.3 Concerns
- 1.4 Glossary
- 1.5 Requirements for Trustworthy Recordkeeping Systems and the Preservation of Electronic Records in a University Setting

PART TWO: INGEST

2.1 Ingest Guide

- 2.2 Ingest Projects
- 2.3 Ingest Tools

PART THREE: MAINTAIN

- 3.1 Maintain Guide
- 3.2 Checklist of Fedora's Ability to Support Maintain Activities

PART FOUR: FINDINGS

- 4.1 Analysis of Fedora's Ability to Support Preservation Activities
- 4.2 Conclusions and Future Directions

TABLE OF CONTENTS

Overview	1
Section A: Negotiate Submission Agreement	5
Overview.....	5
Part A1: Establish Relationship	6
Part A2: Define Project.....	10
Part A3: Collect Information and Assess Value of Records.....	14
Part A4: Assess Record Types.....	18
Part A5: Assess Formats	21
Part A6: Assess Identifier Rules	26
Part A7: Assess Copyright.....	30
Part A8: Assess Access Rights	36
Part A9: Assess Recordkeeping System	41
Part A10: Assess Feasibility	47
Part A11: Finalize Submission Agreement.....	51
Section B: Transfer and Validation.....	57
Overview.....	57
Part B1: Create and Transfer SIPs	58
Part B2: Validate.....	61
Part B3: Transform and Attach Metadata	66
Part B4: Formulate AIPs.....	70
Part B5: Assess AIPs	73
Part B6: Formally Accession	77
Submission Agreement	80
Components, Resources, Products, and Documentation.....	82
Appendix A: Using the Ingest Guide.....	95
Appendix B: Example of a Submission Agreement	97
Appendix C: Producer-Archive Interface Methodology Abstract Standard Crosswalk...	106

OVERVIEW

One of the key challenges to preserving electronic records in a meaningful way is preserving the authenticity and integrity of records during their movement from a recordkeeping system to a preservation system. This Ingest Guide describes the actions needed for a trustworthy ingest process. This process enables an Archive and Producer to move records from a recordkeeping system to a preservation system in a manner that allows a presumption of authenticity.

This Ingest Guide refers to ingest broadly, defining it as the entire process involved in moving records from a recordkeeping system to a preservation system. This process consists of the Producer and Archive agreeing to and defining what records will be transferred and the manner of the transfer, validation, and transformation. Following the Guide should help an Archive and Producer ensure the functional, not just byte-stream preservation, of records. Not only does the Guide articulate steps for ensuring that records are properly tracked and have maintained their structural integrity during ingest, it also provides a way for the Archive to ensure that records remain renderable, functional, and meaningful. Following the Guide should enable an Archive to have a trustworthy ingest process, which would allow a reasonable person to presume that a record has maintained its level of authenticity during ingest.

This guide does not describe the functional or technical requirements for building either a recordkeeping or a preservation system. Instead, this guide presents a detailed description of the complex ingest workflow step by step. For more on authenticity and trustworthy recordkeeping systems and the preservation of records see “Requirements for Trustworthy Recordkeeping Systems and the Preservation of Electronic Records in a University Setting.”¹

The Ingest Guide contains two main sections. Section A, Negotiate Submission Agreement, details how the Producer and the Archive create and arrange a Submission Agreement that defines the terms and conditions of the transfer of records from the Producer to the Archive, and it details the scope of the records along with the nature of their validation and transformation. Section B, Transfer and Validation, details the actual transfer, validation, and transformation of records. Section A contains eleven parts and Section B has six parts. Each part in Section A and B is composed of a number of steps.

Each part includes a narrative summary, a flowchart illustrating all of its steps, and a description of each step. Each description includes an Overview, a list of Components, Resources, Products, and Documentation that each step utilizes and/or produces, and a thumbnail flowchart.

The Ingest Guide also includes a separate section on Components, Resources, Products, and Documentation that describes each of these roles in the ingest process and refers to the steps that use and produce them. The Ingest Guide also has a Submission Agreement section that explains

¹ Fedora and the Preservation of University Records (NHPRC 2004-083), “Requirements for Trustworthy Recordkeeping Systems and the Preservation of Electronic Records in a University Setting,” <http://dca.tufts.edu/features/nhprc/reports/1_5final.pdf>.

the Agreement in further detail. Finally, the Guide includes crosswalks between its own steps of the *Producer-Archive Interface Methodology Abstract Standard*.²

Although the Ingest Guide is a prescriptive guide for a trustworthy ingest process, it is not a detailed manual of procedures. The implementation of the Guide can produce a wide variety of procedures and policies from archive to archive. The Guide describes the actions that must be undertaken to trust the ingest process and prescribes how to undertake these steps at a high level, but it does not prescribe how to proceed in full detail. For example, Step A5.5 calls for the Archive to choose a preservation format for records it chooses to transform, but it does not dictate what those preservation formats should be. An Archive following the Ingest Guide will still have to determine what preservation formats best serve its needs. The Guide points out many tasks that Archives must undertake to have a trustworthy ingest processes, without discussing those tasks in detail. The most prominent of these tasks include the details of the appraisal process in Part A3, the creation of Submission Information Packages in Part B1, and the creation of Resources. For more on the implementation of the Ingest Guide, see Appendix A: Using the Ingest Guide.

The Guide uses the *Open Archival Information System Reference Model (OAIS)* definition of Archive: “An organization that intends to preserve information for access and use.”³ Therefore, while an Archive may be an archives in the sense used by the archival community, it does not necessarily have to be an archives.⁴ In the context of the Ingest Guide, an Archive is any type of office or juridical body that has the responsibility of providing long-term preservation and access to records. Like *OAIS*, the Guide refers to a single ingest or a single set of recurring ingests as an Ingest Project.

The Ingest Guide also uses the *OAIS* definition of Producer: “The role played by those persons, or client systems, who provide the information to be preserved. This can include other OAISs or internal OAIS persons or systems.”⁵ This means that Producer will normally be the custodian of the records—or an entity the Producer has authorized to act on its behalf—that has the authority to transfer the records to the Archive. The Producer may or may not be the individual, group, or organization that is responsible for the creation, production, accumulation, or formation of the records it transfers to the Archive.

The Ingest Guide is based upon the work of the Consultative Committee for Space Data Systems (CCSDS) and builds upon its *OAIS* framework. In particular, Section A of the Ingest Guide is based on the CCSDS’s *Producer-Archive Interface Methodology Abstract Standard*, which is composed of four phases: Preliminary, Formal, Transfer, and Validation. The *Producer-Archive Interface* is a follow-up document to *OAIS*. The Preliminary and Formal phases greatly expand on the Negotiate Submission Agreement activity in the Administration function of *OAIS*. The

² Consultative Committee for Space Data Systems, *Producer-Archive Interface Methodology Abstract Standard*, CCSDS 651.0-B-1, Blue Book, May 2004. <<http://www.ccsds.org/CCSDS/documents/651x0b1.pdf>>.

³ ISO 14721:2003, Space data and information transfer systems -- Open Archival Information System -- Reference model.

⁴ For the “archival community” definition of the term archives see Richard Pearce-Moses, *A Glossary of Archival and Records Terminology* (Archival Fundamentals Series II), (Chicago: Society of American Archivists, 2005).

⁵ ISO 14721:2003, p. 1–12.

Transfer and Validation phases reiterate the Receive Submission and Quality Assurance activities respectively, both of which are in the Ingest function of *OAIS*.

As a product of the Fedora and the Preservation of University Records grant project (NHPRC 2004-083), the Ingest Guide is designed primarily for a university setting. However, its general nature may also make it applicable in other environments. The university orientation of the Ingest Guide differs from the *Producer-Archive Interface*'s orientation. While the PAI treats the creation of a Submission Agreement formally and in two different phases ("Preliminary" and "Formal"), the Ingest Guide does not distinguish preliminary and formal phases, because this level of formality is unnecessary. Such formality would impose unrealistic implementation expectations in a university setting.

The Ingest Guide is based largely on the conceptual underpinnings of the records lifecycle model, presuming that a Producer will create, acquire, utilize, and manage records in a Recordkeeping System to suit its current business needs, and later the Archive will ingest some of those records into a separate Preservation System that the Archive administers. In this model, the Archive acts as a neutral third party to the recordkeeping process acting on behalf of broader societal needs rather than on behalf of the Producer. As a neutral third party the Archive has no stake in the content of the records and no reasons to alter records under its custody, and it should not allow anybody to alter the records either accidentally or on purpose. Many archivists have rejected the lifecycle model in favor of the records continuum concept, where recordkeeping is seen as a continuous process that is not time-based, separated into a series of clearly defined steps, or administered by completely separate juridical entities. Many Producers and Archives operate in a mixed world between these two models. For example, many Archives operate separately from a Producer but are part of same organization as the Producer and do not act as a neutral third party. The Ingest Guide should be useful to most Archives operating in a mixed lifecycle/continuum environment, particularly ones where separate Recordkeeping and Preservation Applications are maintained.⁶

The Ingest Guide assumes that a Producer is submitting managed records to an Archive. Traditionally, an archives might accept boxes of unorganized paper records from a faculty member, for example, with the idea that the archives could add these records to its processing backlog and later impose some sort of order, or arrangement, long after the transfer. It is the assumption of this Guide, and the corresponding preservation requirements, that such delayed arrangement of electronic records is neither scalable nor sustainable. A box of unlabeled disks sent from a faculty member illustrates this point. The Ingest Guide places the activity of imposing order on electronic records outside of the ingest and preservation activities. The work of preparing organized and managed records for transfer to the Archive is the Producer's responsibility and, in the case of the box of unorganized diskettes, the Archive is doing the Producer's job, imposing order on the records after the fact. In a situation like this, during part A of the Ingest Guide, the Archive would either require the Producer to organize the disks and the records they hold before the transfer takes place, or artificially organize the records after accepting them from the Producer but still before the transfer to the preservation repository. By

⁶ The authors do not wish to express any opinion of the relative merit of either the lifecycle or continuum models, but instead simply disclose their inherent bias towards the lifecycle model based on educational background and work experience.

imposing this artificial arrangement the Archive has become a Producer and thus plays both roles in the Ingest process.

SECTION A: NEGOTIATE SUBMISSION AGREEMENT

Overview

The Negotiate Submission Agreement Section of the Ingest Guide describes the actions needed for an Archive and a Producer to generate a Submission Agreement. These agreements define the nature and scope of the records to transfer to the Preservation System and how the Archive will execute transfer, validation, and transformation of these records. All of the work in Section A is undertaken for the production of a Submission Agreement. Actual transfer, validation, and transformation work only occurs in Section B of the Ingest Guide.

This section is composed of eleven parts. During first three parts, Establish Relationships, Define Project, and Collect Information and Assess Value of Records, the Archive conducts an intellectual appraisal of the records under consideration and determines if it should accession them into the Preservation System. During these three parts, the Archive also gathers a variety of information about the records it should accession. This information includes the creator, record type, format, date, and extent of the records and any identifiers associated with them. During the next seven parts, Assess Record Types, Assess Formats, Assess Identifier Rules, Assess Copyright, Assess Access Rights, Assess Recordkeeping System, and Assess Feasibility, the Archive determines if its existing resources for preservation formats, record types, identifier rules, creator records, security procedures, transfer procedures, and system capabilities—referred to as Resources and described in the Components, Resources, Products, and Documentation section of this guide—meets the needs of the records identified in Part A3. The resulting feasibility report should present a gap analysis if the Resources do not reflect the continuing value of any records assessed in A3. The Archive must then determine if it should modify or add to its Resources to meet those assessments or if it should reject or modify the scope of the records involved in the Ingest Project. In Part A11, Finalize Submission Agreement, the Archive and the Producer finalize and agree to a Submission Agreement that is based on the scope of the records defined in Parts A1 through A3 and the decisions made in Parts A4 through A10.

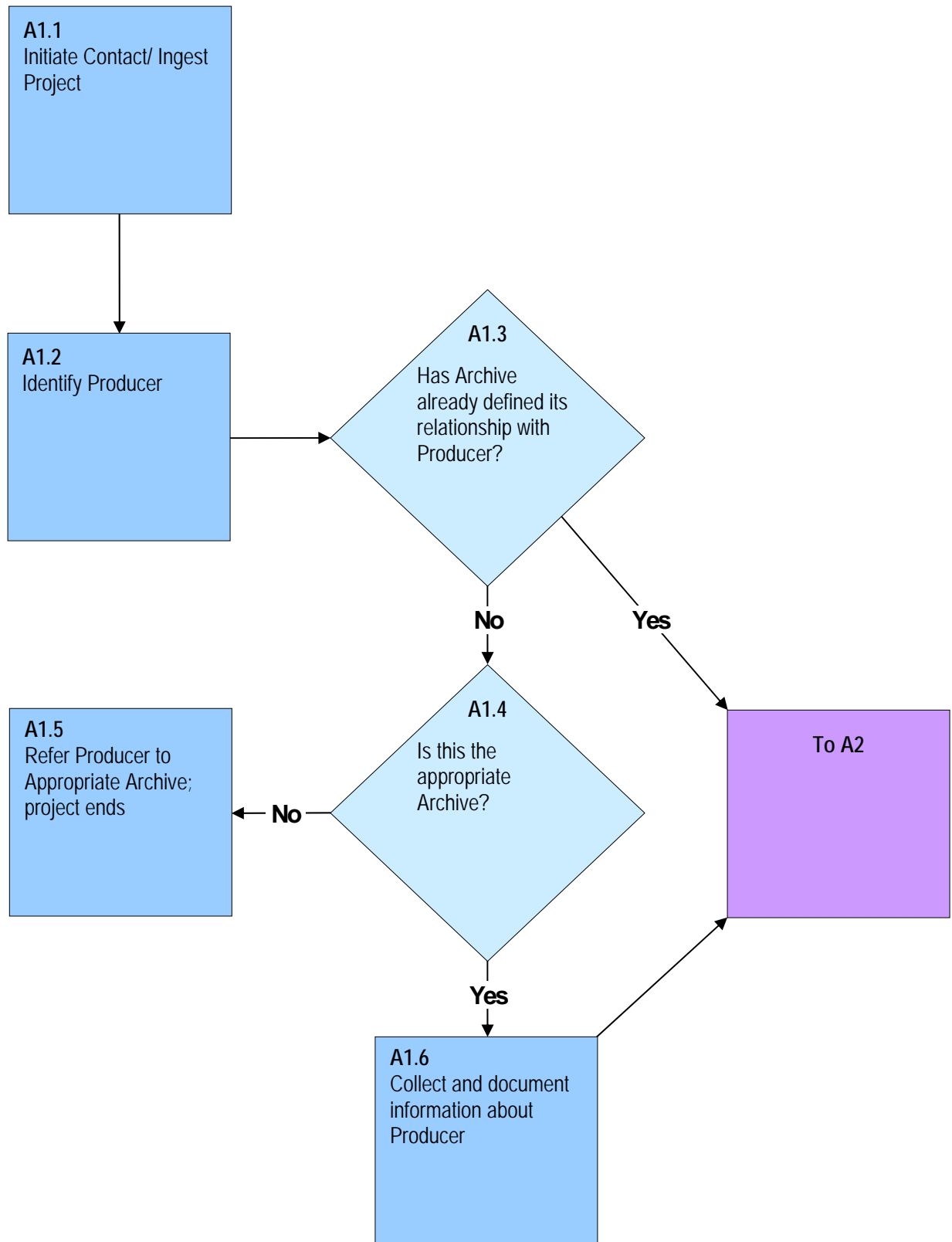
For more information on Submission Agreements, see the Submission Agreement section of this Guide.

SECTION A: NEGOTIATE SUBMISSION AGREEMENT

Part A1: Establish Relationship

Overview

During this Part, either the Archive or the Producer will initiate contact with the other. If the Archive does not already have a relationship with the Producer, the Archive will define its administrative, legal, and/or collecting relationship with the Producer and generate metadata about the Producer. Not every relationship that is established will result in a negotiated Submission Agreement, or even a project definition (Part A2). It is also possible that there will be a time delay between establishing a relationship and defining a potential accession. The Archive should document contacts with Producers to manage potential accession opportunities. How the Archive manages this documentation will depend on the size of the Archive, the number of relationships, the frequency of contact, and the length of the delays.

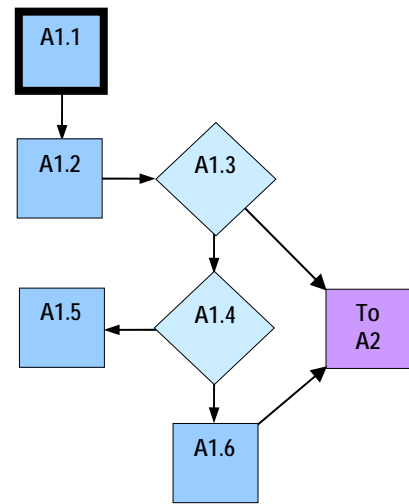


A1.1

Description Either the Archive or the Producer will initiate contact with the other. This is the first step of an Ingest Project. Generally this contact is made informally, although it can be made formally if necessary. Document contact in appropriate Activity Log.

Uses None

Produces/Modifies Activity Log

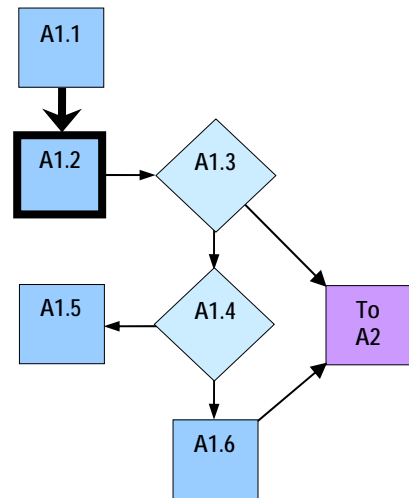


A1.2

Description The Archive determines who the Producer is, in particular determining who he/she/it is as a juridical body.

Uses Institutional Identity Management System

Produces/Modifies None

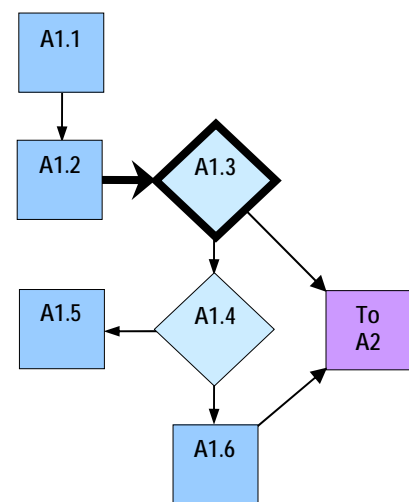


A1.3

Description Based on identifying who or what the Producer is in Step A1.2, the Archive determines if it has already established a Producer-Archive relationship with the Producer in which it can serve as the Archive for at least some of the Producer's records. If this is the case, the Archive documents who the Producer is in the Submission Agreement and goes on to Part A2 of the Ingest Guide. The Archive makes this determination by reviewing Producer Records, Accession Logs; or Activity Logs that describe Producers, their relationships to the Archive, and document its past interactions between the Archive and Producers.

Uses Accession Log, Activity Log, Producer Record

Produces/Modifies Producer Entry

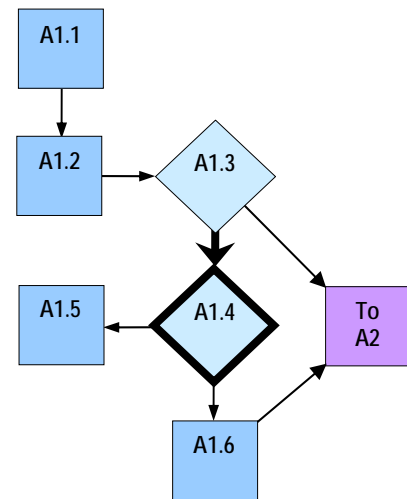


A1.4

Description If the Archive has not already established a relationship with the Producer as the Archive for at least some of its records, the Archive needs to determine if in fact it has the authority to serve as the Archive for at least some of the Producer's records.

Uses Records Authority Statement, Collection Policy

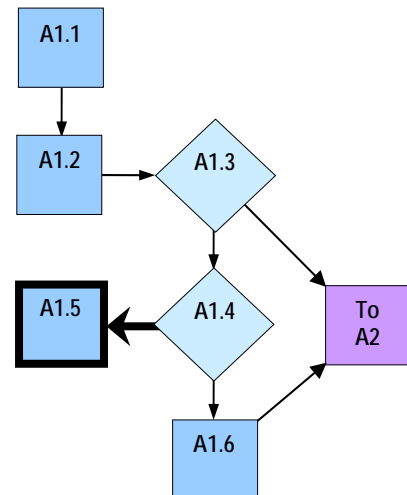
Produces/Modifies None

**A1.5**

Description If the Archive does not have the authority to serve as the Archive for at least some of the Producer's records, it should recommend the Producer to an appropriate Archive. The Ingest Project ends.

Uses Archives Directory

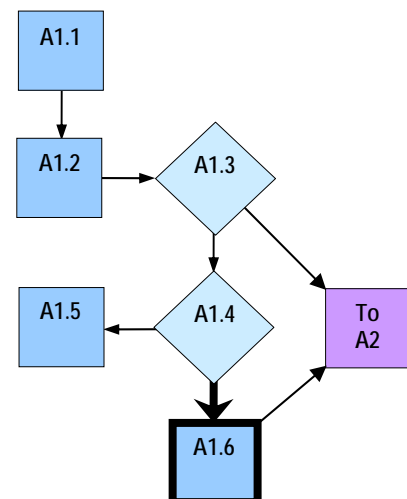
Produces/Modifies Ingest Project Termination Notice

**A1.6**

Description If the Archive does have the authority to serve as the Archive for at least some of the Producer's records, it needs to collect information about the Producer and create a Producer Record and document who the Producer is in the Submission Agreement.

Uses Institutional Identity Management System

Produces/Modifies Producer Record, Producer Entry

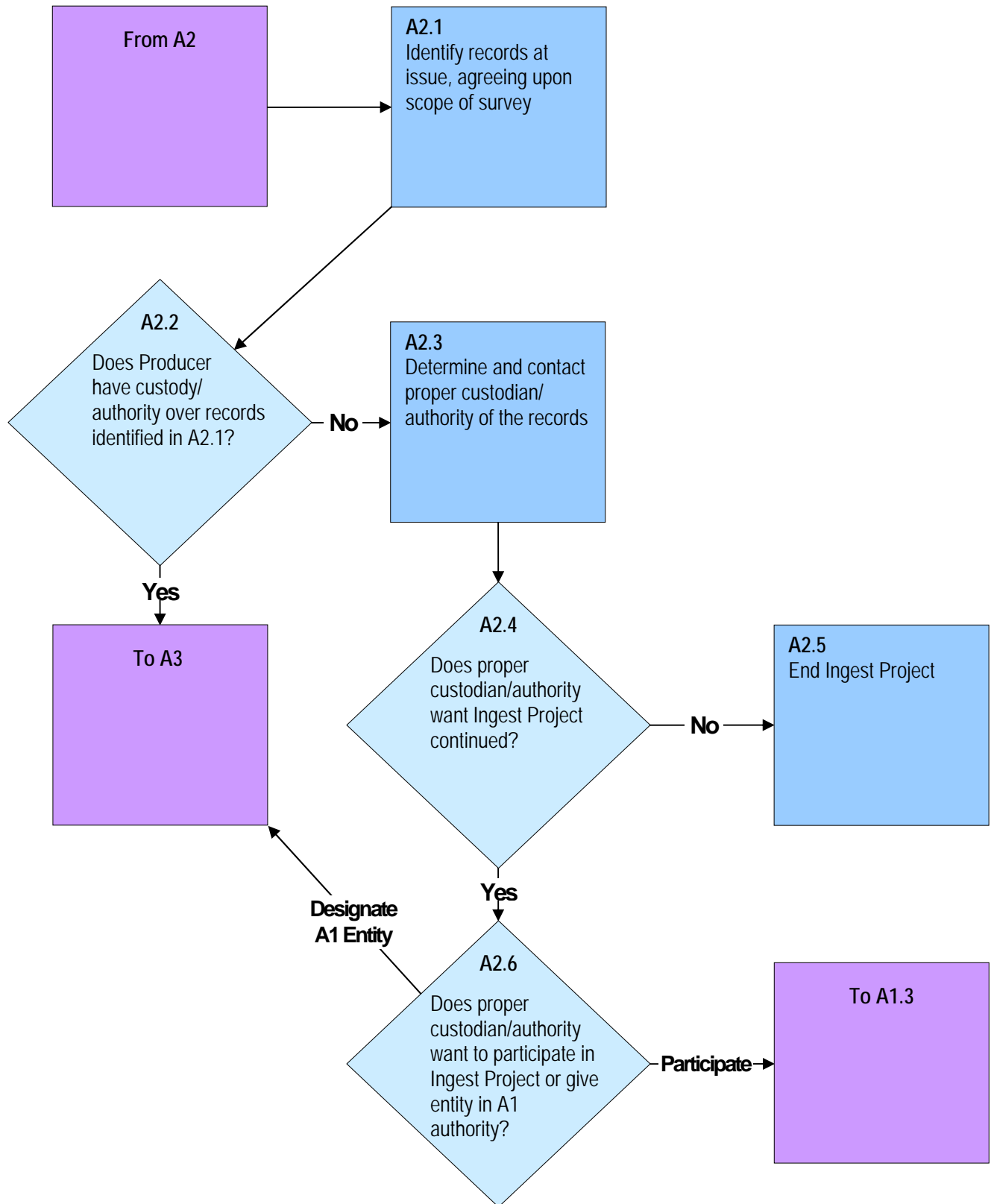


SECTION A: NEGOTIATE SUBMISSION AGREEMENT

Part A2: Define Project

Overview

During this Part, the Archive and the Producer come to an agreement regarding which records the Archive will consider for accession, essentially defining the scope of the Ingest Project. The Archive then verifies that the Producer has proper custody of the records under consideration.

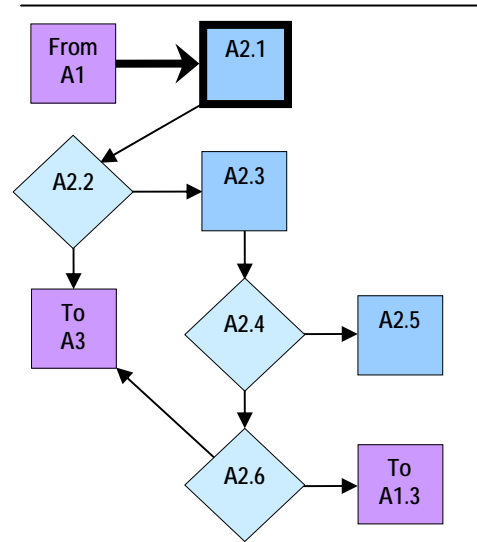


A2.1

Description The Archive and the Producer agree upon the scope of the records that the Archive will survey. The Archive should document this scope of records in what will ultimately become a Survey Report.

Uses None

Produces/Modifies Survey Report

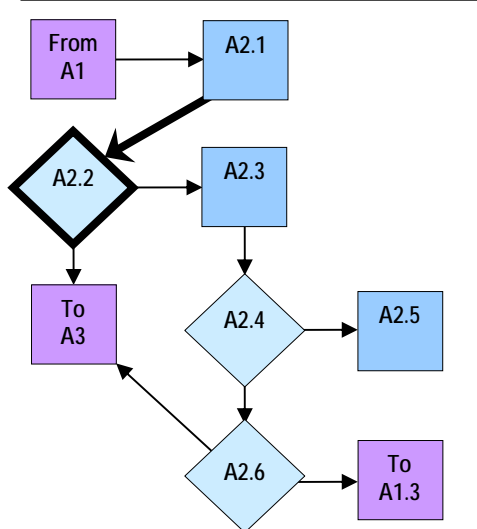


A2.2

Description The Archive determines if the Producer actually has custody or authority over the records the Archive agreed to survey in Step A2.1.

Uses Producer Record

Produces/Modifies None

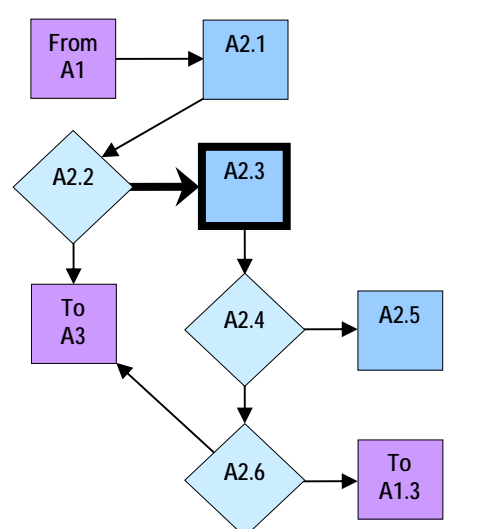


A2.3

Description If the Producer that approached the Archive in Step A1.1 does not have custody or authority over the records, the Archive determines what Producer has proper authority or custody over the records and then contacts that Producer.

Uses Institutional Identity Management System, Producer Record, Activity Logs

Produces/Modifies None

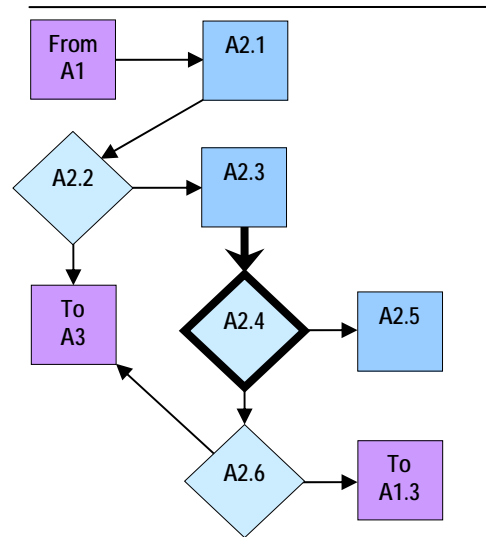


A2.4

Description The Producer who is the proper custodian or authority of the records determines if it wants the Ingest Project to continue.

Uses None

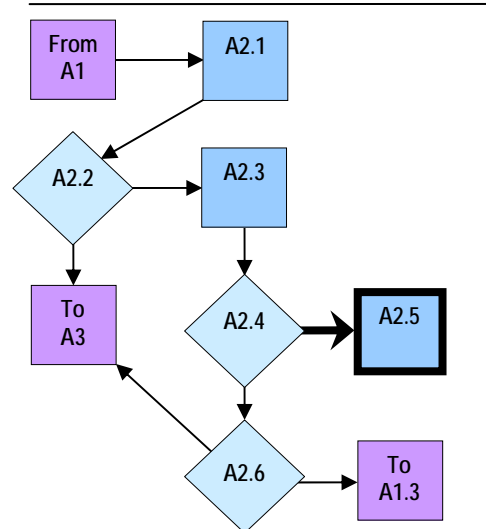
Produces/Modifies None

**A2.5**

Description If the proper custodian or authority of the records does not want the Ingest Project to continue, the Ingest Project ends.

Uses None

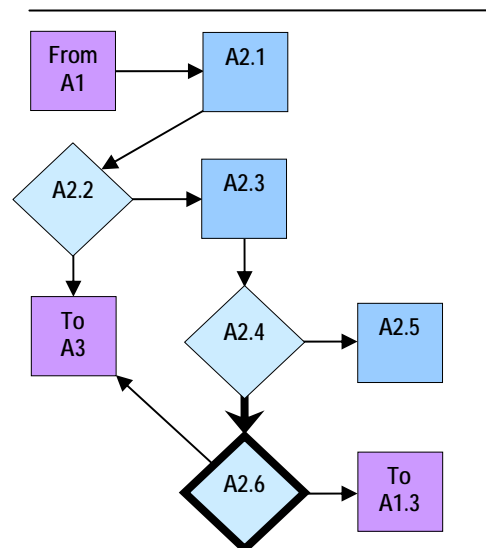
Produces/Modifies Ingest Project Termination Notice

**A2.6**

Description The Producer who is the proper custodian or authority of the records determines if it wants to participate in the Ingest Project or designate the Entity from Step A1.1 as the Producer to continue the Ingest Project. If the Producer who is the proper custodian or authority of the records wants to participate in the Ingest Project, the Archive should return to Step A1.3 to determine if it has already established a Producer-Archive relationship with the Producer.

Uses None

Produces/Modifies None

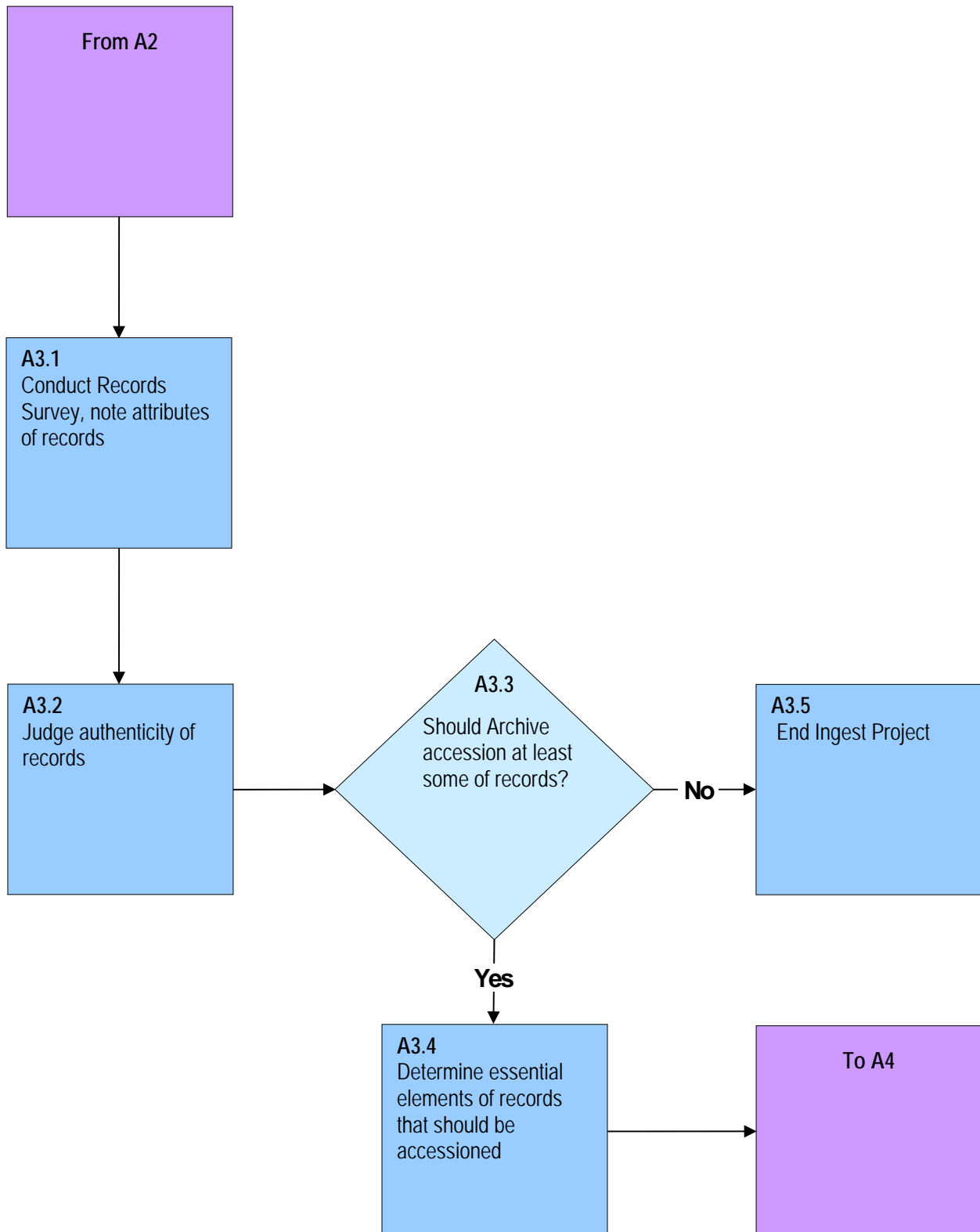


SECTION A: NEGOTIATE SUBMISSION AGREEMENT

Part A3: Collect Information and Assess Value of Records

Overview

Once the Archive has confirmed it is working with the proper custodian, it will conduct a survey of the records identified in Part A2 in order to collect the information needed to undertake the assessments described throughout the rest of Section A. In addition to collecting information the Archive will analyze the records in order to assess their continuing value and authenticity. This assessment will be combined with an assessment of the feasibility of preservation (Parts A4 through A10) in making a final appraisal decision. The appraisal work in this Part focuses on whether the records intellectually belong in the Archive, asking if it *should* accession the records.

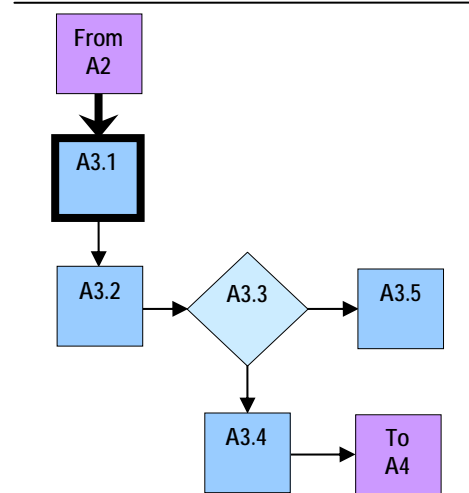


A3.1

Description The Archive conducts a survey of the records identified in Step A2.1, delineating the descriptive data and other information that must be collected in order to make the assessments required in Parts A3 through A10. The Archive will note the record type of the records, the function(s) the records play for the Producer, their recordkeeping environment, format types, file size, Producer-created identifiers, confidentiality requirements, and copyright status in the Survey Report. Such a survey might be undertaken through any combination of a number of different methods, including the Archive interviewing the Producer, the Archiving querying the Producer through some sort of questionnaire, or even the Archive querying the electronic records themselves. The Survey Report should capture all of the information that the Archive will need to complete Parts A4 through A10.

Uses Survey Procedures, Survey Instrument, Survey Report

Produces/Modifies Survey Report

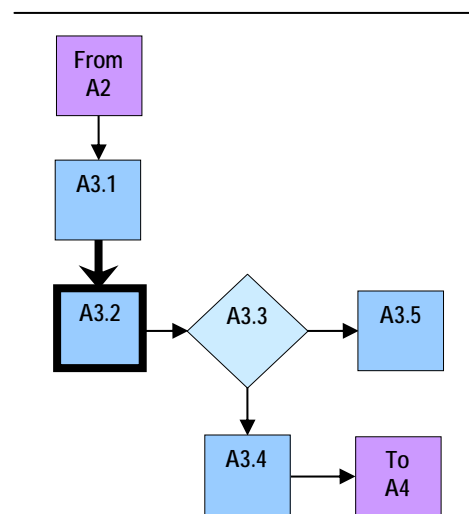


A3.2

Description The Archive analyzes and judges the grounds for presuming the authenticity of the records identified in Step A2.1. This includes determining if the recordkeeping system that maintains the records has the qualities that allows one to presume the records' authenticity. The Archive should base this determination on the "Requirements for Trustworthy Recordkeeping Systems and the Preservation of Electronic Records in a University Setting" or some other set of requirements for trustworthy recordkeeping systems. The Archive may forgo this evaluation if it has already identified the recordkeeping system as a trustworthy system in the Recordkeeping System Report. If the Archive evaluates a new or modified system, it should document the system in a new or updated Recordkeeping System Report. Judging the grounds for presuming the authenticity of the records in the Ingest Project also includes checking if the records are managed according to the rules of the recordkeeping system, and sometimes examining the extrinsic and intrinsic qualities of the records themselves.

Uses Recordkeeping System Report, Recordkeeping System Evaluation Tool

Produces/Modifies Survey Report, Recordkeeping System Report

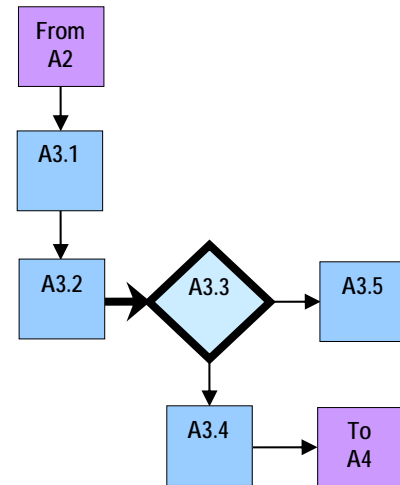


A3.3

Description The Archive determines if it should accession at least some of the records it surveyed. In this Step, this determination is based entirely on policy considerations with no regard to an Archive's capacity or technical considerations, which will come later during the Ingest Project. The Archive makes this appraisal decision by considering the Producer, the record type of the records, the function(s) the records play for the Producer, and the records' authenticity against the Archive's Collection Policy and Records Retention Schedule.

Uses Survey Report, Collection Policy, Records Retention Policy

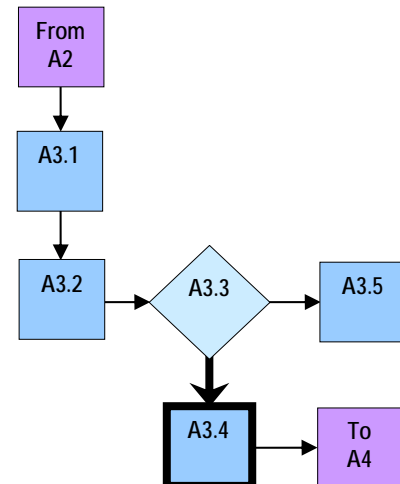
Produces/Modifies Survey Report

**A3.4**

Description The Archive determines the essential elements of the records—documentary components, elements of form, and digital or physical components—it needs to preserve in order to preserve the recordness and authenticity of the records. This Step will guide the Archive's later decisions concerning format transformation.

Uses Survey Report, Collection Policy, Records Retention Policy

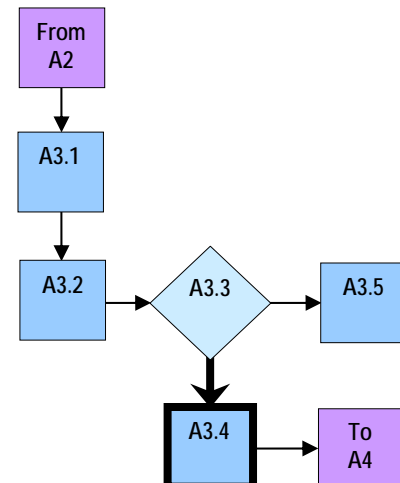
Produces/Modifies Survey Report

**A3.5**

Description If an Archive decides that it should not accession any of the records identified in Step A2.1, the Ingest Project ends.

Uses None

Produces/Modifies Ingest Project Termination Notice

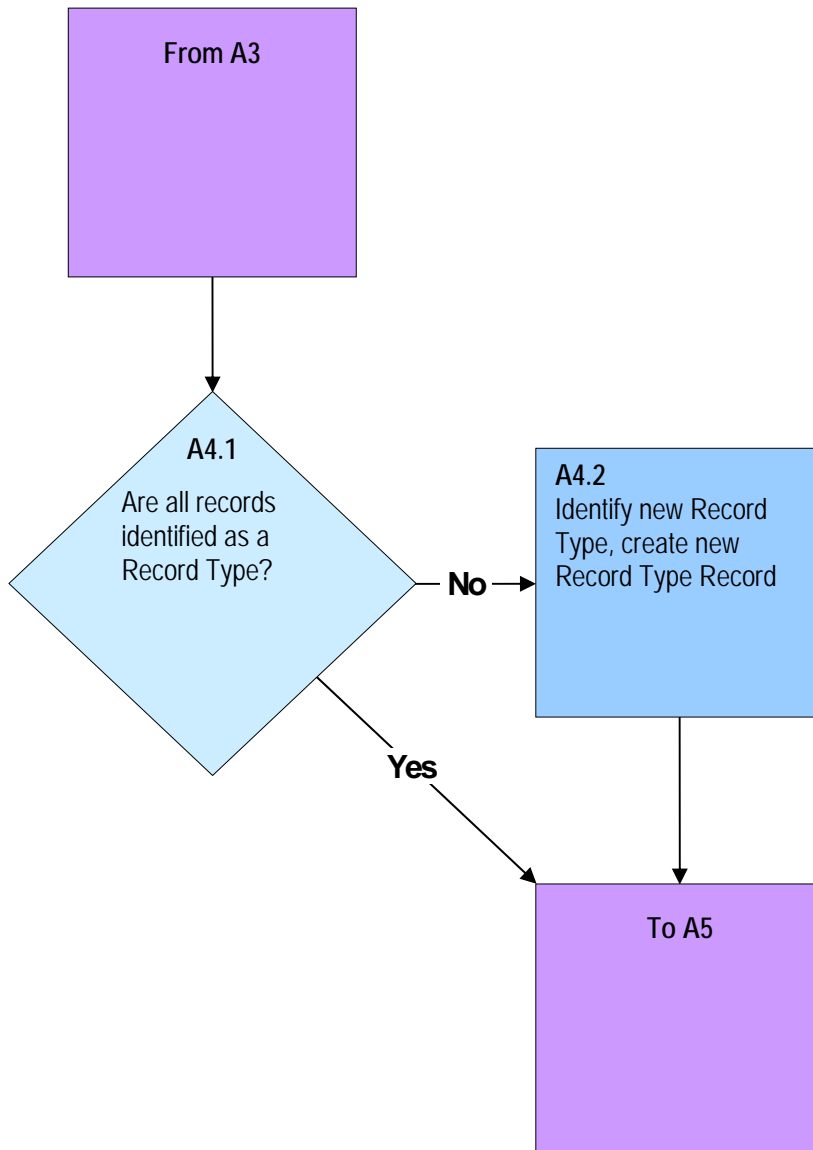


SECTION A: NEGOTIATE SUBMISSION AGREEMENT

Part A4: Assess Record Types

Overview

In this part the Archive determines if any of the records that should be accessioned are record types that the Archive has not previously defined or dealt with. If that is the case, the Archive will define the new record type. Record types define the nature of a class of records. They usually define retention and disposition of records, and, sometimes, their confidentiality status. Definitions of record types usually guide an Archive's appraisal and format transformation decisions.

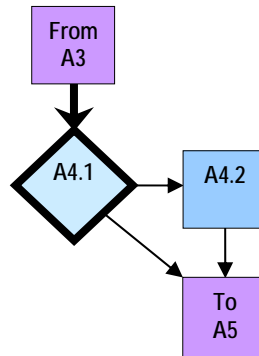


A4.1

Description Based on the information gathered in the Survey Report, the Archive determines if any of the records it should accession are record types that are not one of its established record types.

Uses Survey Report, Record Type Record

Produces/Modifies Record Type List

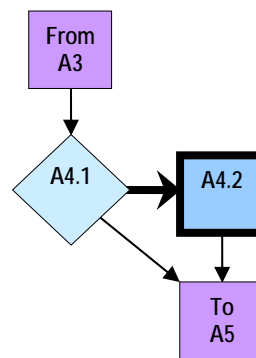


A4.2

Description If there are records in an Ingest Project that are a record type that is not one of the Archive's established record types, then it identifies this new record type and creates a new Record Type Record.

Uses Survey Report

Produces/Modifies Record Type Record

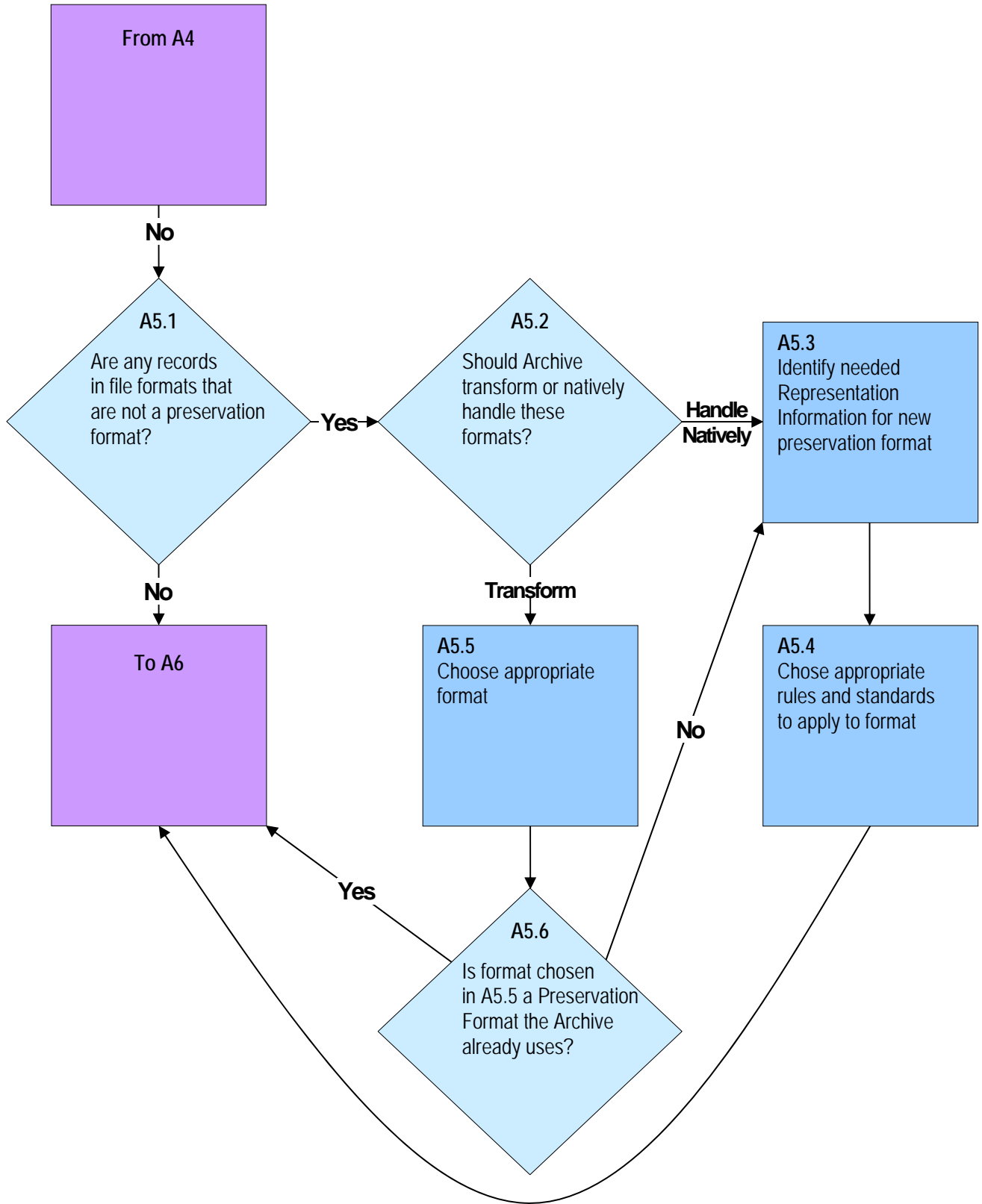


SECTION A: NEGOTIATE SUBMISSION AGREEMENT

Part A5: Assess Formats

Overview

In this Part the Archives appraises the formats of the records that it should accession into the Preservation System and determines if any of these records are in formats that are not one of the formats that the Preservation System supports. The Archives must determine if it will: 1) transform the records into one of the existing preservation formats, 2) transform the records into a new preservation format, or 3) keep the records in their existing formats and make that format a new preservation format. In order to create a new preservation format, the Archive will have to select a variety of rules and standards that apply to the format and identify the Representation Information that is needed for the new format.

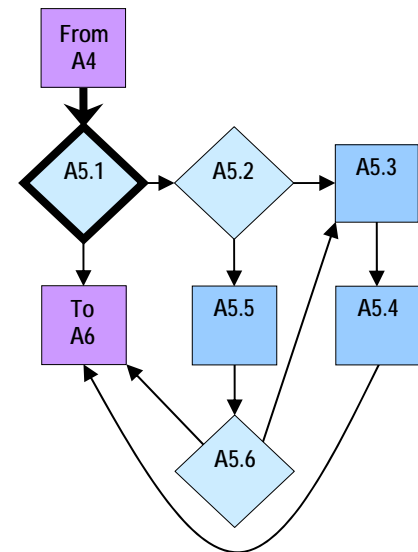


A5.1

Description Based on the information gathered in the Survey Report, the Archive determines if any of the digital components of records earmarked for preservation are in formats that meet the Archive’s Format Standards Policy. If all of the digital components comply with the Format Standards Policy, the Archive proceeds to Part A6.

Uses Survey Report, Format Standards Policy, Format Representation Information System

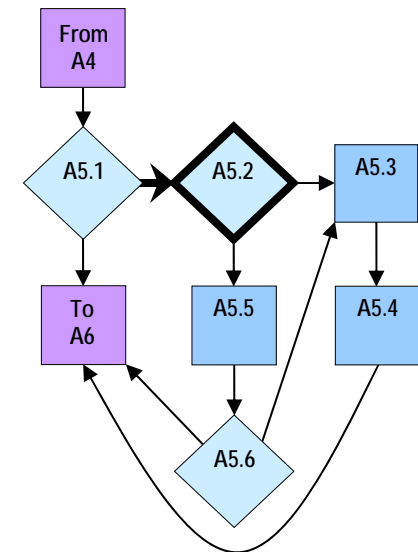
Produces/Modifies Transformation Plan

**A5.2**

Description If there are digital components of records in an Ingest Project that are not compliant with the Archive’s Format Standards Policy, the Archive must decide on the proper preservation strategy for these formats of digital components. The Archive may decide to transform the records into one of its preservation formats, manage the formats natively, or preserve both the native format and some preservation format. Although there are technical considerations to this determination, this is largely an appraisal decision. The Archive must determine how crucial the format is to the structure and ultimately the essential recordness of the record.

Uses Survey Report, Format Representation Information System, Preservation System Capabilities Report

Produces/Modifies None

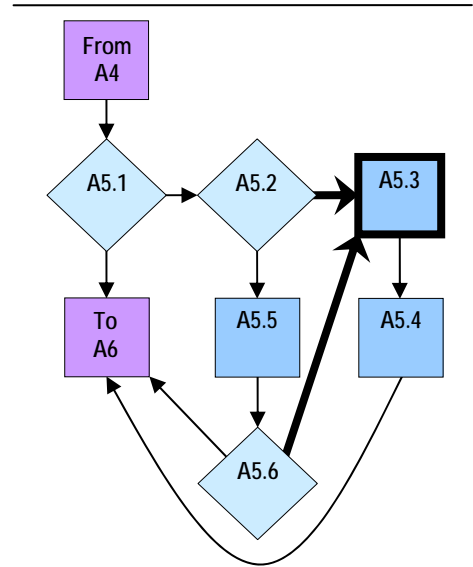


A5.3

Description If the Archive determines that it should preserve the records in the Preservation System in their existing native format(s), then it needs to establish this format in the Format Representation Information System and Formats Standards Policy.

Uses Designated Community Description

Produces/Modifies Format Representation Information System, Representation Information, Formats Standards Policy

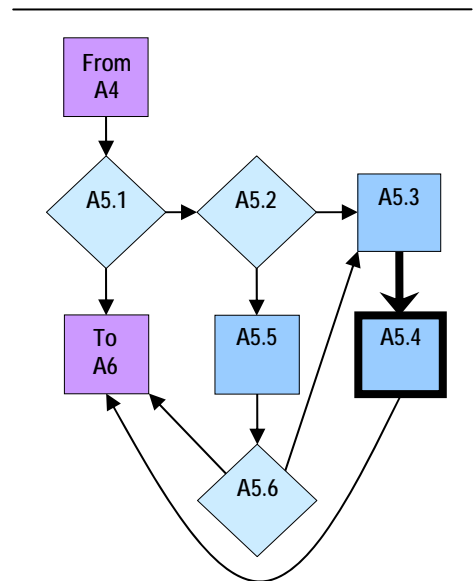


A5.4

Description In addition to the Representation Information, the Archive also needs to establish its rules and select appropriate standards for the new preservation format so it can manage and preserve records in that format successfully.

Uses None

Produces/Modifies Format Standards Policy, Transformation Plan

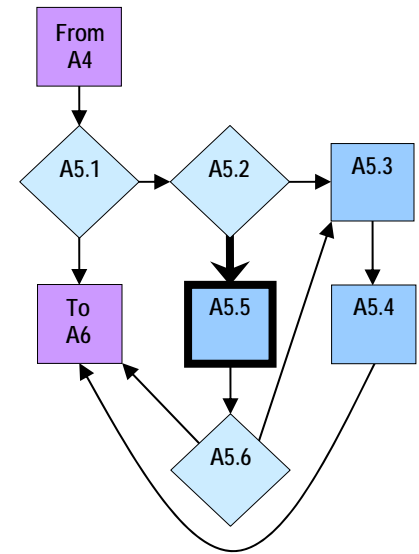


A5.5

Description If the Archive decides to transform the records, it must determine into which preservation format(s) to transform the records. See Step A5.2 for comments about the appraisal considerations involved in this step. If the Archive is dealing with a new combination of record type and format type, it will probably have to add to or modify its Transformation Policy.

Uses Survey Report, Transformation Policy

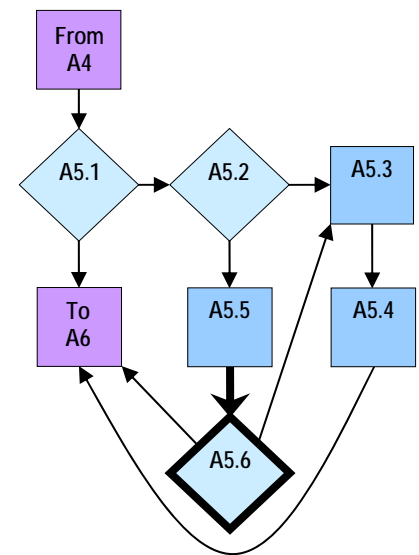
Produces/Modifies Transformation Policy, Transformation Plan

**A5.6**

Description The preservation format into which the Archive chooses to transform a record maybe either an existing preservation format or a new preservation format. If it is a new preservation format, the Archive goes to Step A5.3.

Uses None

Produces/Modifies None

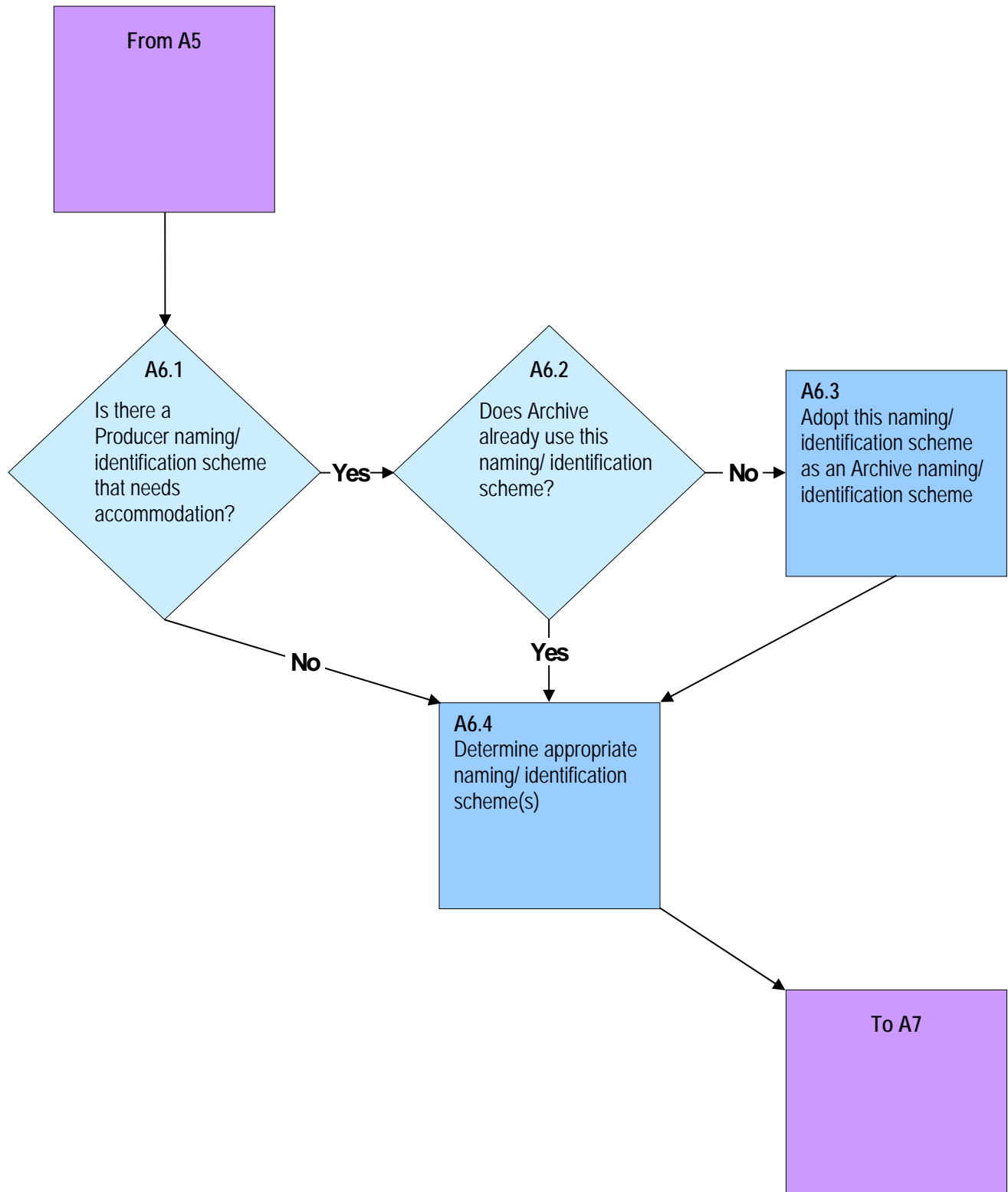


SECTION A: NEGOTIATE SUBMISSION AGREEMENT

Part A6: Assess Identifier Rules

Overview

In this Part the Archive determines if it needs to preserve a Producer identification scheme with the records it should accession. For example, if a Producer's website is involved in an Ingest Project, the Archive may decide to preserve the file and path names of the HTML files to preserve the integrity of the website's internal links. In another instance the Archive may want to preserve a Producer's file and directory names of desktop applications stored on a network file system. If the Archive does need to preserve such a scheme, then the Archive must determine if it already manages the scheme in the Preservation System. If it does not, then it needs to create new rules to accommodate the new scheme.

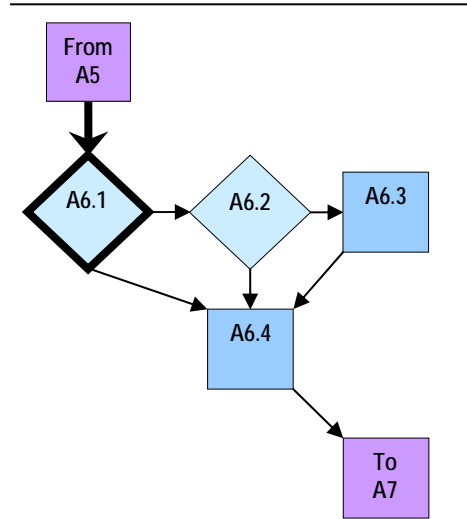


A6.1

Description Based on the information gathered in the Survey Report, the Archive determines if any of the records it should accession need to remain tied to a Producer's naming or identification scheme in the Preservation System.

Uses Survey Report, Producer Naming/Identification Scheme

Produces/Modifies None

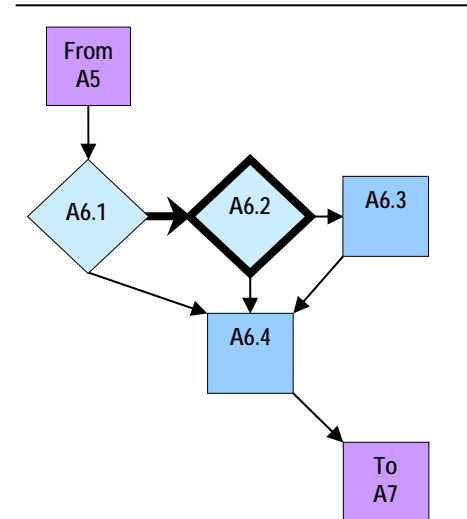


A6.2

Description The Archive determines if it has already adapted that Producer Naming/Identification Scheme as an Archive Naming/Identification Scheme.

Uses Archive Naming/Identification Scheme, Producer Naming/Identification Scheme

Produces/Modifies None

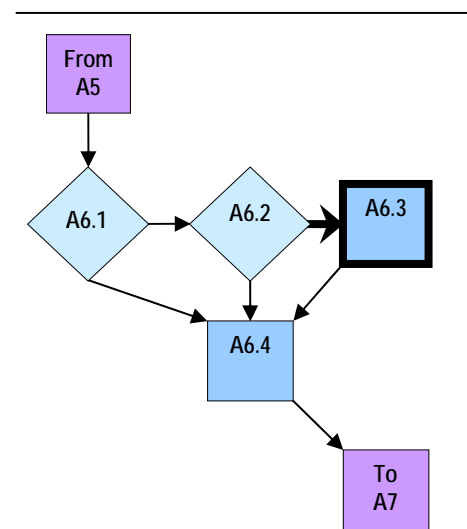


A6.3

Description If the Archive has not adapted the Producer Naming/Identification Scheme as an Archive Naming/Identification Scheme, it will do so.

Uses Producer Naming/Identification Scheme

Produces/Modifies Archive Naming/Identification Scheme

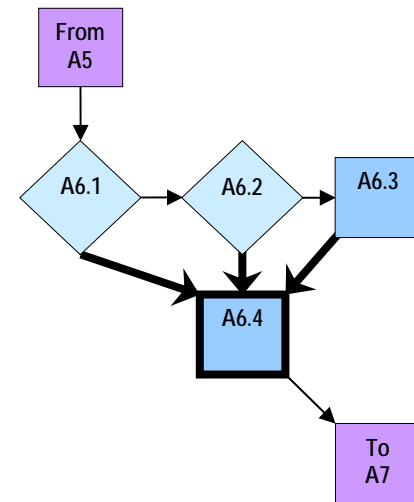


A6.4

Description The Archive determines which of its Naming/Identification Schemes it should use to name or identify the records in an Ingest Project. In many cases the Archive does not really have a choice to make because it only supports one scheme.

Uses Archive Naming/Identification Scheme, Survey Report

Produces/Modifies Archive Naming/Identification Scheme Decision

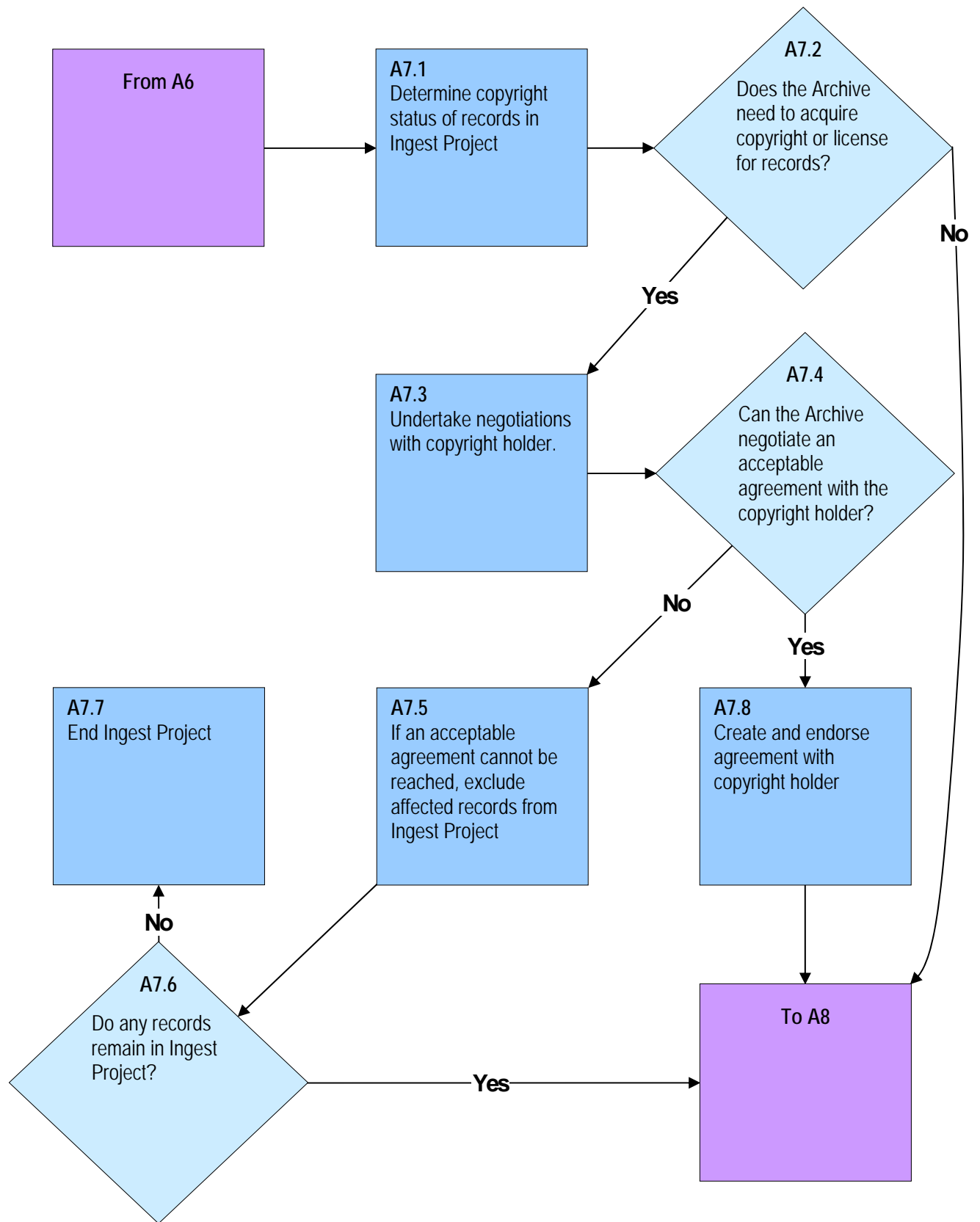


SECTION A: NEGOTIATE SUBMISSION AGREEMENT

Part A7: Assess Copyright

Overview

In this Part the Archive determines the copyright status of the records that it should accession and the copyright status of any associated software. In particular, the Archive determines if it already has the copyright of the records, or needs to obtain the copyrights or a licensing agreement for the records in an Ingest Project. This will allow the Archive to determine the impact copyright has on the feasibility of preserving the records it should accession.

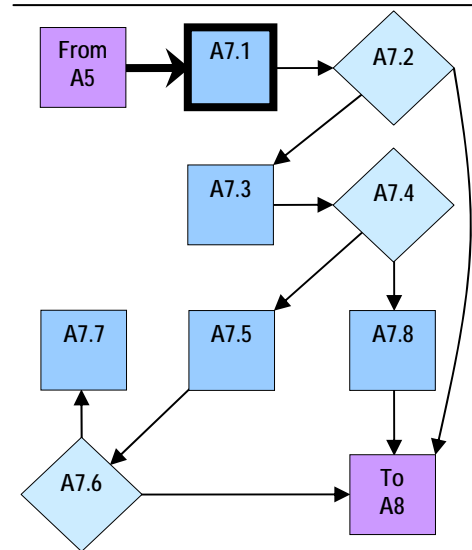


A7.1

Description Based on the information gathered in the Survey Report, the Archive determines the copyright status of the records it should accession. Consideration must be given not only to the copyright of the records' themselves, but also to the copyright of any associated software.

Uses Survey Report

Produces/Modifies Copyright Status

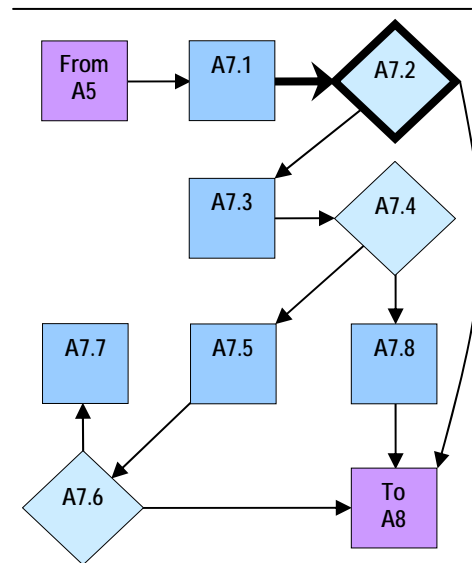


A7.2

Description Based on its finding in Step A7.1, the Archive should determine if it needs to acquire the copyright of the records and/or associated software through a legal transfer or acquire a license for the use of the records and/or associated software from the copyright holder.

Uses Copyright Policy, Copyright Status

Produces/Modifies None

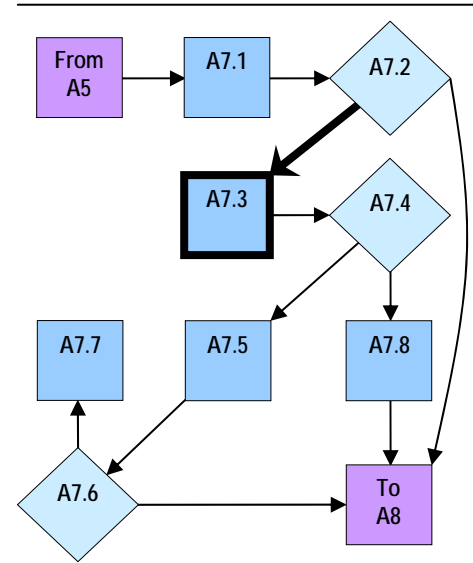


A7.3

Description If the Archive needs to acquire the copyright or licensing rights for at least some of the records and/or associated software it should accession, the Archive should initiate a negotiation process with the rights holder, which may or may not be the Producer. The Archive should use its Copyright Policy as its guide in these negotiations.

Uses Copyright Policy

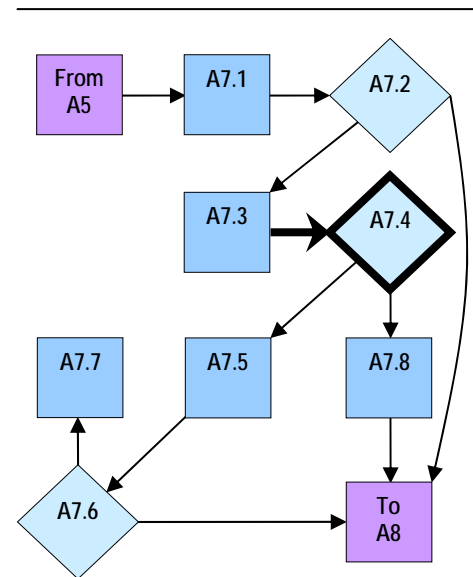
Produces/Modifies None

**A7.4**

Description The Archive determines if it can negotiate a rights transfer or licensing agreement with the rights holder, which it finds reasonable and with which it is capable of complying. Clearing and/or obtaining rights is likely to be resource intensive and difficult. The Archive should make this determination based on its Copyright Policy.

Uses Copyright Policy

Produces/Modifies None

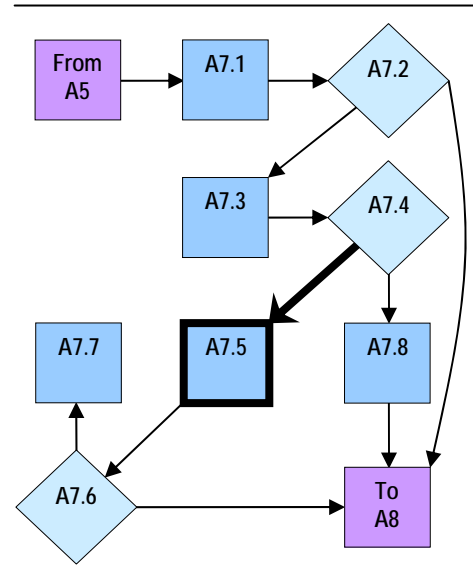


A7.5

Description If the Archive needs to acquire the copyright or licensing rights to at least some of the records and/or associated software it should accession, and the Archive determines it cannot negotiate a reasonable rights transfer or licensing agreement, it should exclude the affected records from the Ingest Project and modify the Survey Report accordingly.

Uses None

Produces/Modifies Survey Report

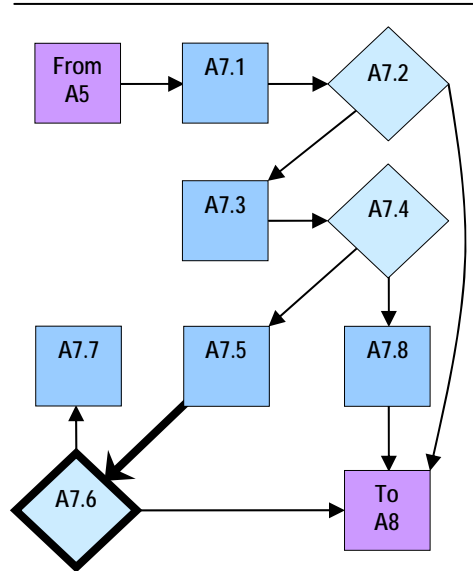


A7.6

Description Based on its actions in Step A7.5, the Archive determines if any records remain in the Ingest Project.

Uses Survey Report

Produces/Modifies None

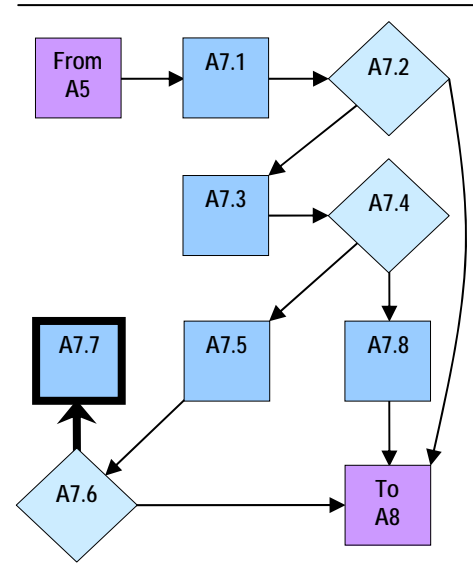


A7.7

Description If no records remain in the Ingest Project, the Archive ends the Project.

Uses None

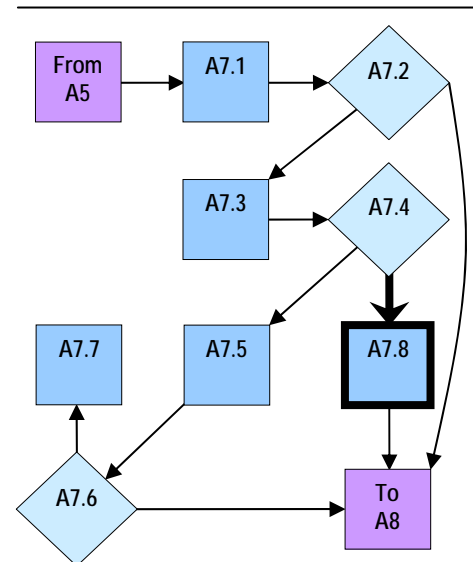
Produces/Modifies Ingest Project Termination Notice

**A7.8**

Description If the Archive determines it can negotiate a reasonable copyright transfer or licensing agreement for at least some of the records in the Ingest Project and/or their associated software; it creates and endorses the agreement it has negotiated with the copyright holder.

Uses None

Produces/Modifies Copyright Transfer/License

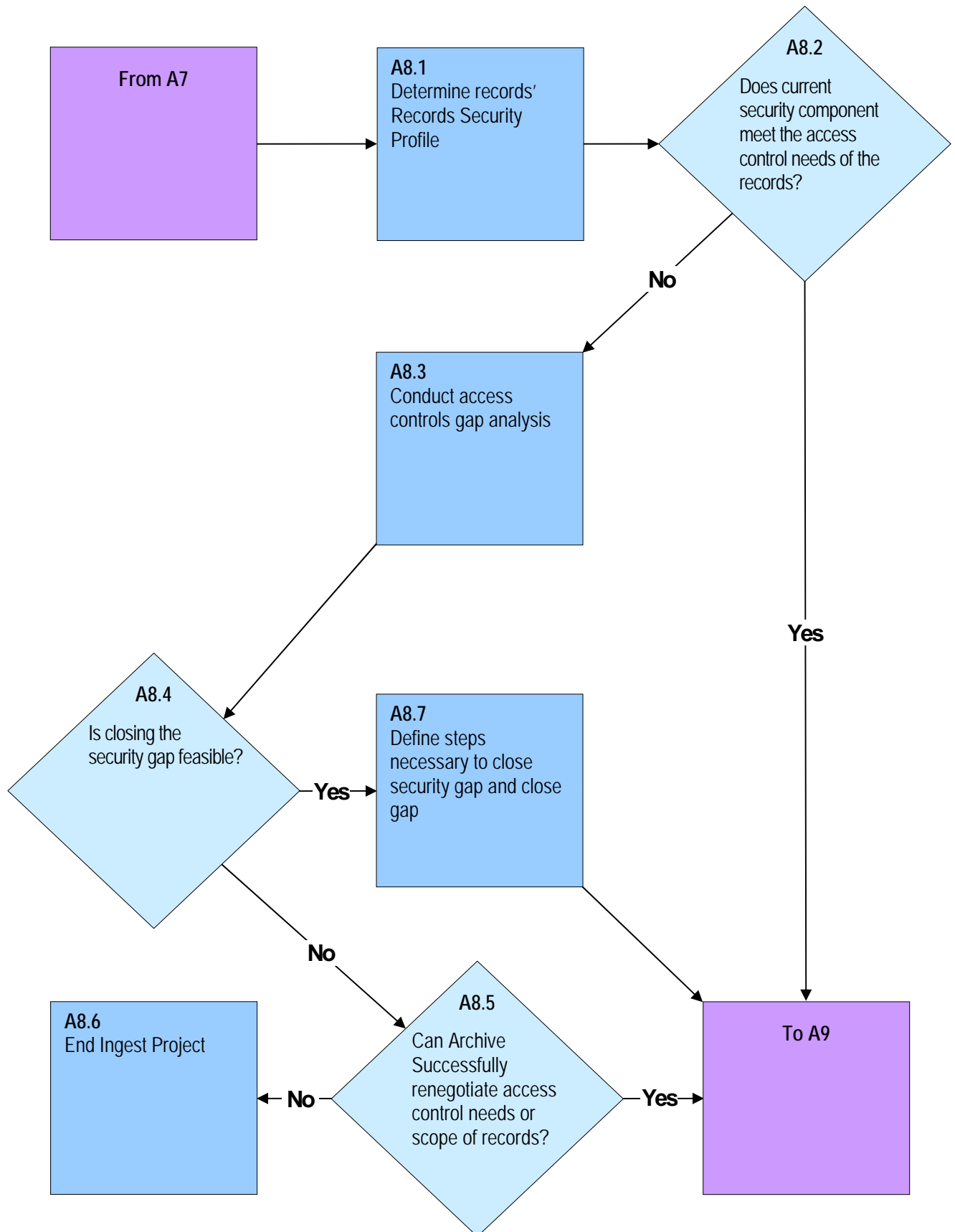


SECTION A: NEGOTIATE SUBMISSION AGREEMENT

Part A8: Assess Access Rights

Overview

In this part, the Archive determines the appropriate Record Security Profile for the records in an Ingest Project. The Archive then determines if the security component of its Preservation System meets the access restriction requirements of the records earmarked for preservation. If the security system does not meet those requirements, the Archive determines if it is feasible to upgrade the security system to meet those requirements. If it is feasible, the Archive will upgrade the system; if it is not feasible, the Archive will not accept the records that require a level of security beyond what the Preservation System can provide.

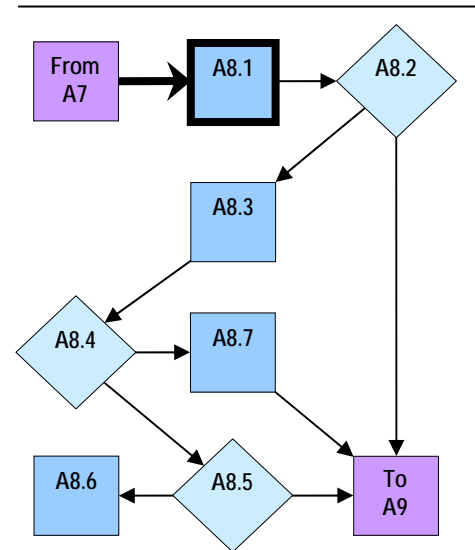


A8.1

Description Based on the information gathered in the records survey, along with its own Access Controls Policy, the Archive determines the appropriate Record Security Profile for the records in an Ingest Project. A Record Security Profile articulates the access control needs of a record. For some records, the Archive may have to create a new Record Security Profile.

Uses Survey Report, Access Controls Policy, Record Security Profile

Produces/Modifies Record Security Profile Decision, Record Security Profile

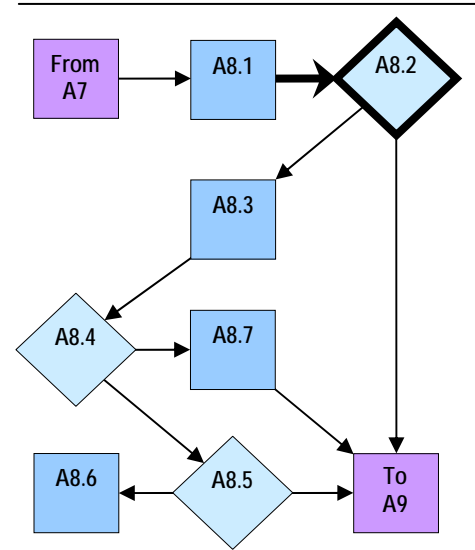


A8.2

Description The Archive determines if the current security component of its Preservation System meets the access control needs of the records in an Ingest Project.

Uses Access Controls Policy, Record Security Profile, Preservation System Capabilities Report

Produces/Modifies None

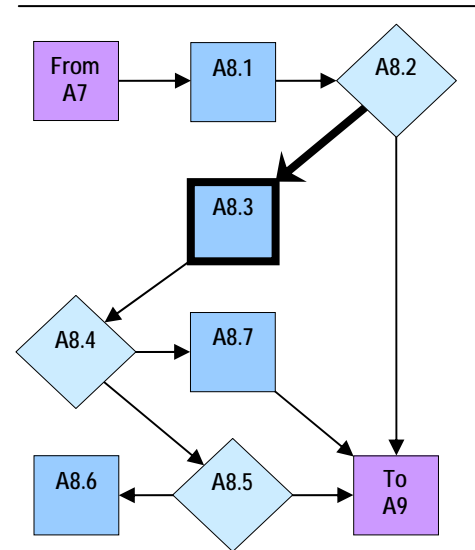


A8.3

Description If the Archive’s security system does not meet the access control needs of the records in an Ingest Project, the Archive conducts a gap analysis of the security system and the records’ access control needs to determine what improvements its security system would need to close the gap.

Uses Access Controls Policy, Record Security Profile, Preservation System Capabilities Report

Produces/Modifies Access Controls Gap Analysis

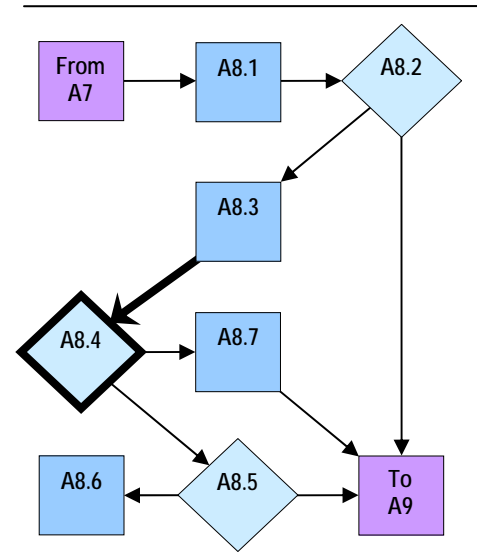


A8.4

Description The Archive determines if it is feasible to take the steps necessary to close the security gap as defined by the findings of A8.3

Uses Access Controls Gap Analysis

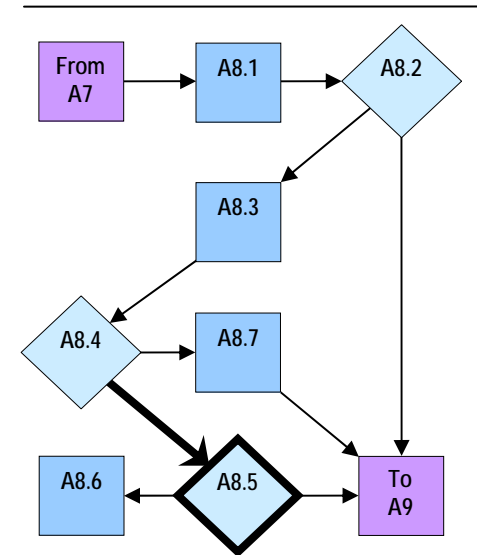
Produces/Modifies Access Controls Gap Analysis Feasibility Statement

**A8.5**

Description: If the Archive determines that it is not feasible to close the security gap, then the Archive must renegotiate with the Producer to define a different set access control needs or a different scope for the Ingest Project, which excludes the affected records.

Uses Access Controls Gap Analysis, Survey Report

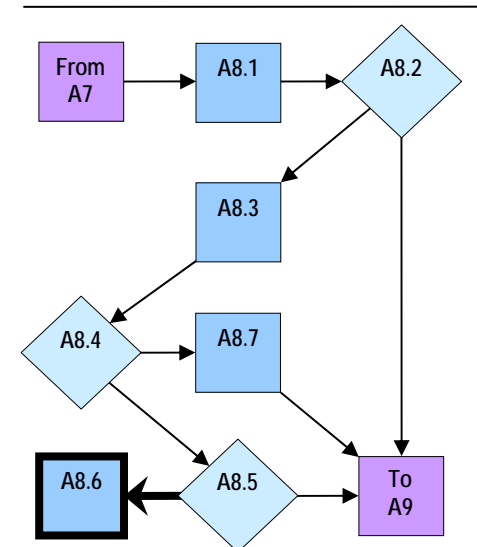
Produces/Modifies Survey Report, Record Security Profile

**A8.6**

Description If the Producer and the Archive are not able to renegotiate satisfactorily, either the Producer or the Archive ends the Project.

Uses None

Produces/Modifies Ingest Project Termination Notice

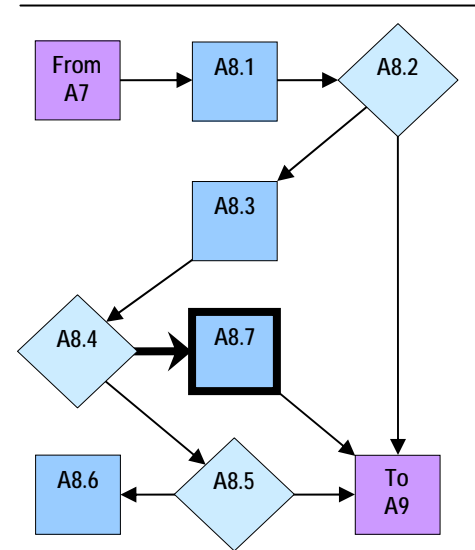


A8.7

Description If the Archive determines that it is feasible to take the steps necessary to close the security gap, then the Archive will define those steps and close the gap.

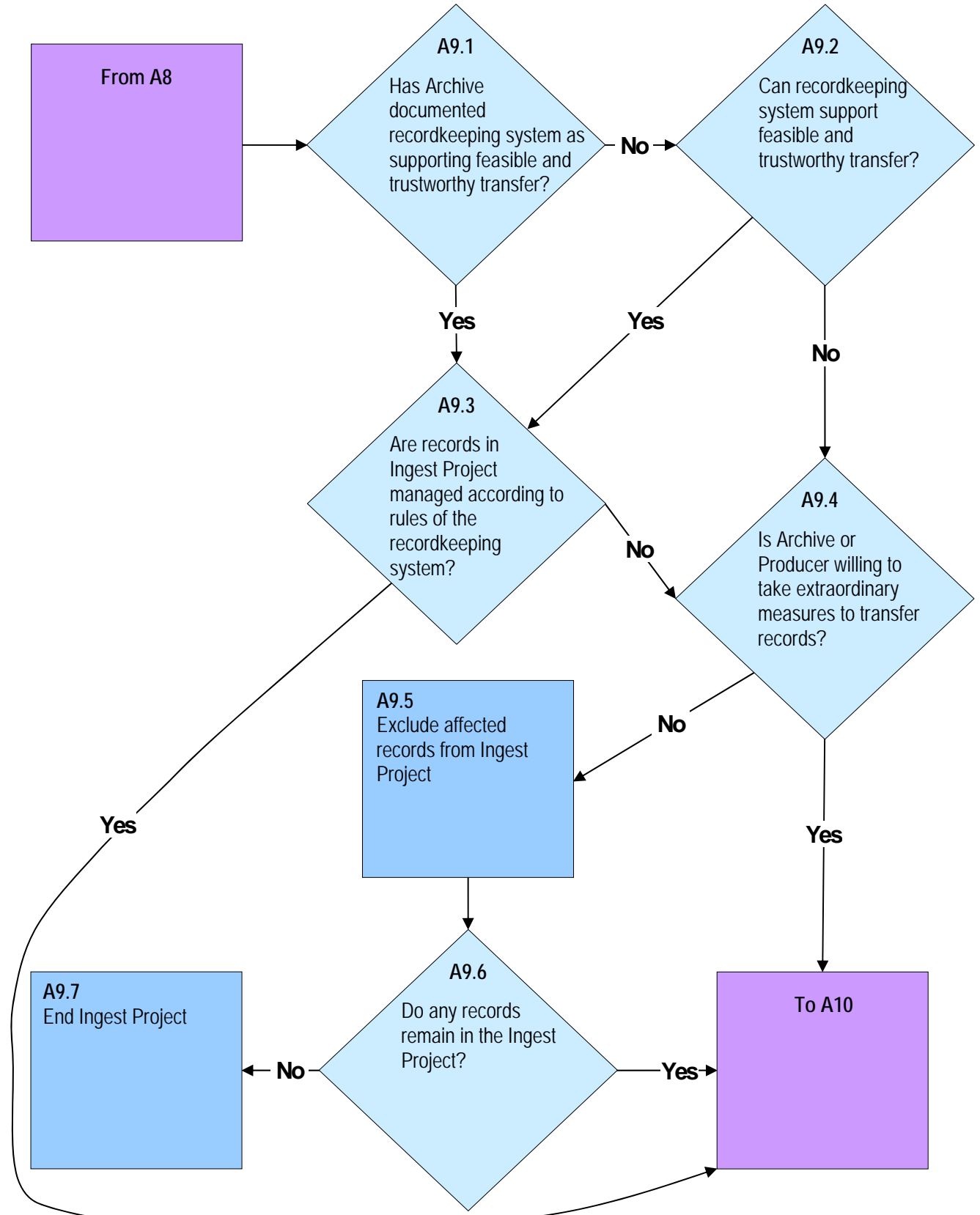
Uses Access Controls Gap Analysis, Access Controls Gap Analysis Feasibility Statement, Access Controls Policy

Produces/Modifies Access Controls Policy



SECTION A: NEGOTIATE SUBMISSION AGREEMENT**Part A9: Assess Recordkeeping System****Overview**

In this part the Archive determines if the active recordkeeping system managing the records in the Ingest Project allows, or is capable of enabling, a trustworthy transfer of records to the Archive in a feasible, scaleable manner. This assessment is entirely separate from any other assessment of the records themselves. Even if the records are authentic and are stored in acceptable record types and formats, and even if there are no problems with existing identifier rules, copyrights, or access rights, it may still simply be too difficult or expensive to facilitate a trustworthy transfer from the recordkeeping system to the preservation system. The Archive should base this determination on the “Requirements for Trustworthy University Electronic Records” or some other set of requirements for trustworthy recordkeeping systems (a Recordkeeping System Evaluation Tool) that evaluate the systems’ ability to transfer records to preservation systems. The Archive may forgo this evaluation if it has already identified the capabilities of the recordkeeping system in a Recordkeeping System Report. If the Archive evaluates a new or modified system, it should document the system in a new or updated Recordkeeping System Report. Determining if records can be transferred in a feasible, scaleable, and trustworthy manner also includes checking that the records are managed according to the rules of the recordkeeping system.

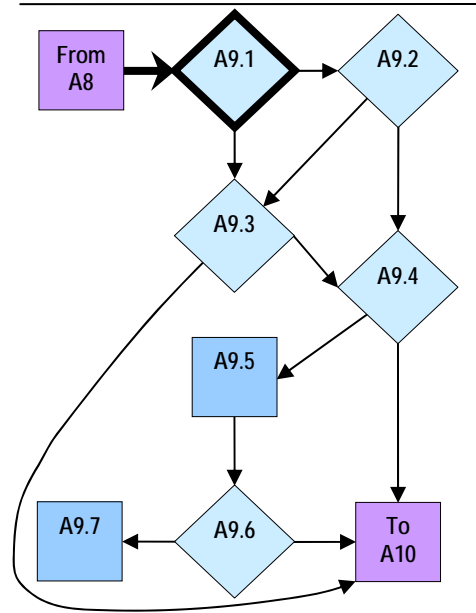


A9.1

Description The Archive determines if has already approved the recordkeeping system managing the records in the Ingest Project as a system from which it can transfer records to the Archive in a feasible, scaleable, and trustworthy manner.

Uses Recordkeeping System Report

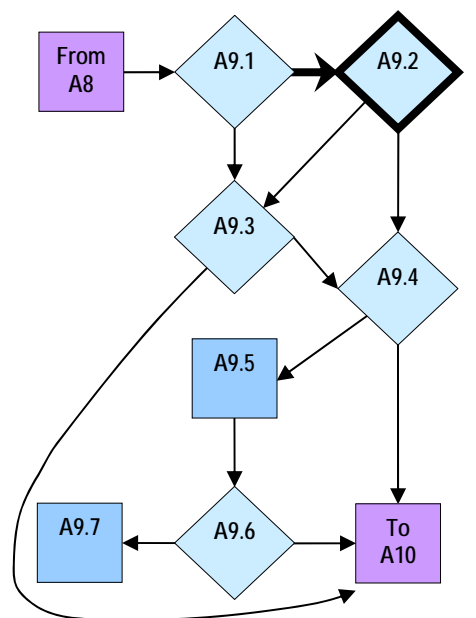
Produces/Modifies None

**A9.2**

Description The Archive evaluates and determines the recordkeeping system's ability to transfer records to the Archive in a feasible, scaleable, and trustworthy manner.

Uses Recordkeeping System Evaluation Tool

Produces/Modifies Recordkeeping System Report

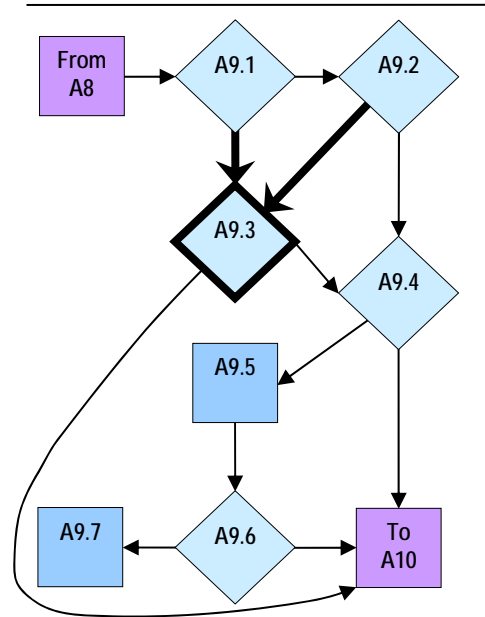


A9.3

Description The Archive determines if the records in the Ingest Project are managed in the recordkeeping system according to the rules of the recordkeeping system.

Uses Recordkeeping System Internal Rules

Produces/Modifies None

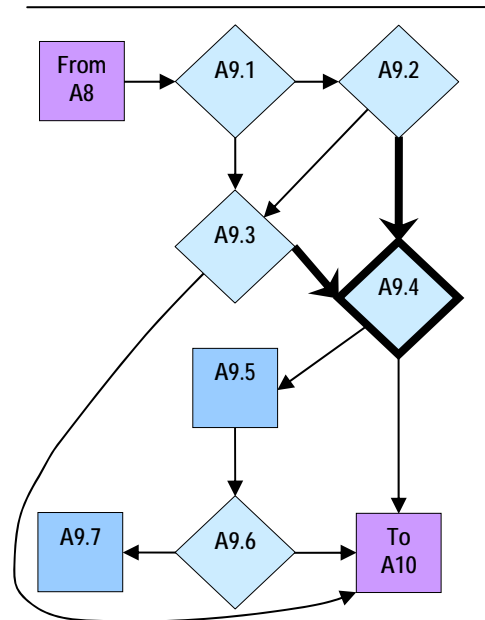


A9.4

Description The Archive and the Producer determine if either or both are willing to take extraordinary measures to transfer the records in the Ingest Project to the Archive in a trustworthy manner. These measures may range from building special tools to undertaking a software reengineering project. Such measures are rarely scaleable efforts that can be turned into regular procedures.

Uses None

Produces/Modifies SIP Creation Procedures

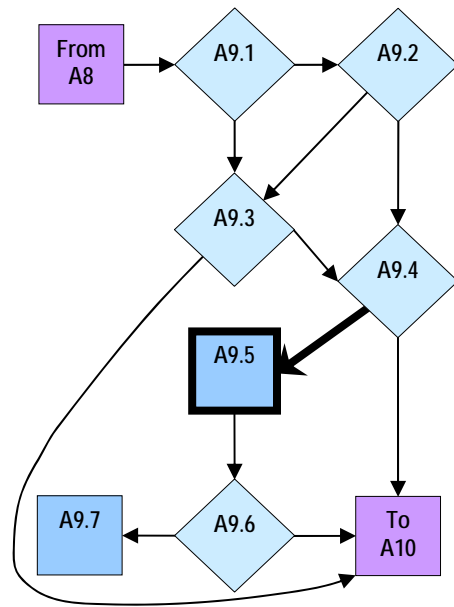


A9.5

Description The Archive excludes from the Ingest Project the records that it and the Producer cannot or are not willing to make the effort to transfer to the Archive in a trustworthy manner. The Archive should modify the Survey Report accordingly.

Uses None

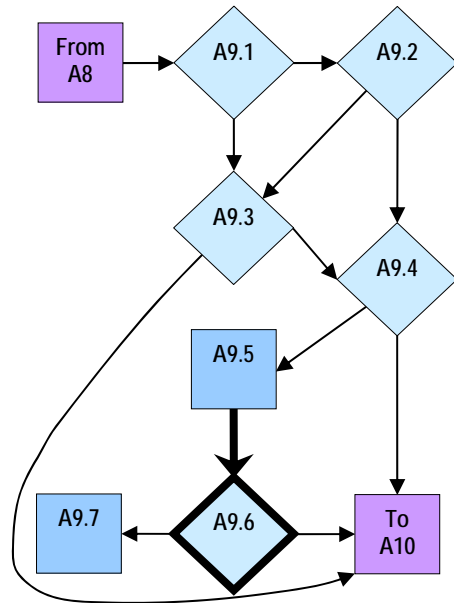
Produces/Modifies Survey Report

**A9.6**

Description The Archive determines if any records remain in the Ingest Project after Step A9.6.

Uses Survey Report

Produces/Modifies None

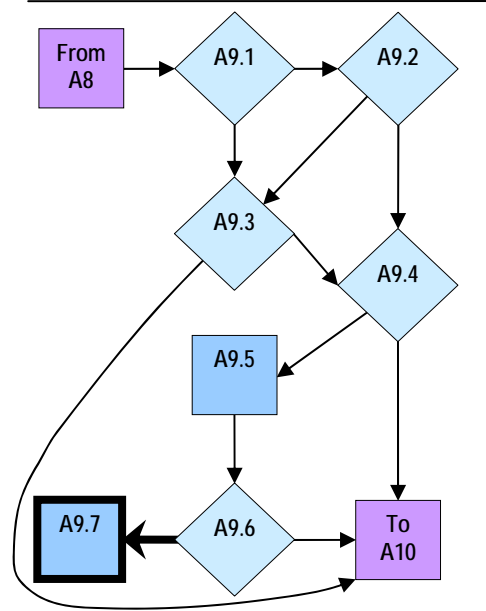


A9.7

Description If no records remain in its Ingest Project, the Archive ends the Project.

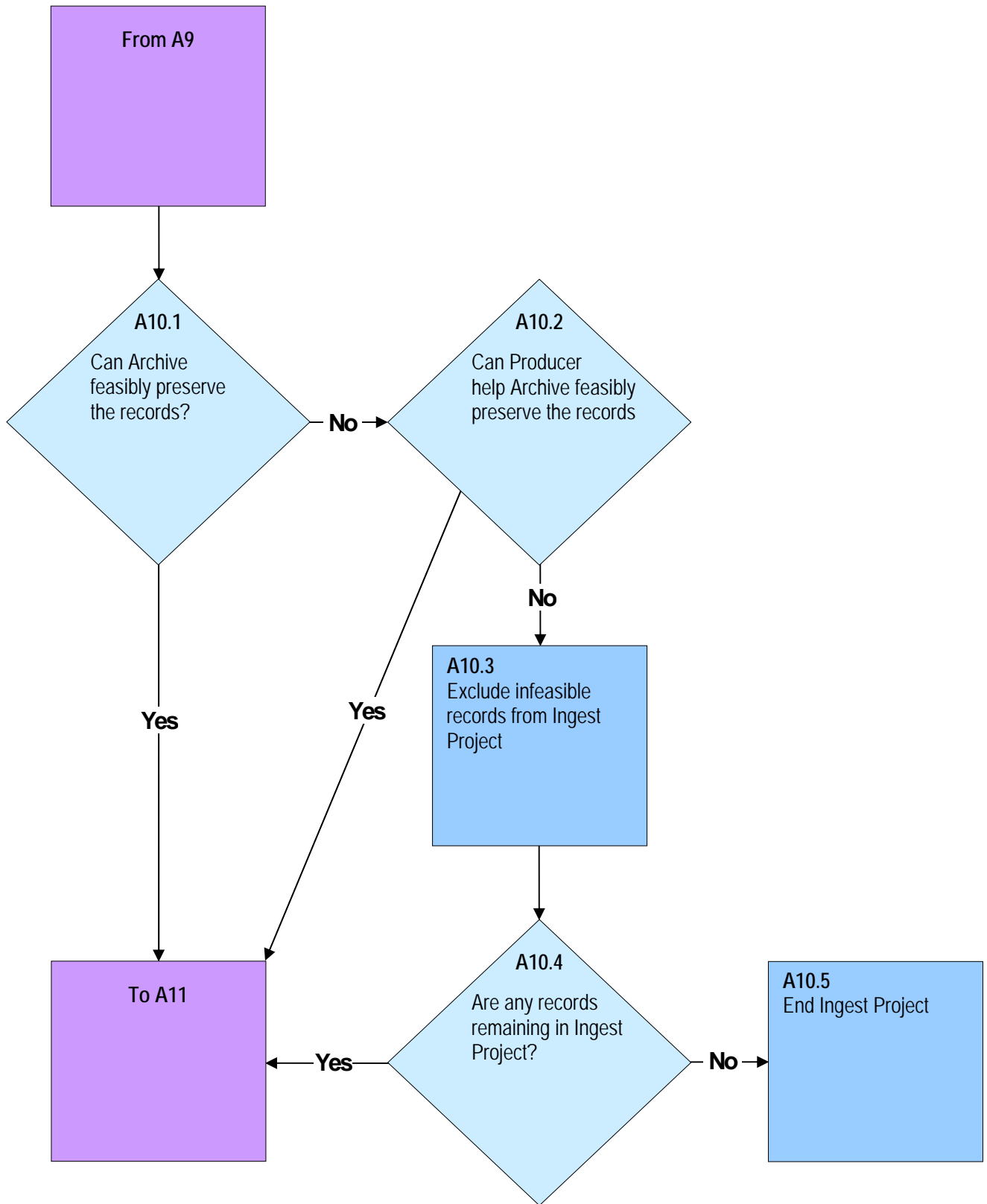
Uses None

Produces/Modifies Ingest Project Termination Notice



SECTION A: NEGOTIATE SUBMISSION AGREEMENT**Part A10: Assess Feasibility****Overview**

During this Part, the Archive assesses if it can accession, manage, and preserve the records it should accession in its Preservation System, either on its own, or with help from the Producer or a third party. This assessment is based on the information collected in the previous parts, including Assess Record Types, Assess Formats, Assess Identifier Rules, Assess Copyright, Assess Access Rights, and Assess Recordkeeping System. The Archive must determine if its existing resources for preservation formats, record types, identifier rules, creator records, security procedures, and system capabilities meets the needs of the records identified in Part A3. The resulting feasibility report should present a gap analysis if the Resources do not reflect the continuing value of any records assessed in A3. The Archive must then determine if it should modify or add to its Resources to meet those assessments or if it should reject or modify the scope of the records involved in the Ingest Project.

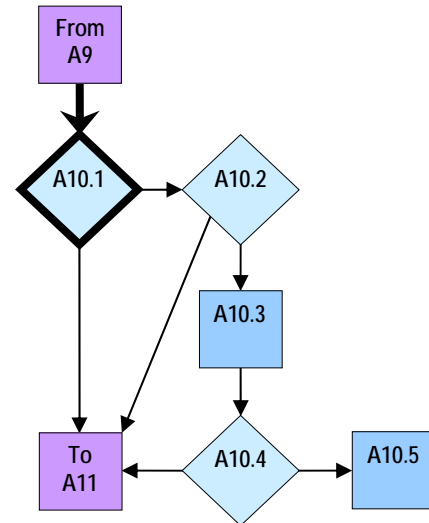


A10.1

Description The Archive assesses if it has the technical and staffing capacity or can draw on the resources of a third party to manage and preserve the records earmarked for preservation. The Archive documents its assessment in a Preservation System Availability Statement.

Uses Survey Report, Preservation System Capabilities Report

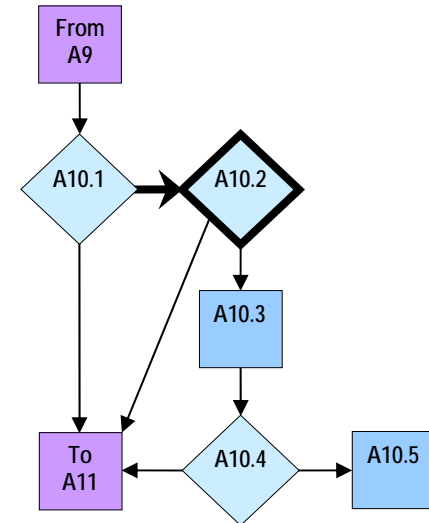
Produces/Modifies Preservation System Availability Statement

**A10.2**

Description If the Archive assesses that it cannot manage at least some of the records it should accession it asks the Producer if it can help the Archive make feasible the transfer of these records. If the Producer can help by producing resources, the Archive amends the System Availability Report to describe the resources that the Producer agrees to provide.

Uses Survey Report, Preservation System Capability Report

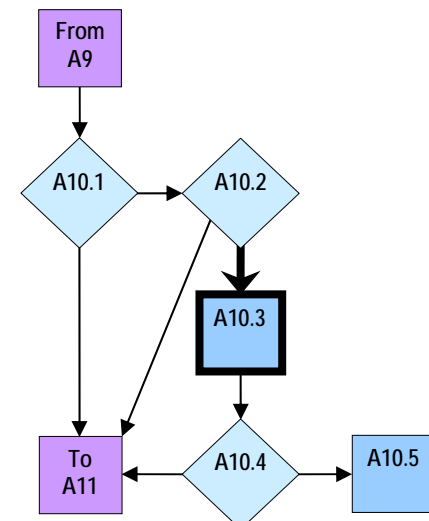
Produces/Modifies Preservation System Availability Statement

**A10.3**

Description If the Producer is unable or unwilling to help the Archive, then the Archive will exclude from the Ingest Project the records it cannot transfer, adjusting its Preservation System Availability Statement and Survey Report accordingly.

Uses Survey Report

Produces/Modifies Preservation System Availability Statement, Survey Report

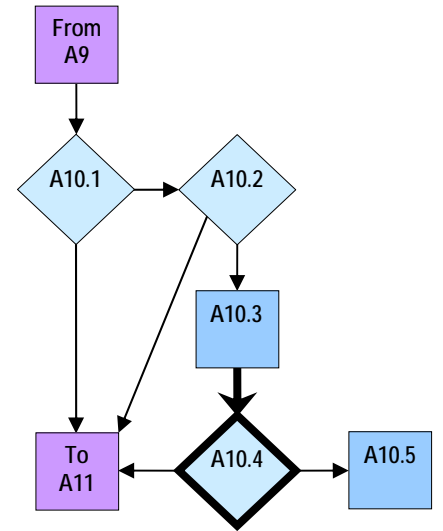


A10.4

Description Based on its actions in Step A10.3, an Archive determines if any records remain in its Ingest Project.

Uses Survey Report

Produces/Modifies None

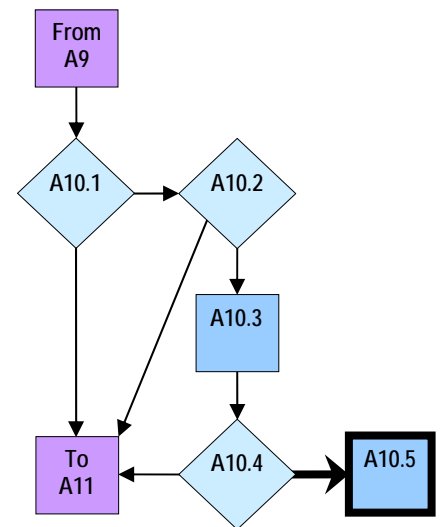


A10.5

Description If no records remain in its Ingest Project, the Archive ends the Project.

Uses None

Produces/Modifies Ingest Project Termination Notice

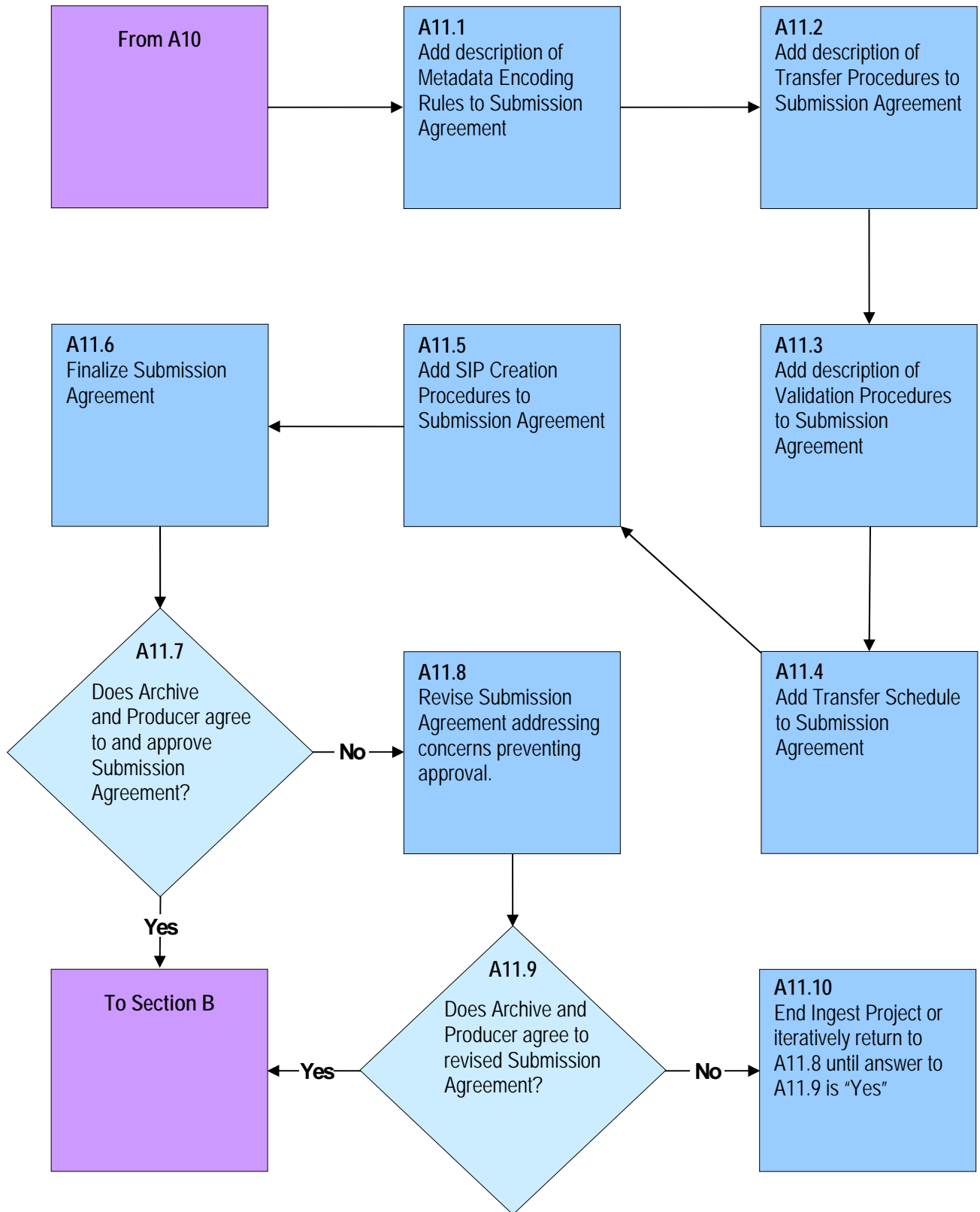


SECTION A: NEGOTIATE SUBMISSION AGREEMENT

Part A11: Finalize Submission Agreement

Overview

During this Part, the Archive adds description of Metadata Encoding Rules, Transfer Procedures and Schedules, and Validation Procedures to the Submission Agreement. Then the Archive and Producer work on finalizing the Submission Agreement until they both agree to endorse it.

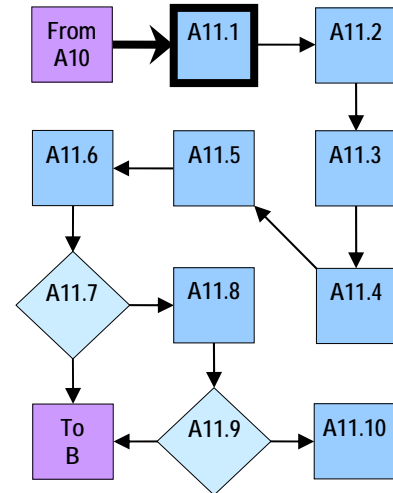


A11.1

Description The Archive attaches a description of its Metadata Encoding Rules to the Submission Agreement.

Uses Metadata Encoding Rules

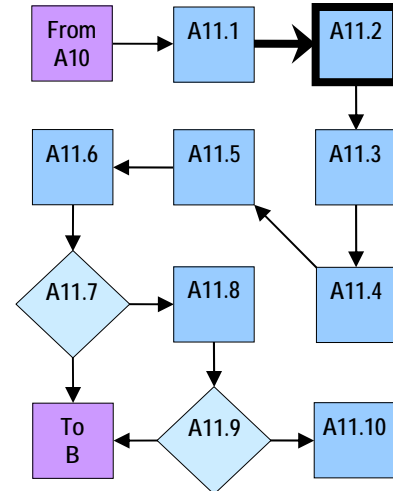
Produces/Modifies Metadata Encoding Rules Decision

**A11.2**

Description The Archive attaches a description of its Transfer Procedures Rules to the Submission Agreement.

Uses Transfer Procedures

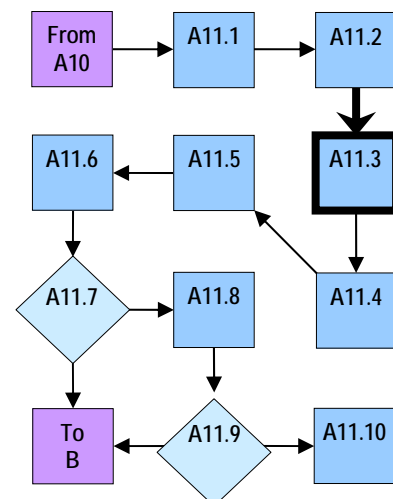
Produces/Modifies Transfer Procedures Decision

**A11.3**

Description The Archive attaches a description of its Validation Procedures to the Submission Agreement.

Uses Validation Procedures

Produces/Modifies Validation Procedures Decision

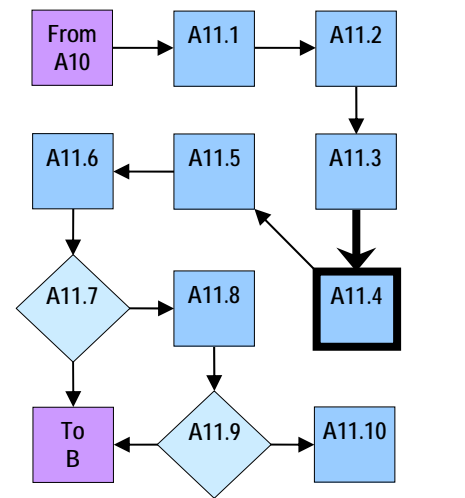


A11.4

Description The Archive determines a schedule for transferring the records in an Ingest Project to the Archive and attaches that schedule to the Submission Agreement.

Uses Transfer Procedures

Produces/Modifies Transfer Schedule

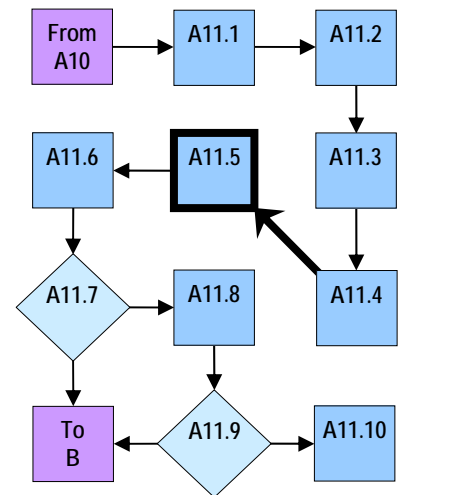


A11.5

Description The Archive attaches a description of its Submission Information Package (SIP) Creation Procedures to the Submission Agreement.

Uses SIP Creation Procedures

Produces/Modifies SIP Creation Procedures Decision

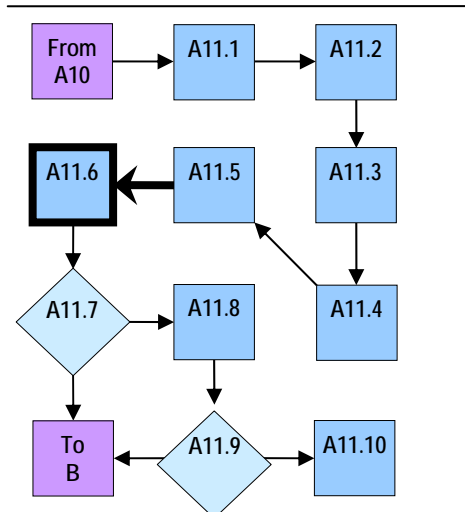


A11.6

Description The Archive takes actions needed to finalize the Submission Agreement that do not substantially change any decisions made earlier in Section A. The Archive constructs an actual draft Submission Agreement, typically using preset language.

Uses None

Produces/Modifies Draft Submission Agreement

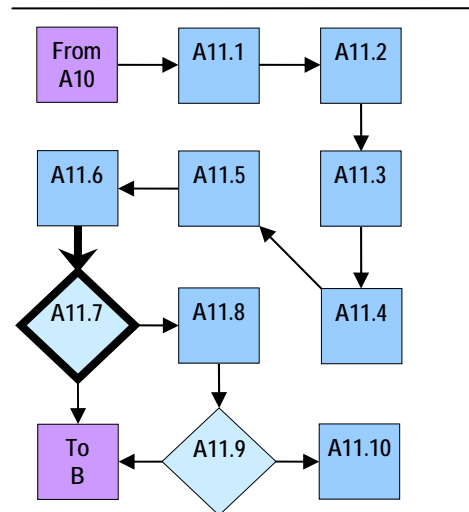


A11.7

Description The Archive and the Producer review the drafted Submission Agreement and determine if they are willing to endorse the Submission Agreement. If both are willing, the Archive and Producer endorse the Submission Agreement. This completes Section A: Negotiate Submission Agreement of the Ingest Guide. Proceed to Section B: Transfer and Validation to complete the Ingest Project.

Uses Draft Submission Agreement

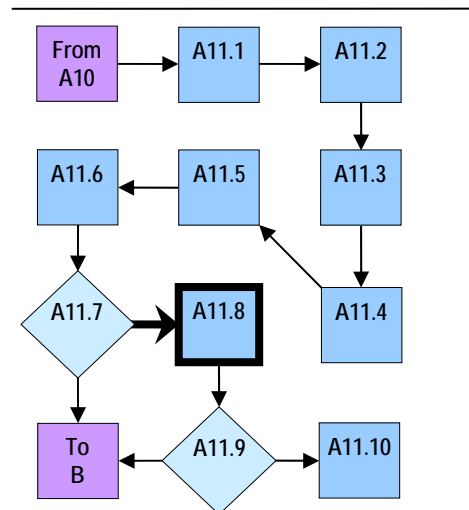
Produces/Modifies Finalized and Endorsed Submission Agreement

**A11.8**

Description If either the Archive or the Producer is unwilling to endorse the Submission Agreement as it is drafted, the Archive, usually in conjunction with the Producer, will revise the Submission Agreement to address the Archive's or Producer's concerns.

Uses Draft Submission Agreement

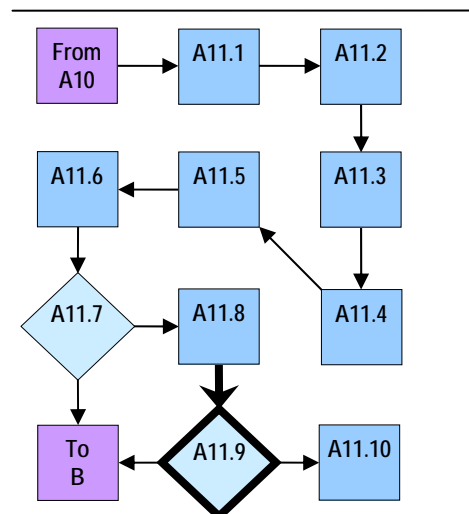
Produces/Modifies Draft Submission Agreement

**A11.9**

Description The Producer and the Archive determine if they are willing to endorse the Submission Agreement with the revisions made in Step A11.8. If they are willing to endorse the Submission Agreement, they will do so. This completes Section A: Negotiate Submission Agreement of the Ingest Guide. Proceed to Section B: Transfer and Validation to complete the Ingest Project.

Uses Draft Submission Agreement

Produces/Modifies Finalized and Endorsed Submission Agreement

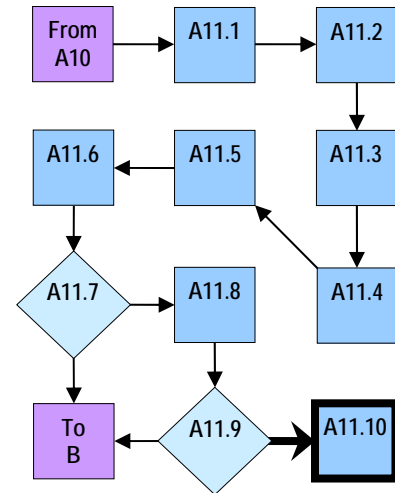


A11.10

Description If either the Producer or the Archive is still not willing to sign the revised submission agreement in Step A11.8, return to Step A11.8 to revise the Agreement again. Repeat this process until the Producer and the Archive agree to sign the Submission Agreement. If the Producer and the Archive cannot agree on a finalized version of a Submission Agreement, the Archive will stop the Ingest Project.

Uses Draft Submission Agreement

Produces/Modifies Ingest Project Termination Notice



SECTION B: TRANSFER AND VALIDATION

Overview

The Transfer and Validation section of the Ingest Guide describes the actions needed for the Archive and a Producer to deposit records into a preservation system. This portion of the Ingest Guide describes the actual transfer, validation, and transformation work of an Ingest Project. The steps for transfer, validation, and transformation are defined by Submission Agreements created in Step A.

This section is composed of six parts. During Part B1 the Producer packages the appropriate records in a Submission Information Package (SIP) with the proper metadata, as stipulated in the Submission Agreement. Once the Archive has received the SIP during Part B2, Automated Validation, it must verify the integrity, completeness, and correctness of the transfer and validate that the transferred records conform to the requirements of the Submission Agreement and to technical file format standards. This includes validating the SIP against the requirements of the Submission Agreement. Then in Part B3, Transform and Attach Metadata, the Archive transforms the records and attaches any needed metadata as prescribed by the Submission Agreement. In Part B4, AIP Formation, the Archive assembles the records involved in the Ingest Project into Archival Information Packages (AIPs). In Part B5, Final Appraisal, the Archive makes a final check of the records and the metadata in the AIP to ensure that they conform to the rules of the Archive and are indeed the records described in the Submission Agreement. Finally in Part B6, Formal Accession, the Archive formally accessions the records into the Preservation System and notifies the Producer of this formal accession.

Part B6 is the final step of the Ingest Guide and is the last step of an Ingest Project.

The Ingest Guide uses the *OAIS* definition of SIP: “An Information Package that is delivered by the Producer to the *OAIS* for use in the construction of one or more AIPs.”⁷ It also uses the *OAIS* definition of AIP: “An Information Package, consisting of the Content Information and the associated Preservation Description Information (PDI), which is preserved within an *OAIS*.”

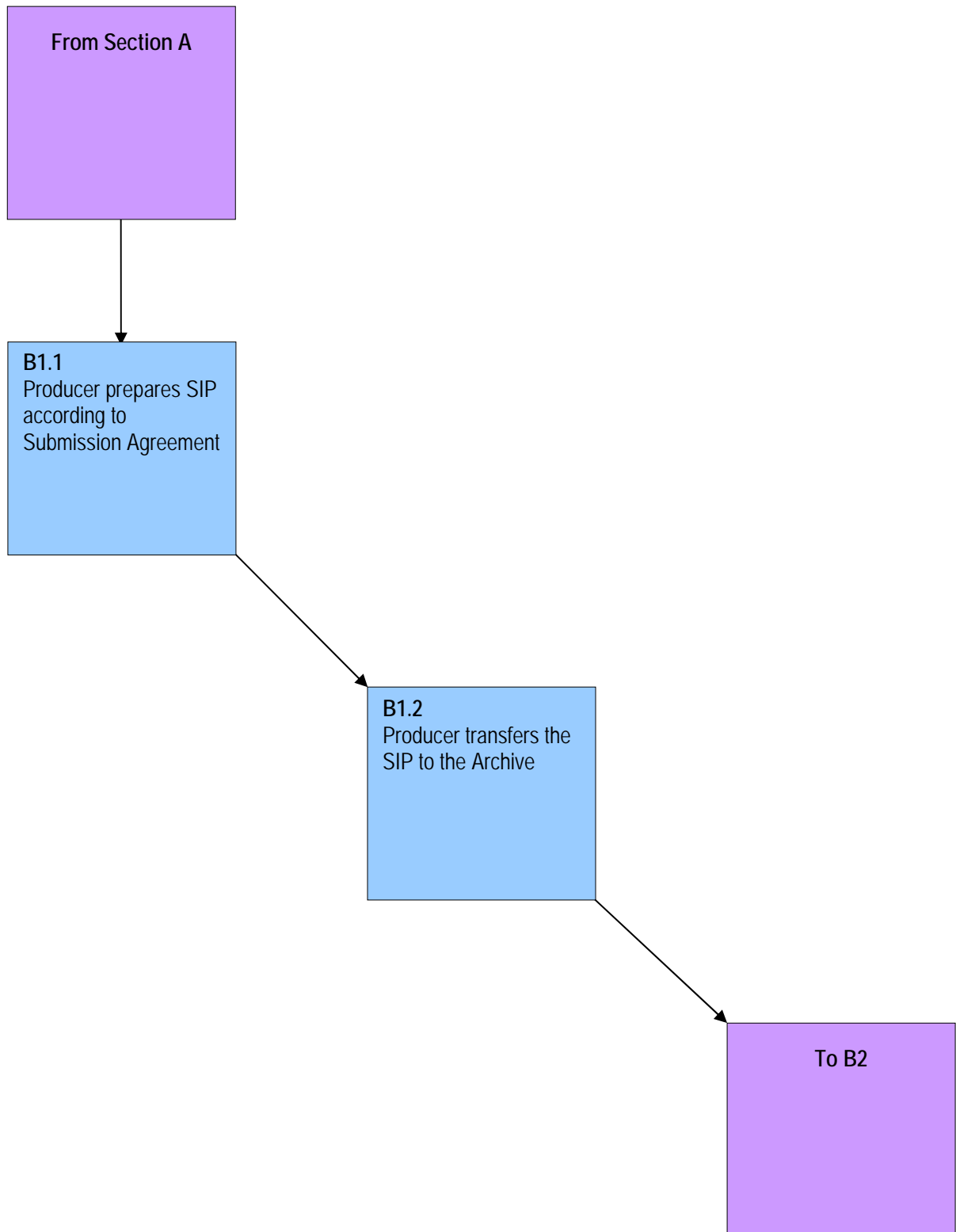
⁷ ISO 14721:2003

SECTION B: TRANSFER AND VALIDATION

Part B1: Create and Transfer SIPs

Overview

During this Part a Producer creates a Submission Information Package (SIP) of the records the Archive will accession according to the terms of a Submission Agreement. A Producer will then transfer the SIP to the Archive.

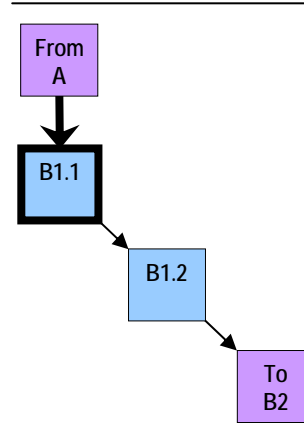


B1.1

Description The Producer prepares the records that it has agreed to transfer in the Submission Agreement by packaging them in a Submission Information Package (SIP) according to the SIP Creation Procedures articulated in the Submission Agreement. The SIP includes any necessary metadata, digital signatures, or Producer-side transformations that the Submission Agreement called for. Creation of the SIP may be technically complex, depending on the number of digital components and the number and complexity of file formats of those components. The Producer may or may not be able to accomplish this task on its own and may require technical assistance from the Archive. The Archive may wish to build a tool set to enable Producers to produce SIPs.

Uses SIP Creation Procedures Decision, SIP Creation Procedure

Produces/Modifies SIP

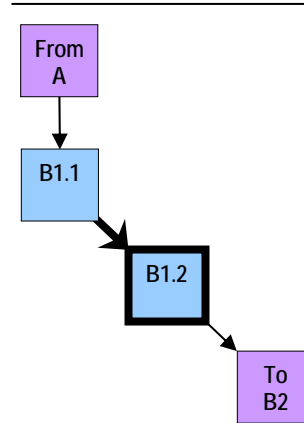


B1.2

Description The Producer transfers the SIP to the Archive. This step represents the actual transfer moment in the ingest process. The transfer may be a physical exchange of storage media containing electronic records or a transfer undertaken over a computer network. None of the steps involved in completing either of these processes is described here.

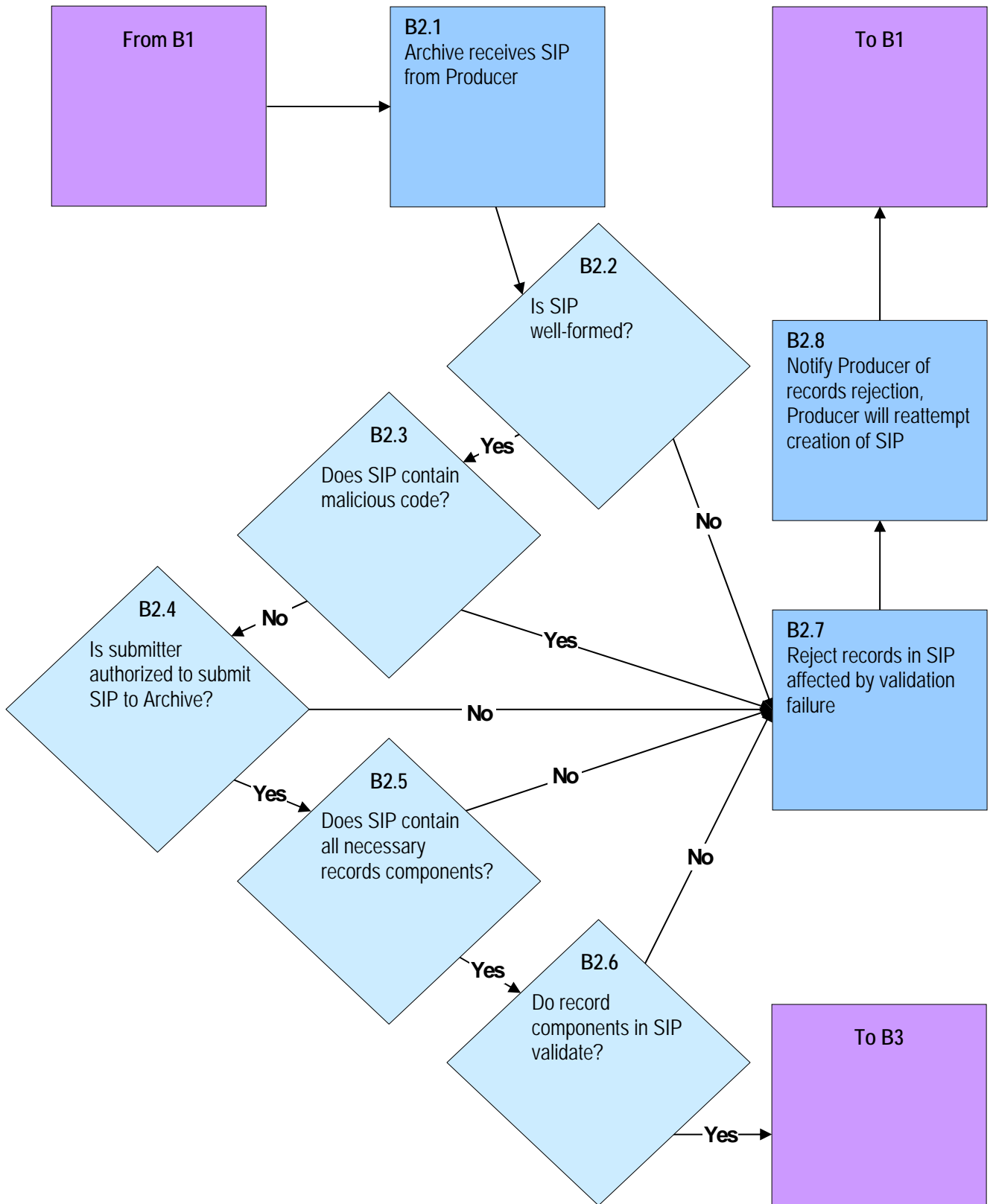
Uses Transfer Schedule

Produces/Modifies None



SECTION B: TRANSFER AND VALIDATION
Part B2: Validate**Overview**

During this Part, the Archive validates the SIP and its included record components received from the Producer. The Archive checks that the SIP and its components are well-formed and whether they contain viruses. It also validates that the Producer was authorized to transfer the SIP and that the SIP conforms to the requirements of the Submission Agreement. If the SIP fails any of these validations, the Archive rejects the SIP and notifies the Producer to generate another SIP. An Archive can carry out all of the steps in this Part in an automated manner. While an Archive does not have to automate these Steps, automation greatly enhances productivity. In practice, an Archive may implement Steps B2.2 through B2.6 in a difference sequence.

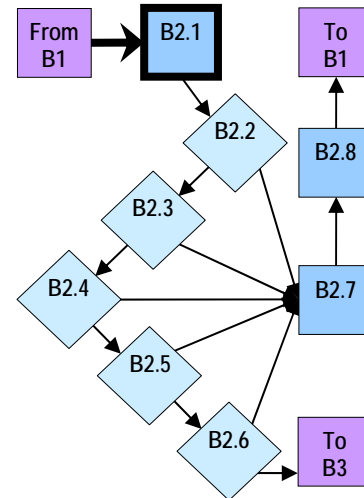


B2.1

Description The Archive receives the SIP from a Producer. The Archive may notify a Producer that it has received the SIP as long as the Archive makes it clear receiving the SIP does not mean accepting it.

Uses Institutional Identity Management System, Producer Record

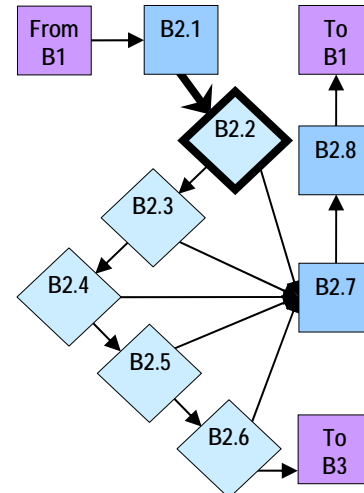
Produces/Modifies Documentation of Receipt

**B2.2**

Description This is the first of five validation steps that the Archive performs on a SIP. In this step, the Archive performs a format check on the SIP package to ensure that it meets the requirements of the Preservation System, as described in the Submission Agreement. This includes confirming the existence of all the records components listed in the SIP manifest.

Uses SIP Creation Procedures

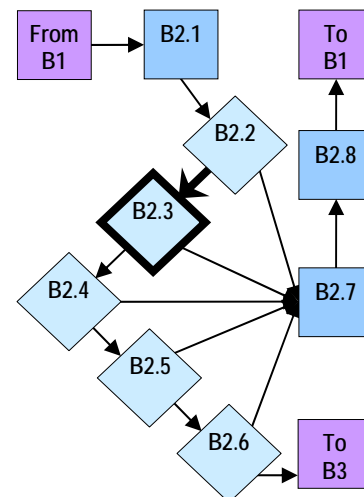
Produces/Modifies SIP Validity Statement

**B2.3**

Description The Archive checks if a SIP contains any malicious code (viruses). Because of the age of the digital components contained in the SIP, the Archive must ensure that malicious code check can recognize very old malicious code.

Uses SIP Creation Procedures, Virus Definition Files

Produces/Modifies SIP Validity Statement

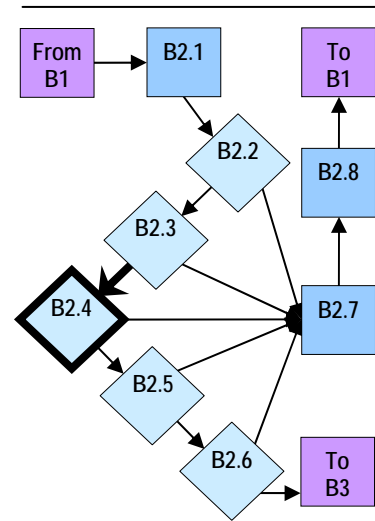


B2.4

Description The Archive checks if the submitter of a SIP is authorized to send the SIP to the Archive.

Uses Producer Entry

Produces/Modifies SIP Validity Statement

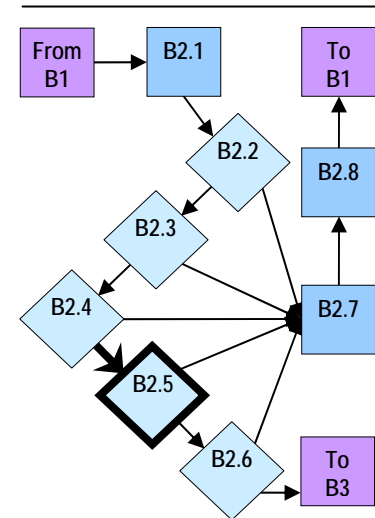


B2.5

Description The Archive checks if the records components in the SIP or SIPs of an Ingest Project properly form the records described in a Submission Agreement.

Uses Transformation Plan, Record Type Record

Produces/Modifies SIP Validity Statement

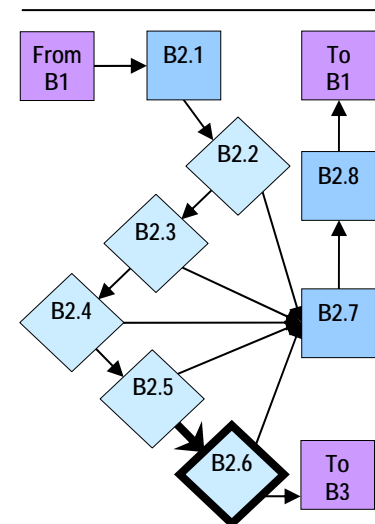


B2.6

Description The Archive validates that the file formats of the records components contained within a SIP conform to technical file format standards.

Uses Transformation Plan, Record Type Record, Format Representation Information System

Produces/Modifies SIP Validity Statement

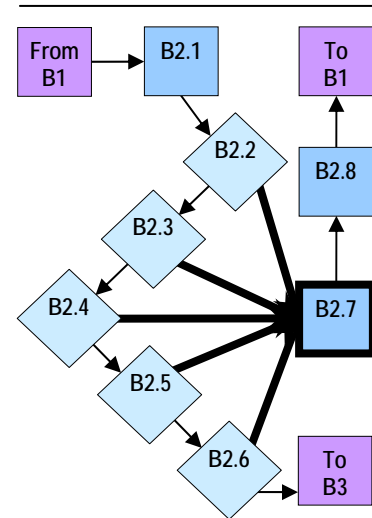


B2.7

Description If the Archive finds any problems with a SIP or its components in Steps B2.2 through B2.6, the Archive rejects all of the records in the SIP affected by the validation failure.

Uses None

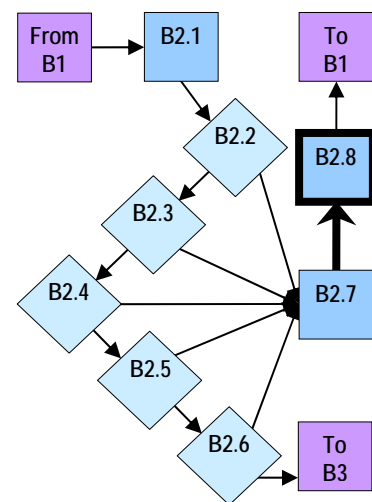
Produces/Modifies None

**B2.8**

Description The Archive notifies a Producer that it has rejected at least some records in a SIP. The Producer will correct the error and repeat Step B1.1, creating another SIP of those records for resubmission.

Uses None

Produces/Modifies SIP Rejection Notification

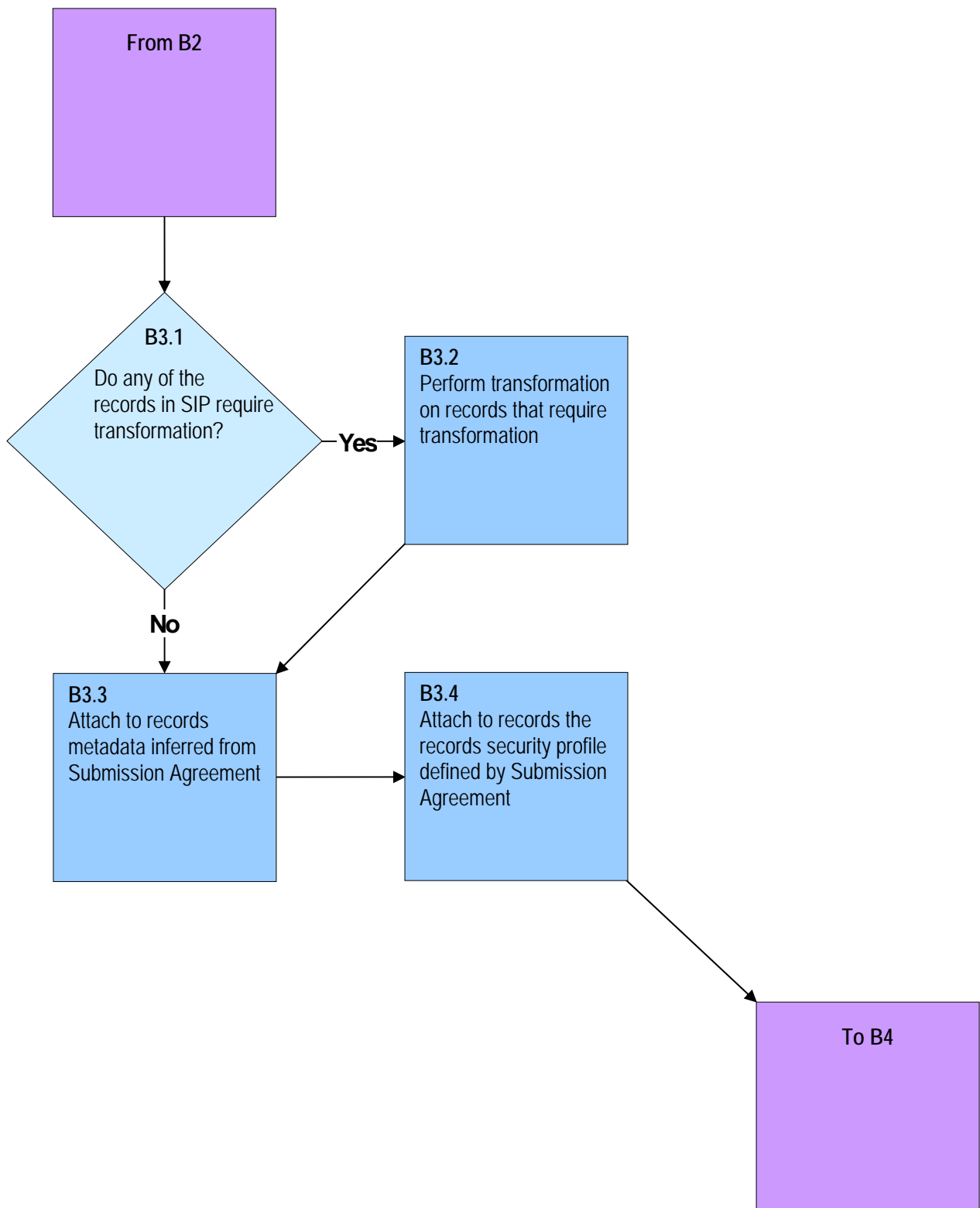


SECTION B: TRANSFER AND VALIDATION

Part B3: Transform and Attach Metadata

Overview

During this Part, the Archive transforms any records in an Ingest Project that require transformation. An Archive also attaches to the records metadata it can automatically infer from the Submission Agreement and its associated documentation. This metadata includes unique identifiers, Records Security Profiles, and information about the Producer, Record Types, formats, and time of transfer. This metadata should allow the Archive to administer the records. The Archive can add descriptive metadata to records after it has completed the Ingest Project.

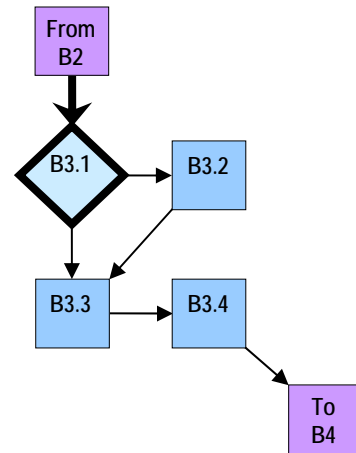


B3.1

Description The Archive determines if any of the records in an Ingest Project requires format transformation. The Archive determines this based on the findings it made in Part A5.

Uses Transformation Plan

Produces/Modifies None

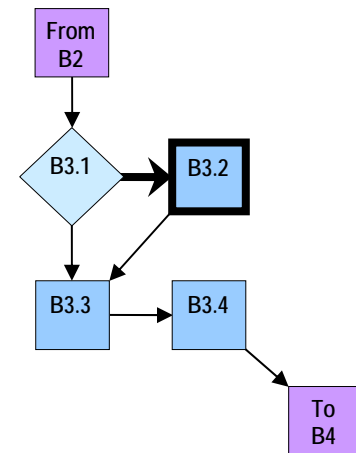


B3.2

Description If the Archive needs to transform any of the records in an Ingest Project, it will transform those records as determined in Part A5.

Uses Transformation Plan

Produces/Modifies Records Transformed

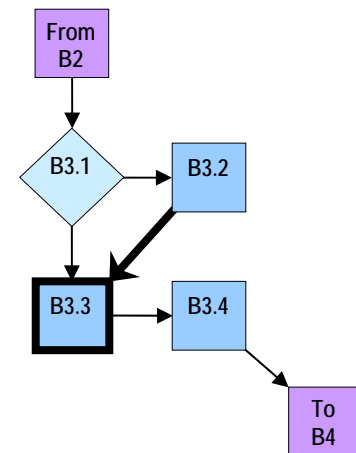


B3.3

Description The Archive attaches metadata it infers from the Submission Agreement to the records. This is an automated application of the metadata whose encoding rules are stipulated in Step A11.1.

Uses Metadata Encoding Rules

Produces/Modifies Records with Attached Metadata

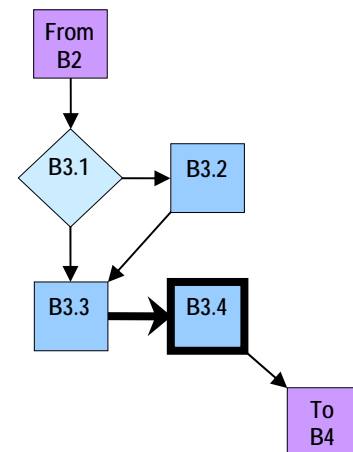


B3.4

Description The Archive attaches the proper records security profile to the records as defined in the Submission Agreement (Part A8).

Uses Records Security Profile

Produces/Modifies Records with Security Profile

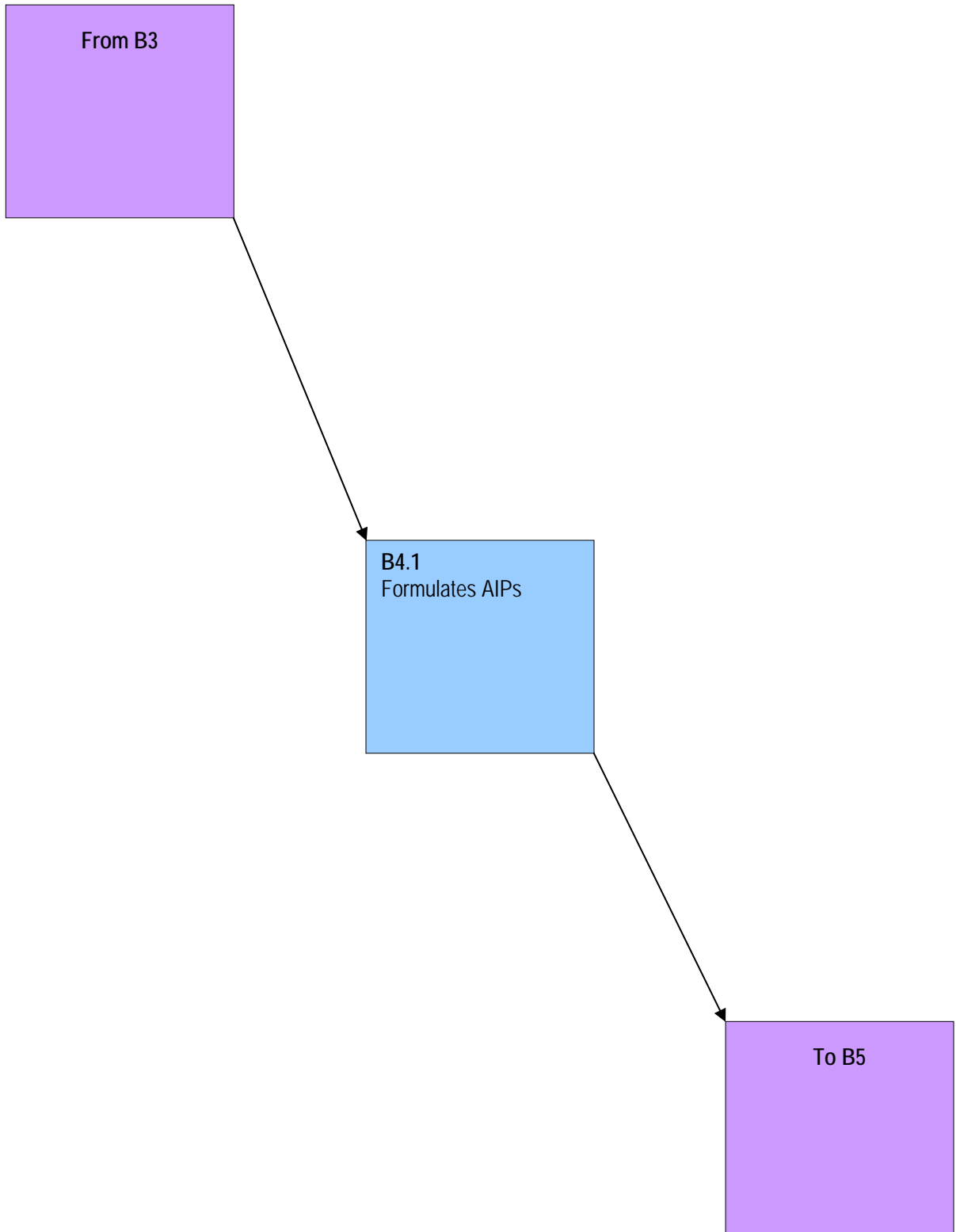


SECTION B: TRANSFER AND VALIDATION

Part B4: Formulate AIPs

Overview

During this Part, the Archive turns the records involved in an Ingest Project into Archival Information Packages (AIP) according to the rules and procedures of the Archive's Preservation System.

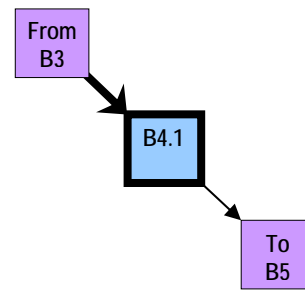


B4.1

Description The Archive turns the records involved in an Ingest Project into AIPs.

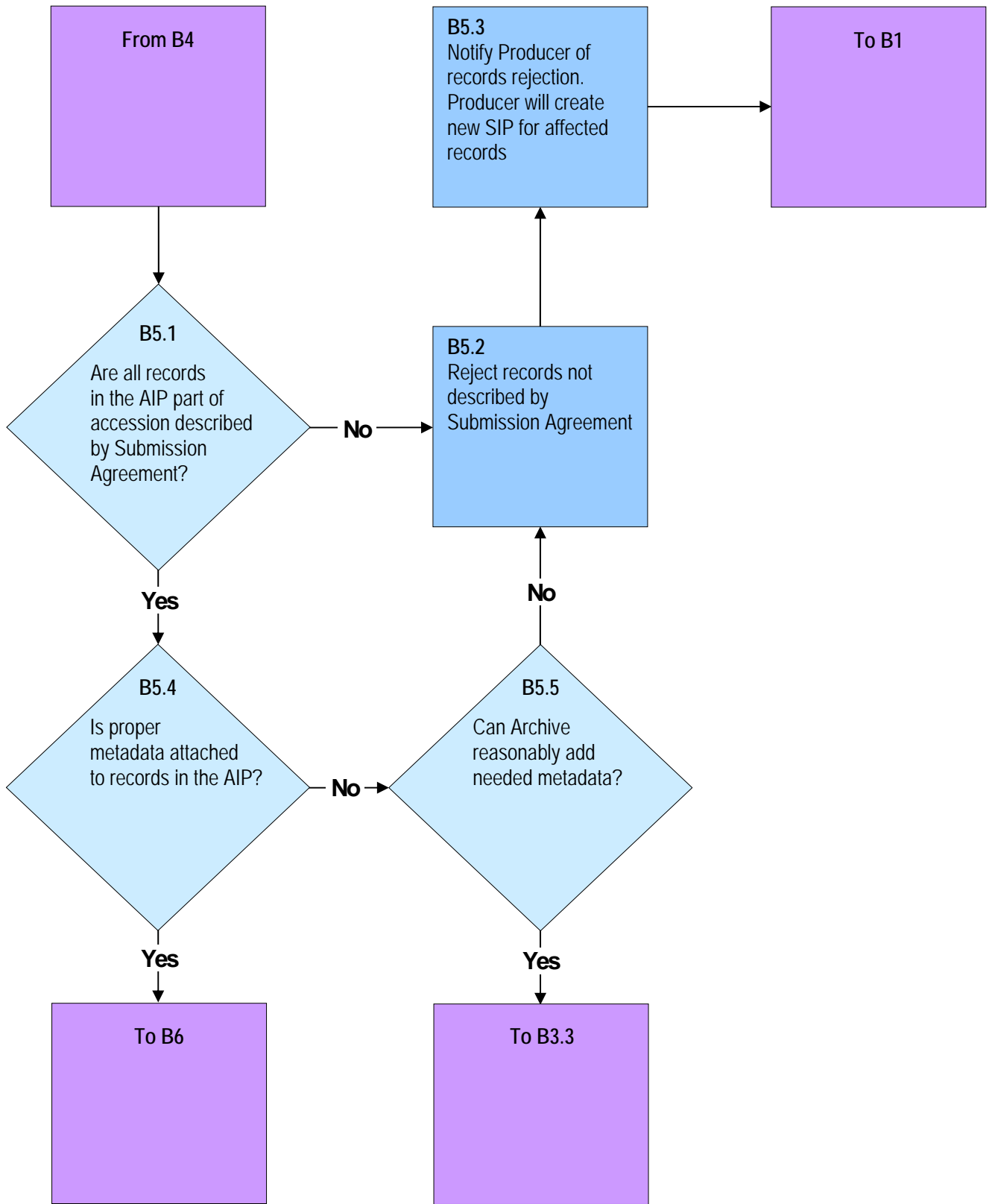
Uses AIP Configuration Rules

Produces/Modifies AIP



SECTION B: TRANSFER AND VALIDATION**Part B5: Assess AIPs****Overview**

During this Part, the Archive conducts a final appraisal of the records involved in the Ingest Project. It ensures that the records in the newly formed AIPs are the records described in the Submission Agreement and that they have the proper metadata associated with them. If the AIP does not contain the correct records, the Archive rejects the records and notifies the Producer to generate a new SIP for the affected records. If the records in the AIP do not have the proper metadata, the Archive determines if they can be added with a reasonable amount of effort. If this effort is too great, the Archive rejects the affected records. Usually, the Archive checks a sample of records involved in the Ingest Project. The appropriate rate of sampling depends on the circumstances of the Archive and individual Ingest Projects.

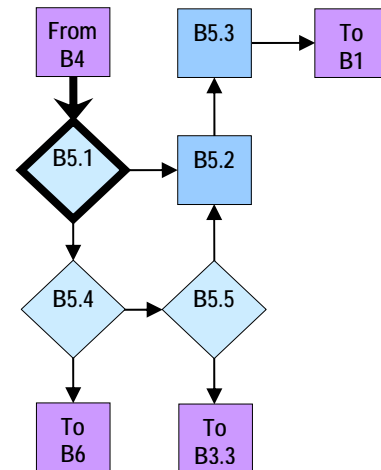


B5.1

Description The Archive manually checks the records in an AIP to determine if they are in fact the records described in the Submission Agreement.

Uses Survey Report

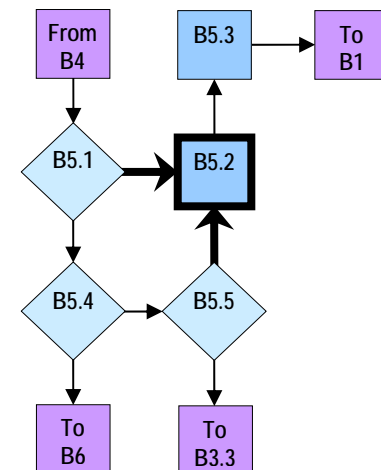
Produces/Modifies AIP Validity Statement

**B5.2**

Description If the records in an AIP do not match the records listed and described in a Submission Agreement, or lacks the necessary metadata the Archive cannot reasonably add on its own, the Archive rejects the affected records in the AIP.

Uses None

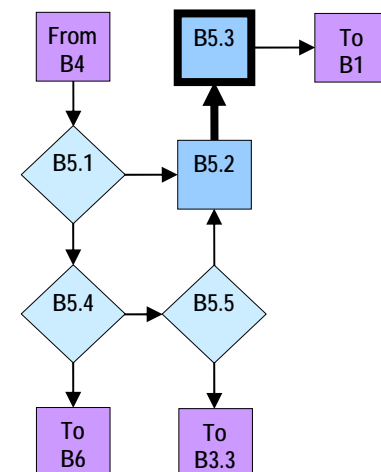
Produces/Modifies None

**B5.3**

Description The Archive notifies a Producer that it has rejected records in an AIP. The Producer corrects the error and repeats Step B1.1, creating a new SIP for the affected records.

Uses None

Produces/Modifies AIP Rejection Notice

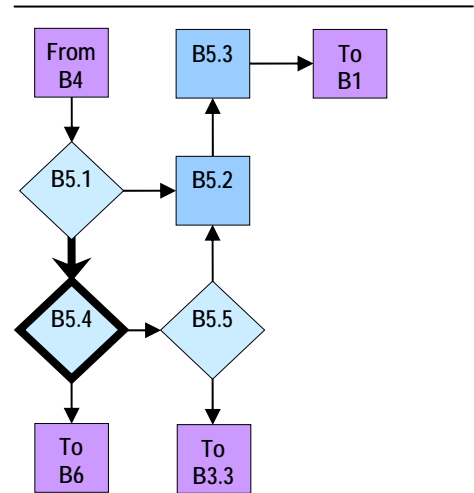


B5.4

Description The Archive manually checks the records in an AIP to determine if they have the properly associated metadata.

Uses Metadata Encoding Rules

Produces/Modifies AIP Validity Statement

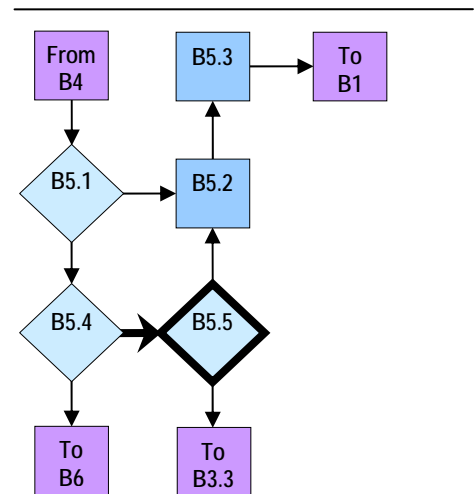


B5.5

Description If at least some of the records in an AIP do not have the proper metadata attached, the Archive determines if it can reasonably attach the proper metadata. If it can, it goes back to Step B3.3 to attach the proper metadata. If it cannot reasonably attach the proper metadata to the records in the AIP, the Archive moves to Step B5.2 and rejects the affected records.

Uses Metadata Encoding Rules

Produces/Modifies None

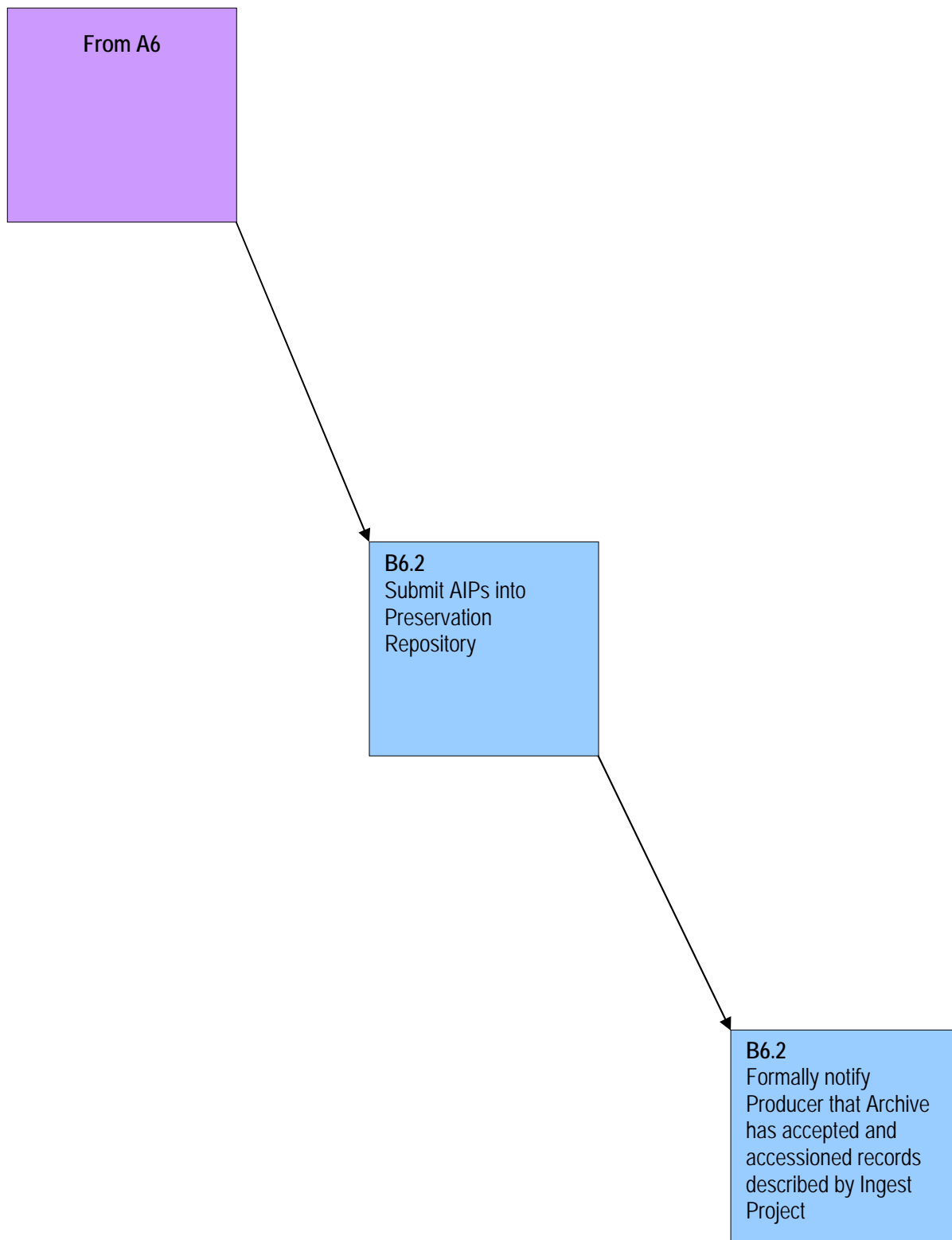


SECTION B: TRANSFER AND VALIDATION

Part B6: Formally Accession

Overview

During this part, the Archive deposits the AIPs it has formulated during an Ingest Project into its Preservation System. Then it formally notifies the Producer that it has accepted and accessioned the records the Producer transferred to the Archive in a SIP or SIPs. This is the moment of formal transfer from the Producer to the Archive.

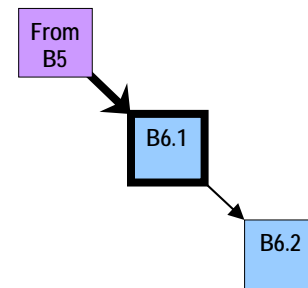


B6.1

Description The Archive deposits the AIP(s) to its Preservation System according to its Preservation System rules.

Uses None

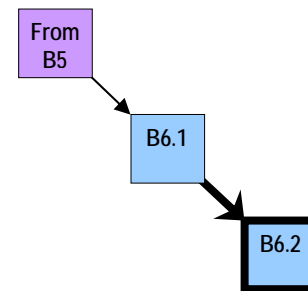
Produces/Modifies Preservation System-Managed AIP

**B6.2**

Description The Archive formally acknowledges that it has accepted and accessioned the records involved in the Ingest Project. This is the moment of formal transfer of the records in a SIP from the Producer to the Archive.

Uses None

Produces/Modifies Transfer Notice, Accession Log



SUBMISSION AGREEMENT

Overview

A Submission Agreement defines the nature and scope of the records involved in an Ingest Project and delineates the manner in which the Archive and the Producer execute the transfer, validation, and transformation of these records. In addition to guiding the work of the Producer and the Archive for transfer and transformation and serving as the benchmark for validation, it also provides both entities a document describing the terms of an Ingest Project that they can endorse and agree to. The Submission Agreement can cover a single Ingest Project or serial Projects.

The Submission Agreement documents the information needed to establish the terms of the scope, transfer, validation, and transformation of an Ingest Project. The Archive documents this as Components in the Submission Agreement. Most of these Components—and the decisions they represent—are tied to standing Resources. For example one of the elements of a Submission Agreement identifies the format types of the records in an Ingest Project. This Component in the Submission Agreement references a Formal Representation Information System on the format types the Archive employs as preservation formats in its Preservation System. Resources are usually policies, procedures, metadata records, or logs of action. While these Resources have an impact on nearly all of the Archive's Ingest Projects, they are not specific to any single Ingest Project.

Ideally, the Archive creates machine-readable and human-readable versions of its Submission Agreements. A human-readable version gives the Producer and the Archive a document both can agree to and endorse. A machine-readable version enables a degree of automated validation and transformation, usually coordinating calls to sets of other machine-readable code that dictate validation and transformation activities. This automation should help the Archive make the size and number of its Ingest Projects scalable. The degree of automation depends largely on the amount of detail an Archive's Resources has. For example, if the Archive's Format Representation Information System contains detailed, machine-readable, technical, and administrative metadata about each Format, it can use that metadata to automatically validate and transform the formats of records during Ingest. If the Archive's Format Representation Information System is a simple paper list and brief narrative description of formats, the Policy will not help the Archive automate the validation or normalization of records.

Some Ingest Projects—usually those with new types of records, formats, creators, or special circumstances—prompt the Archive to create a new version or add to one or more of its Resources. For example, if the Archive decides to preserve a record in a format that is not one of its existing preservation formats, it will have to add that format to its Format Representation Information System, producing technical and administrative metadata about the new format. Although this Resource development demands time and effort, it allows the Archive to automate a greater variety of Ingest Projects in the long run. As the Archive adds detail and breadth to its Resources, it will be able to automate a greater portion of its Ingest Process for a broader range of records.

The Survey Report, created in Parts A2 and A3, is a critical part of the Ingest process: it inventories the records that the Archive should accession in a particular Ingest Project. A Survey Report which a Submission Agreement references, can exist as a separate entity, or it can be embedded into the Submission Agreement.

See Appendix B for an example of a Submission Agreement.

COMPONENTS, RESOURCES, PRODUCTS, AND DOCUMENTATION

Overview

Below are descriptions of the Components, Resources, Products, and Documentation that the Archive uses to undertake Ingest Projects and create Submission Agreements. Each Component, Resource, Product, and Documentation includes a description, an explanation of its role in the Ingest process, and a listing of all the steps that use, produce, and modify it. They are listed alphabetically.

How these Resources and Components manifest themselves varies from Archive to Archive. Therefore, this section of the Ingest Guides describes each Resource and Component generally but does not prescribe their composition in detail. This section essentially highlights the need for the Archive to have these Resources and Components in some form, but does not extensively describe the manner of their existence. This section describes twelve Components, thirty Resources, eight Products, and twelve instances of Documentation.

Components

Expressions or selections of Resources the Archive documents in a Submission Agreement. The Components document decisions the Archive makes for the scope, transfer, validation, and transformation of records in an Ingest Project.

Resources

The Archive's standing policies, procedures, metadata records, or logs of action. Resources dictate the actions the Archive undertakes during its Ingest Projects. Resources are not specific to any single Ingest Project. The Submission Agreement references Resources through Components.

Products

Objects created or modified as a result of the work the Archive and the Producer undertake during an Ingest Project.

Documentation

Expressions of decisions the Archive makes during an Ingest Project that the Archive does not need in its Submission Agreement to undertake transfer, validation, and transformation but does need to document all of its Ingest decisions.

Access Controls Gap Analysis

Description An analysis of the security capabilities of the Preservation System in comparison to the access control needs of the records in an Ingest Project. If the security capabilities fall short of the records' access control needs, the Archive measures this gap and describes it in the Gap Analysis.

Ingest Project Role Documentation

Steps that Use A8.4, A8.5, A8.7

Steps that Produce/Modify A8.3

Access Controls Gap Analysis Feasibility Statement

Description A statement that declares whether the Preservation System is or is not capable of meeting the access control needs of the records in the Ingest Project.

Ingest Project Role Documentation

Steps that Use A8.7

Steps that Produce/Modify A8.4

Access Controls Policy

Description An articulation of an Archive's policy on access to records stored in the Preservation System. This includes security measures an Archive takes to prevent unauthorized access to records. This also usually includes such access restriction profiles as "open access," "restricted, administrative records," and "restricted, personal records" which declare who can gain access to the records, under what circumstances, and when.

Ingest Project Role Resource

Steps that Use A8.1, A8.2, A8.3, A8.7

Steps that Produce/Modify A8.7

Accession Log

Description A record of the accessions an Archive has made. At a minimum, an accession entry should record a basic description or identification of the accessioned records, the date of transfer, and the Producer that transferred the records to the Archive.

Ingest Project Role Resource

Steps that Use A1.3

Steps that Produce/Modify B6.2

Activity Log

Description A record of the work an Archive has done with or for a Producer. Entries in an Activity Log may include accessions, surveys, consultations, or any other type of activities the Archive may engage in. Because Activity Logs vary so greatly among Archives, an Archive may create an entry at any point during an Ingest Project depending on what activities it wants to document. An Archive should only keep an Activity Log to the extent that it wants to document its activities in this form.

Ingest Project Role Resource

Steps that Use A1.1, A1.3, A2.3

Steps that Produce/Modify Varies

Archival Information Package (AIP)

Description An Archival Information Package, commonly referred to as an “AIP,” is the form a record takes when it is managed in a preservation repository. *OAIS* defines an AIP as “An Information Package, consisting of the Content Information and the associated Preservation Description Information (PDI), which is preserved within an OAIS [Preservation System].”⁸

Ingest Project Role Product

Steps that Use None

Steps that Produce/Modify B4.1

Archival Information Package Configuration Rules

Description These rules articulate how an Archive needs to assemble its Archival Information Packages (AIPs) so it can successfully submit them to its preservation repository in order to manage and preserve them over time.

Ingest Project Role Resource

Steps that Use B4.1

Steps that Produce/Modify None

Archival Information Package Rejection Notice

Description The notification an Archive sends to a Producer that it has rejected that Producer’s Archival Information Package (AIP). The Notification should describe why and when the Archive terminated the AIP. An Archive also retains a copy of the Notice for its collection management system.

Ingest Project Role Documentation

Steps that Use None

Steps that Produce/Modify B5.3

Archival Information Package Validity Statement

Description The statement an Archive makes declaring the validity of an AIP.

Ingest Project Role Documentation

Steps that Use None

Steps that Produce/Modify B5.1, B5.4

Archive Naming/Identification Scheme

Description This scheme defines how an Archive names and/or identifies the records it holds in its Preservation System. An Archive may support multiple Naming/Identification schemes, although many Archives only support one such scheme. If an Archive decides to use a Producer Naming/Identification Scheme to name or identify records in its Preservation System, the Archive will adapt that Producer scheme as an Archive Naming/Identification Scheme in Step A6.3.

Ingest Project Role Resource

Steps that Use A6.2, A6.4

Steps that Produce/Modify A6.3

⁸ ISO 14721:2003, p. 1-7.

Archive Naming/Identification Scheme Decision

Description This documents the decision an Archive makes regarding the naming or identification scheme it applies to the records during an Ingest Project.

Ingest Project Role Component

Steps that Use None

Steps that Produce/Modify A6.4

Archives Directory

Description This directory lists archives, their contact information, and a brief summary of their collecting policies. An archive may use the directory to help direct a Producer to an appropriate Archive for its records.

Ingest Project Role Resource

Steps that Use A1.5

Steps that Produce/Modify None

Collection Policy

Description In addition to defining what type of records an Archive collects, a Collection Policy also identifies the Producers from whom an Archive will collect records. Usually, this covers Producers not addressed in a Records Authority Statement, although a Records Authority Statement and a Collection Policy may overlap.

Ingest Project Role Resource

Steps that Use A1.4, A3.3, A3.4

Steps that Produce/Modify None

Copyright Policy

Description The articulation of how an Archive manages the reproduction of records in light of their Copyright Status.

Ingest Project Role Resource

Steps that Use A7.2, A7.3, A7.4

Steps that Produce/Modify None

Copyright Status

Description An indication of the copyright status of the records and associated software in an Ingest Project, usually indicating the copyright holder and any applicable licensing agreement the Archive has with the copyright holder. This may include a description of the records' and associated software's moral rights.

Ingest Project Role Component

Steps that Use A7.2

Steps that Produce/Modify A7.1

Copyright Transfer/License

Description An agreement between an Archive and a copyright holder that is not the Producer in an Ingest Project. It either transfers copyright ownership of the records in an Ingest Project from the original copyright holder to the Archive, or it licenses to the Archive the right to reproduce the records under a set of conditions defined by the copyright holder.

Ingest Project Role Resource

Steps that Use None

Steps that Produce/Modify A7.8

Designated Community Description

Description A description of “an identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities.”⁹ Usually the description of a community’s ability to understand a set of information focuses on the community’s technical capability to functionally use the formats of a set of records.

Ingest Project Role Resource

Steps that Use A5.3

Steps that Produce/Modify None

Documentation of Receipt

Description A notification sent from an Archive to a Producer that it has received Submission Information Package(s) (SIP(s)) from the Producer. The notification is not a declaration that the Archive has formally accepted the records contained in the SIP(s) it has received.

Ingest Project Role Documentation

Steps that Use None

Steps that Produce/Modify B2.1

Draft Submission Agreement

Description A completed Submission Agreement that neither the Producer nor the Archive has endorsed.

Ingest Project Role Product

Steps that Use A11.7, A11.8, A11.9, A11.10

Steps that Produce/Modify A11.6, A11.8

Finalized and Endorsed Submission Agreement

Description A completed Submission Agreement that both the Producer and the Archive have endorsed.

Ingest Project Role Product

Steps that Use None

Steps that Produce/Modify A11.7, A11.9

Format Representation Information System

Description: A repository of information about all formats used by the Preserver along with documentation of verification, validation, and rendering tools for each format. The system should be able to associate digital components with corresponding file format specifications and the tools to work with them. These systems may be homegrown, created and maintained externally (PRONOM, Global Digital File Format Registry), or a hybrid of local and external representation information. Some of this information might be stored in paper documentation, but for the system to be useful it needs to be electronic and integrated into the Ingest process.

⁹ ISO 14721:2003, p. 1-10.

Ingest Project Role Resource

Steps the Use A5.1, A5.2, A5.3, B2.6

Steps the Produce/Modify None

Format Standards Policy

Description A Policy that declares the formats an Archive is able to process at Ingest, and the formats the Archive uses to preserve records in its Preservation System. A preservation format is a format that an Archive is capable of functionally preserving over the long term. Each time an Archive decides to employ a new preservation format, it needs to add that format to its Format Standards Policy.

Ingest Project Role Resource

Steps that Use A5.1

Steps that Produce/Modify A5.3, A5.4

Ingest Project Termination Notice

Description A notification an Archive sends to a Producer that it has ended the Ingest Project. The Notification should describe why and when the Archive ended the Ingest Project. An Archive also retains a copy of the Notice for its collection management system.

Ingest Project Role Documentation

Steps that Use None

Steps that Produce/Modify A1.5, A2.5, A3.5, A7.7, A8.6, A9.7, A10.5, A11.10

Institutional Identity Management System

Description An institution-wide identity management system that the Archive may use to help confirm the identity of a Producer. The system may manifest itself in a variety of ways, from a simple paper directory to a Lightweight Directory Access Protocol (LDAP)-based system.

Ingest Project Role Resource

Steps that Use A1.2, A1.6, A2.3, B2.1

Steps that Produce/Modify None

Metadata Encoding Rules

Description An articulation of the schemas for the metadata associated to the records in an Archive's Preservation System.

Ingest Project Role Resource

Steps that Use A11.1, B3.3, B5.4, B5.5

Steps that Produce/Modify None

Metadata Encoding Rules Decision

Description This articulates the metadata schema(s) an Archive decides to use as the encoding standard for the metadata it associates to the records during an Ingest Project.

Ingest Project Role Component

Steps that Use None

Steps that Produce/Modify A11.1

Preservation System Availability Statement

Description A statement of the institutional, financial, staffing, and technical capabilities of the Preservation System to manage and preserve the records during an Ingest Project.

Ingest Project Role Documentation

Steps that Use None

Steps that Produce/Modify A10.1, A10.2, A10.3

Preservation System Capabilities Report

Description A serial report that states the institutional, financial, staffing, and technical capabilities of the Preservation System.

Ingest Project Role Resource

Steps that Use A5.2, A8.2, A8.3, A10.1, A10.2

Steps that Produce/Modify None

Preservation System-Managed Archival Information Package

Description An AIP that is stored, maintained, and preserved in a Preservation System.

Ingest Project Role Product

Steps that Use None

Steps that Produce/Modify B6.1

Producer Entry

Description A reference to a specific Producer Record about a particular Producer.

Ingest Project Role Component

Steps that Use B2.4

Steps that Produce/Modify A1.3, A1.6

Producer Naming/Identification Scheme

Description A scheme that defines how a Producer names and/or identifies its records.

Ingest Project Role Documentation

Steps that Use A6.1, A6.2, A6.3

Steps that Produce/Modify None

Producer Record

Description A record that authoritatively identifies and describes a Producer. It also describes the relationship of a Producer to an Archive and its relationship with other Producers. Records of producers may describe individuals or departments, offices, or units. An Archive should have a record of every Producer that has transferred records to the Archive. This way, the Archive can associate all of its records in its holdings with a Producer. A Producer Record is usually encoded as machine-readable metadata. For example, a Producer Record may be written to ISAAR(CPF) (International Standard Archival Authority Record for Corporate Bodies, Persons, and Families, 2nd ed.) and encoded in EAC (Encoded Archival Context). The more details a Producer Record has, the more an Archive can utilize that record to automate its management and description of the records in its holdings. An Archive might populate a Producer Record with data from its Institutional Identity Management System.

Ingest Project Role Resource

Steps that Use A1.3, A2.2, A2.3, B2.1

Steps that Produce/Modify A1.6

Record Security Profile

Description A description of a class of security characteristics that an Archive assigns to the records in its Preservation System. A Profile articulates the access control needs of a record. Profiles are usually records-specific articulations of an Archive's Access Controls Policy.

Ingest Project Role Resource

Steps that Use A8.1, A8.2, A8.3, B3.4

Steps that Produce/Modify A8.1, A8.5

Record Security Profile Decision

Description A reference to a particular Records Security Profile.

Ingest Project Role Component

Steps that Use None

Steps that Produce/Modify A8.1

Record Type List

Description A listing of the Record Types of the records in an Ingest Project.

Ingest Project Role Component

Steps that Use None

Steps that Produce/Modify A4.1

Record Type Record

Description Documentation that describes the type of records that an Archive may accession into its Preservation System. A Record Type Record may describe various properties of a record type, including its general composition, function, confidentiality status, its Producers, and their management by Producers and Archives. The more details a Record Type Record has, the more an Archive can utilize it to automate its management and description of the records in its holdings.

Ingest Project Role Resource

Steps that Use A4.1, B2.5, B2.6

Steps that Produce/Modify A4.2

Recordkeeping System Evaluation Tool

Description A set of requirement delineating the attributes a recordkeeping system needs to be considered a trustworthy system. These requirements usually include requirements to facilitate the trustworthy transfer of records to a Preservation System.

Ingest Project Role Resource

Steps that Use A3.2, A9.2

Steps that Produce/Modify None

Recordkeeping System Internal Rules

Description A recordkeeping system's set of rules delineating how it must manage records.

Ingest Project Role Resource

Steps that Use A9.3

Steps that Produce/Modify None

Recordkeeping System Report

Description Records that identify and describe recordkeeping systems. These records should document an Archive's determination of a system's trustworthiness—its characteristics that allows a person to presume the authenticity of the records it manages—and its ability to support the scaleable, feasible, and trustworthy transfer of records to the Archive. An Archive should update the Report of a recordkeeping system often enough to accurately reflect its current configuration.

Ingest Project Role Resource

Steps that Use 3.2, A9.1

Steps that Produce/Modify 3.2, A9.2

Records Authority Statement

Description A statement that gives an Archive the authority to serve as the Archive for Producers. If applicable, it provides evidence that an Archive has the right to serve as the Archive for the records in the Ingest Project.

Ingest Project Role Resource

Steps that Use A1.4

Steps that Produce/Modify None

Records Retention Policy

Description An articulation of the disposition and period of retention of an institution's records. A Records Retention Policy may declare some records to have a disposition of permanent retention in the institution's Archive.

Ingest Project Role Resource

Steps that Use A3.3, A3.4

Steps that Produce/Modify None

Records Transformed

Description Records that an Archive has normalized into one its Preservation Formats.

Ingest Project Role Product

Steps that Use None

Steps that Produce/Modify B3.2

Records with Attached Metadata

Description Records to which an Archive has associated metadata during Part B3.

Ingest Project Role Product

Steps that Use None

Steps that Produce/Modify B3.3

Records with Security Profile

Description Records to which an Archive has associated security profile(s) during Part B3.

Ingest Project Role Product

Steps that Use None

Steps that Produce/Modify B3.4

Representation Information

Description The details needed to make the content of a record (defined by *OAIS* as a Content Data Object) understandable to a Designated Community.¹⁰ Representation Information allows for the full interpretation of the content into meaningful concepts. Generally, Representation Information is the technical information needed to fill in the gap between the configuration of a record's format and the knowledge base of a Designated Community. Each time an Archive decides to employ a new preservation format, it will need to generate Representation Information for that preservation format in Step A4.3. An Archive may also have to produce new Representation Information in response to a change in a Designated Community's knowledge base.

Ingest Project Role Resource

Steps that Use None

Steps that Produce/Modify A5.3

Submission Information Package (SIP)

Description A Submission Information Package, commonly referred to as a "SIP," is created by a Producer to prepare records for transfer to an Archive. *OAIS* defines a SIP as "an Information Package that is delivered by the Producer to the OAIS [Preservation System] for use in the construction of one or more AIPs."¹¹

Ingest Project Role Product

Steps that Use None

Steps that Produce/Modify B1.1

Submission Information Package Creation Procedures

Description A set of procedures that describes how a Producer should create a SIP.

Ingest Project Role Resource

Steps that Use A11.5, B1.1, B2.2, B2.3

Steps that Produce/Modify A9.4

Submission Information Package Creation Procedures Decision

Description An articulation of an Archive's decision to use a set SIP Creation Procedures in an Ingest Project.

Ingest Project Role Component

Steps that Use B1.1

Steps that Produce/Modify A11.5

Submission Information Package Rejection Notification

Description The notification an Archive sends to a Producer that it has rejected the Producer's SIP. The Notification should describe why and when the Archive rejected the SIP.

Ingest Project Role Documentation

Steps that Use None

Steps that Produce/Modify B2.8

¹⁰ ISO 14721:2003, p. 1-13.

¹¹ ISO 14721:2003, p. 1-13.

Submission Information Package Validity Statement

Description A statement that an Archive makes that it has validated a SIP.

Ingest Project Role Documentation

Steps that Use None

Steps that Produce/Modify B2.2, B2.3, B2.4, B2.5, B2.6

Survey Instrument

Description A tool that an Archive uses to gather information about the records it surveys. This tool should be designed to help the Archive execute its Survey Procedures and help ensure that the Archive gathers the information it needs to capture during a survey. The Survey Instrument usually manifests itself as a form but can be any type of tool or set of tools that the Archive uses to gather information about records.

Ingest Project Role Resource

Steps that Use A3.1

Steps that Produce/Modify None

Survey Procedures

Description A plan or set of methods established to guide an Archive in undertaking a Records Survey. Usually these procedures consist of the Archive interviewing a Producer, the Archive querying the Producer through some sort of questionnaire, the Archive querying the electronic records themselves, or some combination of these methods.

Ingest Project Role Resource

Steps that Use A3.1

Steps that Produce/Modify None

Survey Report

Description A Report that identifies the records an Archive should accession during the Ingest Project. Survey Reports can vary greatly in detail, from a general description of the records the Archive should accession to an item-level inventory of those records. An Archive may create an early working draft of the Report in Step A2.1, after it and the Producer agree on the scope of the records that will be surveyed. In Steps A3.1 and A3.2, the Archive describes in the Survey Report the records it surveyed to the level of detail it requires. In Steps A3.3 and A3.4, the Archive documents in the Report its decisions on which, if any, records in the survey it should accession and what essential elements of these records it needs to preserve. To guide the Archive's appraisal decisions in Steps A3.3 and A3.4 and to be useful in Parts A3 through A10, the Survey Report needs to identify the records' Producer, Record Types, format type, file size, confidentiality requirements, copyright status, and any Producer-created identifiers.

Ingest Project Role Documentation

Steps that Use A2.1, A3.1, 3.3, A3.4, A4.1, A4.2, A5.1, A5.2, A5.5, A6.1, A6.4, A7.1, A7.6, A8.1, A8.5, A9.6, A10.1, A10.2, A10.3, A10.4, B5.1

Steps that Produce/Modify A3.1, A3.2, A3.3, A3.4, A7.5, A8.5, A9.5, A10.3

Transfer Notice

Description A notice an Archive creates and sends to a Producer declaring that it has received and accessioned the records of an Ingest Project.

Ingest Project Role Documentation

Steps that Use None

Steps that Produce/Modify B6.1

Transfer Procedures

Description A set of procedures that articulates how a Producer transfer records to an Archive.

Ingest Project Role Resource

Steps that Use A11.2, A11.4

Steps that Produce/Modify None

Transfer Procedures Decision

Description An articulation of an Archive's decision to use a set Transfer Procedures in an Ingest Project.

Ingest Project Role Component

Steps that Use None

Steps that Produce/Modify A11.2

Transfer Schedule

Description A schedule that articulates when a Producer transfers a SIP or set of SIPs to an Archive during an Ingest Project.

Ingest Project Role Component

Steps that Use B1.2

Steps that Produce/Modify A11.4

Transformation Plan

Description An articulation of how an Archive plans to transform the records involved in an Ingest Project, detailing the process of accepting format types from the Producer and transforming them into appropriate Preservation Formats.

Ingest Project Role Component

Steps that Use B2.5, B2.6, B3.1, B3.2

Steps that Produce/Modify A5.1, A5.4, A5.5

Transformation Policy

Description A policy that articulates an Archive's approach to making appraisal decisions concerning its transformation of records. It guides an Archive's decision to transform a record from one format to another. This Policy can exist as very general policy stating broad appraisal principles, or it can exist as a very detailed policy mapping the transformation of a specific format to a preservation format for a specific record type. This Policy can exist as information embedded in the Records Retention Policy, Format Standards Policy, Format Representation Information System, and/or Collection Policy. If the Archive is dealing with a new combination of record type and format type in Step A4.5, it probably has to add to or modify its Transformation Policy.

Ingest Project Role Resource

Steps that Use A5.5

Steps that Produce/Modify A5.5

Validation Procedures

Description A set of procedures that articulates how an Archive validates records.

Ingest Project Role Resource

Steps that Use A11.3

Steps that Produce/Modify None

Validation Procedures Decisions

Description The articulation of an Archive's decision to use a set Validation Procedures during an Ingest Project.

Ingest Project Role Component

Steps that Use None

Steps that Produce/Modify A11.3

Virus Definition Files

Description These files define viruses.

Ingest Project Role Component

Steps that Use B2.3

Steps that Produce/Modify None

APPENDIX A: USING THE INGEST GUIDE

Overview

The Ingest Guide is a prescriptive guide. It describes the steps Archives need to undertake to have a trustworthy ingest process; the Guide is not a detailed procedure manual: it does not explain how an Archive should precisely execute these steps or construct its Resources. How the Guide is used will vary greatly from Archive to Archive depending on circumstances and needs.

At first reading, the Ingest Guide may appear to prescribe a daunting process, but much of what the Guide asks Archives to undertake they already do, often intuitively or informally. Archives already appraise records; determine their appropriate access restrictions; make transformation decisions (like photocopy newspaper clippings) and apply metadata (like finding aids) to records. The Ingest Guide calls for Archives to document their decisions carefully and to base them on well-documented procedures, policies, and standards. Because electronic records are less forgiving about preservation than paper records, this documentation is imperative. In addition, traditional archival practices have often produced less than ideal results. What archivist has not contended with poorly documented terms of use, transfer, or preservation decisions?

Although following the Ingest Guide entails more work than most traditional archival accessioning methods do, this more carefully documents the accession process. It also regularizes and streamlines many decision-making steps and offers the potential to automate a considerable portion of accessioning, preservation, and description. The Guide is geared towards enabling an Archive to ingest records in a semi-automated and scalable manner. The more an archive articulates its Resources as machine-readable objects, the more it will be able to automate its ingest process. Obviously, expressing Resources as machine-readable objects can take a considerable investment of effort. Each Archive will have to determine the degree of automation that is appropriate for its operations. Archives will get the biggest payoff from automating their Ingest process when they have serial Ingests Projects from the same Producer sending the same type of records. An Archive can rapidly repeat the Ingest decisions they previously made and the more it has automated its SIP creation, transfer, validation, and transformation steps, the more quickly it can turn its SIPs into AIPs and complete its Ingest Projects. This work slows down when an Archive undertakes an Ingest Project with a new producer or record type or format since it has to add to its stable of resources to accommodate this new type of accession.

However, reaching a highly automated, trustworthy Ingest process requires individual Archives or the records and digital preservation communities to undertake a number of tasks that this Guide points out but does not fully address. Three of the most significant of these tasks are:

- 1 Develop Resources

In order for an Archive to have an automated Ingest process it needs to have fully developed rules for creating the Resources and a schema for articulating them in a machine-readable manner. Currently, most of the Resources do not have the needed rules or schemas.

2 Create SIPs (Part B1)

Creating the Resource “Submission Information Package Creation Procedures” alone represents a substantial amount of work for an Archive. Not only does a Producer have to configure records into a SIP properly, it has to extract those records from a recordkeeping environment in a trustworthy manner which is not a trivial task.

3 Appraise Records (Part A3)

An Archive has to use fully developed methodologies for surveying records and for determining a record’s authenticity, disposition, and essential elements. Implementing these methodologies is a major undertaking for any Archive.

APPENDIX B: EXAMPLE OF A SUBMISSION AGREEMENT

Overview

Below are the machine-readable and the human-readable versions of a single Submission Agreement. Each version is a different expression of the same information. The data in machine-readable Submission Agreement are Components that reference Resources. The Components are expressed as underlined text in the human-readable version.

The Archive uses the human-readable version to document its own and the Producer's endorsements of the Submission Agreement. It uses the machine-readable version to automate its validation and transformation work.

In this example the Archive, the Digital Collections and Archives at Tufts University, accessions the website of the Task Force on the Undergraduate Experience. It has previously collected paper records from the Task Force.

This is a detailed illustration of a Submission Agreement; it is not a guide for constructing them. This is one example of the many ways an Archive can organize its Submission Agreements. A standard syntax for expressing Submission Agreements does not yet exist. In order for Archives to use these Agreements successfully, that syntax will need to be created.

2.1 Ingest Guide

```
<submission-agreement id="SA00023">

  <record-survey id="RS00023" />

  <archive>
    <identifier>US::TUFTSU::Central::0001</identifier>
    <name>Digital Collections and Archives</name>
  </archive>

  <producer>
    <identifier>US::TUFTSU::Taskforce::0001</identifier>
    <name>Task Force on the Undergraduate Experience</name>
  </producer>

  <circumstance>
    This Ingest Project occurs after a previous Ingest Project with the Task Force
    concerning its paper records.
  </circumstance>

  <description>
    The records in this Ingest Project are records on the website
    (http://ugtaskforce.tufts.edu) of the Task Force on the Undergraduate Experience
    (CID US::TUFTSU::Taskforce::0001). These include the President's charge, reports
    generated by the Task Force, and the additional records describing the Task Force's
    activities and findings.
  </description>

  <subscribed-users>
    <user>user1</user>
    <user>user2</user>
  </subscribed-users>

  <workflow>University Records</workflow>

  <sipcreation status="009" elements="ALL" />

  <transfer status="008" elements="ALL" />

  <descriptionstandard status="001" elements="ALL" />

  <copyright status="001" elements="ALL">
    Copyright held by Tufts University.
  </copyright>

  <confidentiality status="001" elements="ALL">
    Open access to the general public.
  </confidentiality>

  <metadata elements="ALL">
    <tag tagNS ="http://purl.org/dc/elements/1.1/"
      tagName="publisher"
      value = "Task Force on the Undergraduate Experience"/>

    <tag tagNS ="http://purl.org/dc/elements/1.1/"
      tagName="rights"
      value = "Copyright Tufts University 2001"/>
  </metadata>

  <relationships>
    <sa-relationship/>

    <relationship urn="tufts:central:dca:UA088:00001"
      relNS="info:fedora/fedora-system:def/relations-external"
```



```

        relName="isMemberOf"/>

    <relationship urn="tufts:central:dca:PROD:00001"
        relNS="http://dca.tufts.edu/ns/relations/"
        relName="producedBy"/>
</relationships>

<element id="SA00023:001">
    <description>
        University president's charge to the Task Force.
    </description>

    <urnpool>
        <list>
            <urn>tufts:central:dca:nhprc-erec:UGT:00001</urn>
        </list>
    </urnpool>

    <date range begin="1999" end="1999" />

    <object-profiles>
        <profile name="0011"/>
    </object-profiles>

    <relationships>
        <relationship urn="tufts:central:dca:RTD:00019"
            relNS="http://dca.tufts.edu/ns/relations/"
            relName="hasRecordType"/>
    </relationships>
</element>

<element id="SA00023:002">
    <description>
        Various interim, status, and final reports created by the Task Force.
    </description>

    <urnpool>
        <range>
            <begin>tufts:central:dca:nhprc-erec:UGT:RP001</begin>
            <end>tufts:central:dca:nhprc-erec:UGT:RP999</end>
        </range>
    </urnpool>

    <date range begin="2000" end="2003" />

    <object-profiles>
        <profile name="0011"/>
    </object-profiles>

    <relationships>
        <relationship urn="tufts:central:dca:RTD:00011"
            relNS="http://dca.tufts.edu/ns/relations/"
            relName="hasRecordType"/>
    </relationships>
</element>

<element id="SA00023:003">
    <description>
        List of outreach activities undertaken by the Task Force
    </description>

    <urnpool>
        <list>

```

2.1 Ingest Guide

```
<urn>tufts:central:dca:nhprc-erec:UGT:00002</urn>
</list>
</urnpool>

<date range begin="2003" end="2003" />

<object-profiles>
  <profile name="0031"/>
</object-profiles>

<relationships>
  <relationship urn="tufts:central:dca:RTD:00021"
               relNS="http://dca.tufts.edu/ns/relations/"
               relName="hasRecordType"/>
</relationships>
</element>

<element id="SA00024:004">
  <description>
    List of links to online news stories concerning the Task Force
  </description>

  <urnpool>
    <list>
      <urn>tufts:central:dca:nhprc-erec:UGT:00003</urn>
    </list>
  </urnpool>

  <date range begin="2003" end="2003" />

  <object-profiles>
    <profile name="0043"/>
  </object-profiles>

  <relationships>
    <relationship urn="tufts:central:dca:RTD:00035"
                 relNS="http://dca.tufts.edu/ns/relations/"
                 relName="hasRecordType"/>
  </relationships>
</element>

<element id="SA00024:005">
  <description>
    Digitized print news stories concerning the Task Force
  </description>

  <urnpool>
    <range>
      <begin>tufts:central:dca:nhprc-erec:UGT:00100</begin>
      <end>tufts:central:dca:nhprc-erec:UGT:00199</end>
    </range>
  </urnpool>

  <date range begin="2003" end="2003" />

  <object-profiles>
    <profile name="0011"/>
  </object-profiles>

  <relationships>
    <relationship urn="tufts:central:dca:RTD:00035"
                 relNS="http://dca.tufts.edu/ns/relations/"
                 relName="hasRecordType"/>
  </relationships>
</element>
```

```
</element>

<element id="SA00024:006">
  <description>
    List of Task Force members.
  </description>

  <urnpool>
    <list>
      <urn>tufts:central:dca:nhprc-erec:UGT:00004</urn>
    </list>
  </urnpool>

  <date range begin="2003" end="2003" />

  <object-profiles>
    <profile name="0031"/>
  </object-profiles>

  <relationships>
    <relationship urn="tufts:central:dca:RTD:00015"
      relNS="http://dca.tufts.edu/ns/relations/"
      relName="hasRecordType"/>
  </relationships>
</element>

<element id="SA00024:007">
  <description>
    List of links to studies concerning undergraduates at other institutions
    the Task Force used as benchmarks
  </description>

  <urnpool>
    <list>
      <urn>tufts:central:dca:nhprc-erec:UGT:00005</urn>
    </list>
  </urnpool>

  <date range begin="2003" end="2003" />

  <object-profiles>
    <profile name="0043"/>
  </object-profiles>

  <relationships>
    <relationship urn="tufts:central:dca:RTD:00015"
      relNS="http://dca.tufts.edu/ns/relations/"
      relName="hasRecordType"/>
  </relationships>
</element>

<element id="SA00024:008">
  <description>
    Content of the Task Force website as a whole.
  </description>

  <urnpool>
    <list>
      <urn>tufts:central:dca:nhprc-erec:UGT:00006</urn>
    </list>
  </urnpool>
```

2.1 Ingest Guide

```
<date range begin="ca. 2000" end="2003" />

<object-profiles>
  <profile name="0064"/>
</object-profiles>

<relationships>
  <relationship urn="tufts:central:dca:RTD:00017"
    relNS="http://dca.tufts.edu/ns/relations/"
    relName="hasRecordType"/>
</relationships>
</element>

</submission-agreement>
```

SUBMISSION AGREEMENT

Submission Agreement ID SA00023

This Submission Agreement defines the terms of the transfer of records described in this Agreement and in the Records Survey (RS00023) (hereafter known as “The Records”) from the Task Force on the Undergraduate Experience (US::TUFTSU::Taskforce::0001) to the Digital Collections and Archives (DCA).

By agreeing to this Submission Agreement the Task Force on the Undergraduate Experience (US::TUFTSU::Taskforce::0001) declares that it has the proper authority to transfer the records to the DCA.

Special circumstances of Ingest Project

This Ingest Project occurs after a previous Ingest Project with the Task Force concerning its paper records.

General Description of the Records

The records in this Ingest Project are records on the website (<http://ugtaskforce.tufts.edu>) of the Task Force on the Undergraduate Experience (CID US::TUFTSU::Taskforce::0001). These records are the President’s charge, reports generated by the Task Force, and the additional records describing the Task Force’s activities and findings.

Detailed Description the Records

All of the records

Will be prepared for transfer according to the DCA SIP Creation rule for web-based records. (SIP Creation Rule 009)

All of the records

Will follow the DCA standard transfer procedures for university records on web server. (Transfer Rule 008)

All of the records

Will be part of Task Force on Undergraduate Experience Collection (UA088)

All of the records

Will be described according to the DCA standard descriptive rule for university records. (Metadata Descriptive Rule 01)

All of the records

Copyright held by Tufts University. (Copyright Status 001)

All of the records

Open access to the general public. (Records Security Profile 01)

President's Charge (SA00023:001)

These record(s) are the University President's charge to the Task Force. They are Charges (Record Type 00019) created from 1999 through 1999 in the form of PDF documents that the DCA will keep and preserve as PDF documents (Object-Profile 0011).

Various Reports (SA00023:002)

These record(s) are the Various interim, status, and final reports created by the Task Force. They are Reports (Record Type 00011) created from 2000 through 2003 in the form of PDF documents that the DCA will keep as PDF documents (Object-Profile 0011).

Outreach Activities List (SA00023:003)

These record(s) are the List of outreach activities undertaken by the Task Force. They are Event Records (Record Type 00021) created from 2003 through 2003 in the form of an HTML file that the DCA will normalize into plain text (Object-Profile 0031).

Links to News Stories (SA00023:004)

These record(s) are the List of links to online news stories concerning the Task Force. They are News Clippings (Record Type 00035) created from 2003 through 2003 in the form of an HTML file that the DCA will normalize into an XBEL file (Object-Profile 0043).

News Stories (SA00023:005)

These record(s) are the Digitized print news stories concerning the Task Force. They are News Clippings (Record Type 00035) created from 2003 through 2003 in the form of PDF documents that the DCA will keep as PDF documents (Object-Profile 0011).

Membership List (SA00023:006)

These record(s) are a List of Task Force members. They are Subject Files (Record Type 00015) created from 2003 through 2003 in the form of an HTML file that the DCA will normalize into plain text (Object-Profile 0031).

Benchmarking Studies (SA00023:007)

These record(s) are the List of links to studies concerning undergraduates at other institutions the Task Force used as benchmarks. They are Subject Files (Record Type 00015) created from 2003 through 2003 in the form of an HTML file that the DCA will normalize into an XBEL file (Object-Profile 0043).

Website (SA00023:008)

These record(s) are the content of the Task Force website as a whole. They are Publications (Record Type 00017) created in ca. 2000 through 2003 in the form of HTML files, PDF files, and JPEG files that the DCA will normalize together into an ZIP Tidy file (Object-Profile 0064).

Endorsement of Submission Agreement

The Digital Collections and Archives agrees to the conditions of this Submission Agreement, which commits the DCA to accession the records into its holdings according to the terms of the Submission Agreement.

Eliot Wilczek
University Records Manager
Authorized Representative of the Digital Collections and Archives

The Task Force on the Undergraduate Experience, agrees to the conditions of this Submission Agreement, which commits the Task Force on the Undergraduate Experience to transfer the records to the DCA according to the terms of the Submission Agreement.

Armand Greene
Director
Authorized Representative of Task Force on the Undergraduate Experience

APPENDIX C: PRODUCER-ARCHIVE INTERFACE METHODOLOGY ABSTRACT STANDARD CROSSWALK

Overview

The Ingest Guide builds on the work of the Consultative Committee for Space Data Systems, particularly its 2004 *Producer-Archive Interface Methodology Abstract Standard*,¹² which consists of four phases: Preliminary, Formal, Transfer, and Validation. In its Preliminary and Formal phases, the *Producer-Archive Interface* proposes a detailed description of the steps needed to produce a submission agreement between a Producer and an Archive. These steps form a foundation for the development of Section A of the Ingest Guide, Negotiate Submission Agreement. While the *PAI* separates the submission agreement process into separate Preliminary and Formal phases, the Ingest Guide does not make such a distinction. The *PAI*'s Transfer, and Validation phases are much less detailed than its Preliminary and Formal phases and only guided in a general way the development of Section B of the Ingest Guide, Transfer and Validation.

The crosswalks from the Ingest Guide to the *PAI* should help users understand how these two documents relate to each other.

¹² Consultative Committee for Space Data Systems, *Producer-Archive Interface Methodology Abstract Standard*, CCSDS 651.0-B-1, Blue Book, May 2004. <<http://www.ccsds.org/CCSDS/documents/651x0b1.pdf>>.

Crosswalks

Ingest Guide To Producer-Archive Interface Methodology Abstract Standard Crosswalk	
Ingest Guide	Producer-Archive Interface Methodology Abstract Standard
A1.1	P1
A1.2	P1
A1.3	None
A1.4	None
A1.5	None
A1.6	None
A2.1	P2, P3
A2.2	None
A2.3	None
A2.4	None
A2.5	None
A2.6	None
A3.1	P2, P9, P12, P16, P23, P24, P29, P30
A3.2	None
A3.3	P3
A3.4	None
A3.5	None
A4.1	None
A4.2	None
A5.1	P10
A5.2	P9, P10
A5.3	P4, P5
A5.4	P10, P12, P14
A5.5	P10, P12, P14
A5.6	None
A6.1	P16
A6.2	P17
A6.3	P17
A6.4	P17
A7.1	P29
A7.2	P29
A7.3	P29
A7.4	P29
A7.5	P29
A7.6	None
A7.7	F13
A8.1	P23, P24, P25, P30
A8.2	P23, P24, P25, P30
A8.3	P23, P24, P25, P30
A8.4	P42
A8.5	P42

2.1 Ingest Guide

A8.6	None
A8.7	None
A8.8	TBD
A8.9	P6
A10.1	P8, P15, P18, P22, P27, P33, P36, P40, P42, P43
A10.2	P43
A10.3	None
A10.4	None
A10.5	None
A11.1	F4
A11.2	F14, F15, F16, F17, F18, F19
A11.3	F20, F21, F22, F23, F24, F25
A11.4	F16, F26
A11.5	F15
A11.6	F36
A11.7	None
A11.8	F36
A11.9	F36
A11.10	F36
B1.1	T2
B1.2	T2
B2.1	T2
B2.2	V2
B2.3	V2
B2.4	V2
B2.5	V2
B2.6	V2
B2.7	V3
B2.8	V3
B3.1	None
B3.2	None
B3.3	None
B3.4	None
B4.1	None
B5.1	V2
B5.2	V3
B5.3	T3
B5.4	V2
B5.5	None
B6.1	T2
B6.2	T2