



InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

Title: General Study 06 Final Report:
A Bayesian Belief Network: Supporting the
Assessment of the Degree of Belief that a
Recordkeeping System Maintains Authentic
Digital Records

Status: Final (public)

Version: 2.0

Submission Date: July 2007 (Revised February 2008)

Release Date: February 2008

Author: The InterPARES 2 Project

Writer(s): William Underwood and Sheila Isbell
Computer Science and Information Technology Division,
Information and Telecommunications Laboratory,
Georgia Tech Research Institute

Project Unit: Focus 2

URL: http://www.interpares.org/display_file.cfm?doc=ip2_gs06_final_report.pdf

TABLE OF CONTENTS

1. Introduction	2
1.1 Background.....	2
1.2 Purpose.....	2
1.3 Scope.....	2
2. The Benchmark Requirements and Method of Assessment	3
3. Subjective Probability and Bayesian Inference	7
3.1 Bayesian Belief Networks.....	7
3.2 Genie and Smile.....	7
4. A BBN Supporting the Assessment of the Degree of Belief that a Recordkeeping System is Trusted and Contains Authentic Digital Records	9
5. Using the BBN to Support Assessment	19
5.1 The Mars Global Explorer Data Records in the Planetary Data System.....	19
5.2 Assessment of the Trustworthiness of the Planetary Data System.....	19
6. Conclusion and Areas of Future Research	27
References	28

Table of Figures

Figure 1. The Benchmark Requirements for a Presumption of Authenticity of a Creator’s Records.....	4
Figure 2. The Genie Interface	8
Figure 3. Definition of Conditional Dependencies Among Nodes in a BBN	8
Figure 4. Graph Showing the Dependencies Between The Hypothesis and the Requirements	10
Figure 5. Dependence of Requirement A1 on Requirements A1a and A1b	10
Figure 6. Dependence of Requirement A.1.A on Other Requirements.....	11
Figure 7. Dependence of Requirement A.1.B on Other Requirements.....	12
Figure 8. Dependence of Requirement A.2 on Other Requirements	13
Figure 9. Dependence of Requirement A.3 on Other Requirements	14
Figure 10. Dependence of Requirement A.4 on Other Requirements	14
Figure 11. Dependence of Requirement A.5 on Other Requirements	15
Figure 12. Truth Table Showing Dependence of Requirement A.6 on its Component Statements.....	15
Figure 13. Conditional Requirement A.6.....	16
Figure 14. Conditional Requirement A.7.....	16
Figure 15. Conditional Requirement A.8.....	18
Figure 16. The BBN Model for the Benchmark Requirements	18
Figure 17. Correspondence of Verbal Expressions of Degrees of Belief to Numerical Expressions	20
Figure 18. A PDS Label for a Mars Orbital Camera Image of the Martian Surface.....	21
Figure 19. Using Genie to Set Evidence of Nodes in the BBN.	26

ABSTRACT

One of the research results of InterPARES 1 was the Authenticity Task Force's (ATF's) formulation of the *Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records*. Quoting the ATF report: "A presumption of authenticity will be based upon the number of requirements that have been met and the degree to which each has been met. The requirements are, therefore, cumulative: the higher the number of satisfied requirements, and the greater the degree to which an individual requirement has been satisfied, the stronger the presumption of authenticity." The purpose of this General Study is to investigate just how these requirements and method of assessment can actually be applied to assess the authenticity of a creator's digital records.

The approach is based on an interpretation of probability suggested by Bayesian theory, which holds that the concept of probability can be defined as the degree to which a person believes a proposition. Bayesian theory also suggests that Bayes' theorem can be used as a rule to infer or update the degree of belief in light of new information.

First, the Benchmark Requirements are represented as a Bayesian Belief Network (BBN). A BBN is a graphical notation with an associated set of probability tables. The graph consists of nodes and arcs representing conditional dependencies. BBN's provide an intuitive visual representation that can aid in clarifying implicit assumptions made by an expert. BBNs can also expose and overcome some of the common psychological biases and fallacies in reasoning due to misunderstanding of probability. However, the most important use of BBNs is in revising probabilities in light of actual observations of events. A software tool, GeNIe, is used that implements algorithms for propagating the results of new evidence through the BBN, as well as providing a graphical user interface to draw the graphs and fill in the probability tables.

Among the results of this investigation is the identification of a number of conditional dependencies among the requirements that are not explicitly represented in the original Benchmark Requirements. Other factors influencing the Benchmark requirements such as user authentication and denial of service to non-users are identified. Additional consideration needs to be given to the relative, quantitative contribution of the degree of belief in an individual requirement to the overall degree of belief in the trustworthiness of the recordkeeping system.

An approach to testing the BBN model is described using data from a recordkeeping system for scientific data records. The next step is to specify the a priori probabilities and conditional probabilities of the model through consultations with the developers of the Benchmark Requirements.

Keywords: trusted recordkeeping system; authentic copies of digital records, decision support

1. Introduction

1.1 Background

Archivists who must preserve digital records may need to attest to the authenticity of the records in their custody. This can be problematic in the case of digital records, because of the technological environment in which the records were created and maintained. Unless electronic recordkeeping systems are used to maintain essential metadata, the contextual information necessary for understanding a record is at risk of loss. Unless integrity and access control mechanisms were in place during maintenance of the records, an archivist may not be able to conclude that records were not altered, deleted or replaced.

The Authenticity Task Force (ATF) of the InterPARES Project investigated the essential attributes of authentic digital records and the technological and procedural factors impacting their authenticity. They formulated the conceptual requirements for presuming the authenticity of the digital records maintained by a records creator, which they terms Benchmark Requirements. They also formulated similar requirements for records maintained by the institution responsible for the long-term preservation of the digital records, which they termed the Baseline Requirements. They formulated a method for assessing a presumption of authenticity of records based on the degrees of belief that the Benchmark or Baseline Requirements are supported by evidence [Duranti 2005, MacNeil 2000].

1.2 Purpose

The next step is to test the method and the Benchmark Requirements in an assessment of the authenticity of a creator's digital records. The research question arises:

How can the method for assessment of authenticity based on the Benchmark Requirements be more precisely specified and tested so that a preserver could be confident that he could apply the method and Requirements and be confident in the result?

The purpose of this paper is to describe the results of an investigation of representing the Benchmark Requirements as a Bayesian Belief Network. To demonstrate the use of the BBN, data is used from a case study specifically designed for this purpose.

1.3 Scope

In section 2, the ATF's Benchmark Requirements for a presumption of the authenticity of a creators records and the method of assessment are reviewed. In addition, the difficulties and complexities involved in reasoning with degrees of belief are discussed. In section 3, the Bayesian approach to reasoning with degrees of belief and Bayesian Belief Networks (BBNs) are discussed. In section 4, the construction of a BBN based on the Benchmark Requirements is described. In section 5, data collected in an InterPARES 2 Science Focus case study is used to demonstrate the use of the Bayesian Belief Network in supporting the assessment of the degree of belief in the trustworthiness of a recordkeeping system. In section 6, the conclusions of this research are summarized and future work described.

2. The Benchmark Requirements and Method of Assessment

The Authenticity Task Force points out in its report that

In both archival theory and jurisprudence, records that the creator relies on in the usual and ordinary course of business are presumed authentic. However, digital information technology creates significant risks that electronic records may be altered, either inadvertently or intentionally. Therefore, in the case of records maintained in electronic systems, the presumption of authenticity must be supported by evidence that a record is what it purports to be and has not been modified or corrupted in essential respects. [Duranti 2005, Appendix 2]

The Benchmark Requirements for a presumption of authenticity of a records creator’s records are shown in Figure 1. [Duranti 2005, Appendix 2]

To support a presumption of authenticity the preserver must obtain evidence that:	
REQUIREMENT A.1: Expression of Record Attributes and Linkage to Record	<p>the value of the following attributes are explicitly expressed and inextricably linked to every record. These attributes can be distinguished into categories, the first concerning the identity of records, and the second concerning the integrity of records.</p> <p><i>A.1.a</i> Identity of the record:</p> <p><i>A.1.a.i</i> Names of the persons concurring in the formation of the record, that is:</p> <ul style="list-style-type: none"> • name of author • name of writer (if different from the author) • name of originator (if different from name of author or writer) • name of addressee <p><i>A.1.a.ii</i> Name of action or matter</p> <p><i>A.1.a.iii</i> Date(s) of creation and transmission, that is:</p> <ul style="list-style-type: none"> • chronological date • received date • archival date • transmission date(s) <p><i>A.1.a.iv</i> Expression of archival bond (e.g., classification code, file identifier)</p> <p><i>A.1.a.v</i> Indication of attachments</p> <p><i>A.1.b</i> Integrity of the record:</p> <p><i>A.1.b.i</i> Name of handling office</p> <p><i>A.1.b.ii</i> Name of office of primary responsibility (if different from handling office)</p> <p><i>A.1.b.iii</i> Indication of types of annotations added to the record</p> <p><i>A.1.b.iv</i> Indication of technical modifications;</p>
REQUIREMENT A.2: Access Privileges	the creator has defined and effectively implemented access privileges concerning the creation, modification, annotation, relocation, and destruction of records;
REQUIREMENT A.3: Protective Procedures: Loss and Corruption of Records	the creator has established and effectively implemented procedures to prevent, discover, and correct loss or corruption of records;
REQUIREMENT A.4: Protective Procedures: Media and Technology	the creator has established and effectively implemented procedures to guarantee the continuing identity and integrity of records against media deterioration and across technological change;

REQUIREMENT A.5: Establishment of Documentary Forms	the creator has established the documentary forms of records associated with each procedure either according to the requirements of the juridical system or those of the creator;
REQUIREMENT A.6: Authentication of Records	if authentication is required by the juridical system or the needs of the organization, the creator has established specific rules regarding which records must be authenticated, by whom, and the means of authentication;
REQUIREMENT A.7: Identification of Authoritative Record	if multiple copies of the same record exist, the creator has established procedures that identify which record is authoritative;
REQUIREMENT A.8: Removal and Transfer of Relevant Documentation	if there is a transition of records from active status to semi-active and inactive status, which involves the removal of records from the electronic system, the creator has established and effectively implemented procedures determining what documentation has to be removed and transferred to the preserver along with the records.

Figure 1. The Benchmark Requirements for a Presumption of Authenticity of a Creator’s Records

The Benchmark and Baseline Requirements are based on the notion of trust in a recordkeeping system.

Both the benchmark and the baseline requirements are based on the notion of trust in record keeping and record preservation. The benchmark requirements draw specifically on the notion of a *trusted record-keeping system*, and the baseline requirements are predicated on the role of the preserver as a *trusted custodian*.

A *trusted record-keeping system* has been defined as “a type of system where rules govern which documents are eligible for inclusion in the record-keeping system, who may place records in the system and retrieve records from it, what may be done to and with a record, how long records remain in the system, and how records are removed from it.” [Duranti 2005, Part 1]

Duranti [2003] reformulated the Benchmark Requirements in terms of belief in the trustworthiness of a recordkeeping system.

It is essential to ensure that the electronic records are clearly identifiable and of demonstrable integrity and that accidental corruption or purposeful tampering have not occurred since their creation. How do we do so? [By] Maintaining the records in a trusted recordkeeping system.

A trusted record-keeping system comprises the whole of the rules that control the creation, maintenance, and use of the records of the creator and that provides a circumstantial probability of the authenticity of the records within the system.

The method for inferring a degree of belief in the authenticity of a creators records based on the degree of belief that each of the benchmark requirements is warranted by the evidence is stated in part one of the report and in Appendix 2 [Duranti 2005].

A presumption of authenticity is an inference that is drawn from known facts about the manner in which a record has been created, handled, and maintained. The evidence that supports the presumption that the creator created and maintained its electronic records authentic are enumerated in the Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records (Requirements Set A). A presumption of authenticity will be based upon the number of requirements that have been met and the degree to which each has been met. The requirements are, therefore, cumulative: the higher the number of satisfied requirements, and the greater the degree to which an individual requirement has been satisfied, the stronger the presumption of authenticity. This is why these are termed ‘benchmark’ requirements.

Although the method of assessment is expressed in simple terms, there are substantial pitfalls inherent in subjective probability assessment due to psychological biases and common misunderstandings of probabilistic reasoning [Kahneman et al 1982, Wright and Ayton 1994]. Furthermore, the conditional dependencies between requirements and between the evidence needed to conclude that a requirement is met can be quite complex.

- The method described for inferring a degree of belief in the authenticity of a creator’s records does not specify quantitatively how the degrees of belief in the individual requirements contribute to the degree of belief in the authenticity of the records.
- The method described for inferring a degree of belief in the authenticity of a creator’s records misleads those who might apply the method by stating “The evidence that supports the presumption that the creator created and maintained its electronic records authentic are enumerated in the Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records.” For instance, requirement A.5 is “the creator has established the documentary forms of records associated with each procedure ...” Evidence that this requirement is met might consist of seeing and reading a document that shows for each business procedure the corresponding documentary forms. Alternatively, one might be told by a competent person, that there was a document specifying the documentary forms. These different types of evidence might lead to different degrees of belief that requirement A.5 was met. In any case, the requirement itself is not the evidence.
- Suppose a preserver has just seen that the name of the addressee was not included in the metadata associated with a record. His subjective estimate of the likelihood that the record system will contain other records in which the name of the addressee is not associated with a record will temporarily rise. The event is salient, and therefore more available. But the judgment may be in error.
- When assessing the authenticity of digital records, does the uncertainty concerning the authenticity of the records lie in the preserver, or is it a property of the records system? Does uncertainty come from within the preserver, or is it an intrinsic property of events in the environment? If you chose the second option, that uncertainty is a property of events in the environment, then you are subject to the fallacy of denying uncertainty. You

believe you can control it. For example, you may be a skillful preservation manager who believes she can avoid uncertainty and reduce risk by your action. But the uncertainty is attributable to you. Recognition of this fallacy is essential to any attempt to quantify subjective probabilities about uncertain events.

- Suppose that the preserver observes the metadata attributes associated with a record. The metadata should contain the name of the author and the name of the addressee of a record, but not all metadata for records includes the name of the author and the name of the addressee. Which is more probable: The metadata contains both the name of the author and the name of the addressee, or the metadata contains only the name of the addressee? If you selected the first option you are incorrect. The probability of a conjunction $P(A \& B)$ cannot exceed the probability of either $P(A)$ or $P(B)$.
- Overconfidence occurs when accumulating evidence, for example, from case-study material about authentic records, from which certain predictions are then made. Confidence in one's conclusions continues to rise as more information is received. Towards the end of the evidence-gathering process, most people are overconfident about their judgments.
- Two preservers applying the Benchmark Requirements to a recordkeeping system, and who have the same evidence can infer degrees of belief as to the trustworthiness of the system. One reason this can occur is that they have different preferences with regard to risk. Risk takers will tend to overestimate and risk adverse people will tend to underestimate.

The Bayesian approach to reasoning under uncertainty is an approach to reasoning with degrees of belief while dealing with the complexity of conditional dependencies. Combined with Bayesian Belief Networks, it can also expose and overcome some of the common psychological biases and fallacies in reasoning due to misunderstanding of probability.¹ Hence, the original research question might be reformulated as:

Can the method for assessment of authenticity based on the Benchmark Requirements be more precisely specified and tested using Bayesian Probability and Bayesian Belief Networks so that a preserver could be confident that he could apply the method and Requirements and be confident in the result?

¹ There are other methods of modeling decision making under uncertainty, for example, Dempster-Shafer, Truth Maintenance, Fuzzy Logic, and Logical Probability, that would lead to similar research questions.

3. Subjective Probability and Bayesian Inference

Bayesian probability is a formal notation and theory that allows one to reason about beliefs under conditions of uncertainty. A person's subjective belief in a statement H will depend on some body of evidence E . This can be represented as the conditional probability $P(H|E)$ that a hypothesis H is true (e.g., that a requirement is met) given available evidence E . The expression $P(H|E)$ is a measure of a person's belief in the truth of H warranted by the evidence E [Howson and Urbach 1993].

The definition of the conditional probability of A given that B is true or known is the joint probability of A and B divided by the probability of B .

$$P(A|B) = P(A, B)/P(B)$$

It follows as a theorem (known as Bayes rule) that

$$P(A|B) = P(B|A) P(A) / P(B)$$

Bayes rule can be thought of as a means of updating ones belief about a hypothesis A in light of new evidence B . Specifically, ones posterior belief $P(A|B)$ is calculated by multiplying their prior belief $P(A)$ by the likelihood $P(B|A)$ that B will occur if A is true.

3.1 Bayesian Belief Networks

A Bayesian Belief Network (BBN) is a graphical notation with an associated set of probability tables [Pearl 1988]. The graph consists of nodes and arcs representing conditional dependencies $P(A_1|A_2, \dots, A_n)$. The key feature of BBNs is that they enable one to model conditional dependencies of variables and to reason using degrees of belief. BBN's provide an intuitive visual representation that can aid in clarifying implicit assumptions made by an expert. With BBNs, it is possible to articulate expert beliefs about the dependencies between different variables. BBNs can also expose and overcome some of the common psychological biases and fallacies in reasoning due to misunderstanding of probability. However, the most important use of BBNs is in revising probabilities in light of actual observations of events.

3.2 GeNIe and SMILE

There are software tools that implement the algorithms for propagating the results of new evidence through the BBN, as well as providing a graphical user interface to draw the graphs and fill in the probability tables. GeNIe (Graphical Network Interface) is a graphical user interface to SMILE (Structural Modeling, Inference, and Learning Engine) used to create models for probabilistic reasoning and decision-making under uncertainty. It was developed by the Decision Systems Laboratory of the School of Information Sciences at the University of Pittsburg.²

² <http://genie.sis.pitt.edu/about.html>

Figure 2 shows the GeNIe interface for constructing BBN models. The icons support the creation of nodes and dependencies among nodes.

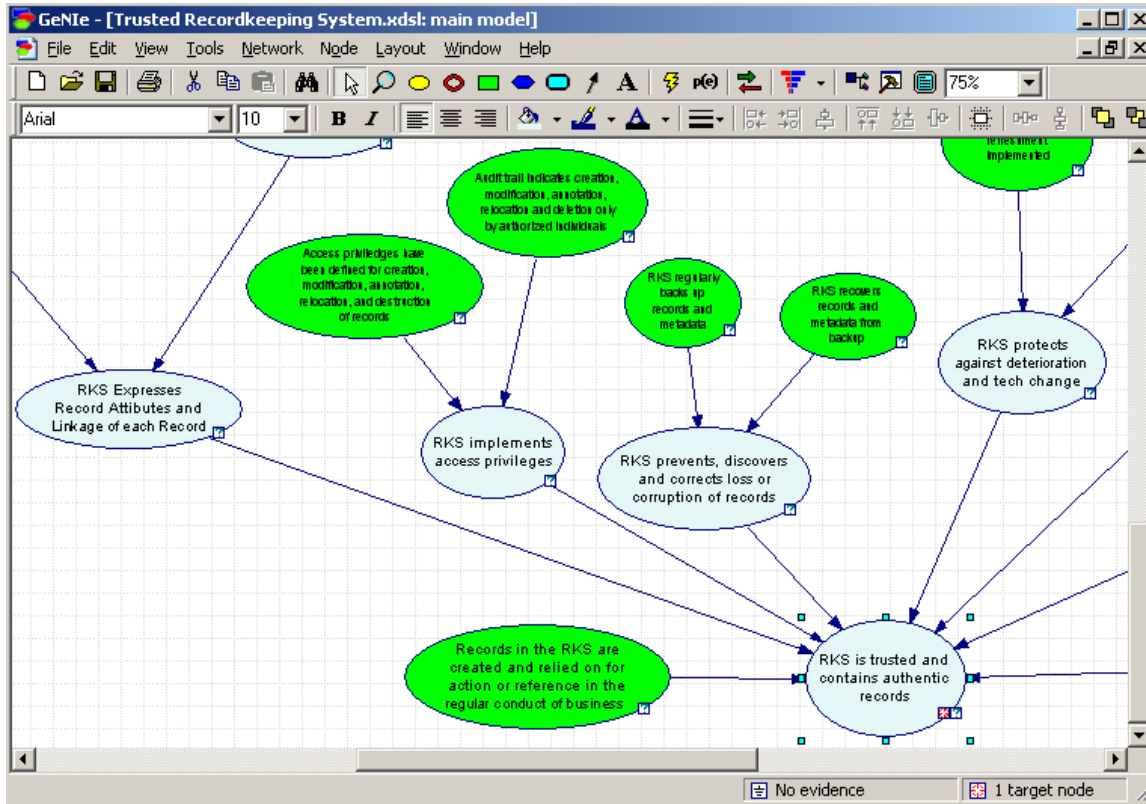


Figure 2. The GeNIe Interface

Figure 3 shows the properties of a node and, specifically, a table for defining conditional probabilities among nodes.

Node properties: RKS is trusted and contains authentic records

	Satisfied		Unsatisfied		Satisfied		Unsatisfied	
Records in the RKS are created and relied on for action or...								
RKS Expresses Record Attributes and Linkage of each Re...								
RKS implements access privileges								
RKS prevents, discovers and corrects loss or corruption of...								
RKS protects against deterioration and tech change								
The creator has established the documentary forms of rec...								
If authentication is required, the creator has established sp...								
If semi-active or inactive records are removed from the RK...								
trusted	1	0.95	0.95	0.9	0.95	0.9	0.9	0.85
not_trusted	0	0.05	0.05	0.1	0.05	0.1	0.1	0.15

Figure 3. Definition of Conditional Dependencies Among Nodes in a BBN

GeNIe was used to define the BBN Model for the Benchmark Requirements that is described in the next section.

4. A BBN Supporting the Assessment of the Degree of Belief that a Recordkeeping System is Trusted and Contains Authentic Digital Records

To develop a BBN, one must elicit expert knowledge of the concepts and causal connections in the Benchmark Requirements that will comprise the topology of the network, and the probabilities comprising each node's probability table [Meyer and Booker 1991]. A test is described in which a human subject acting as a preserver/assessor used case study data from an actual records system to estimate the degree of belief in the terminal nodes of the BBN.

Following Duranti [2003] we will state the primary hypothesis as:

A Recordkeeping System (RKS) is trusted and contains authentic records

The degree of belief in this hypothesis depends on the degree of belief in the requirements (subhypotheses) listed below

A.1 RKS expresses Record Attributes and Linkages to Record

A.2 RKS expresses access privileges

A.3 RKS prevents, discovers and corrects loss or corruption of records

A.4 RKS protects identity and integrity of records against media deterioration and technological change

A.5 Documentary form associated with each business procedure has been established

A.6 If authentication is required by the juridical system or the needs of the organization, the creator has established specific rules regarding which records must be authenticated, by whom, and the means of authentication

A.7 If multiple copies of the same record exist, the creator has established procedures that identify which record is authoritative;

A.8 Procedures are established and implemented for removal and transfer of semi-active or inactive records and associated documentation

Since, as the ATF pointed out, "records that the creator relies on in the usual and ordinary course of business are presumed authentic," an additional requirement will be added that "Records in the RKS are created and relied on for action or reference in the regular conduct of business." Not to do so would imply that the presumption of authenticity did not depend on the fact that the records were those of the creator.

Using GeNIe, these dependencies are represented graphically as shown in Figure 4.

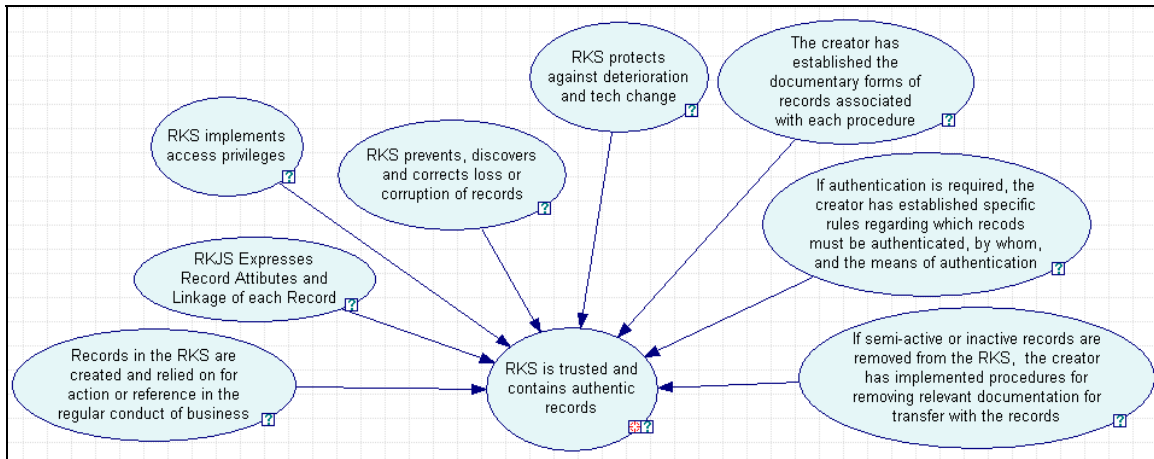


Figure 4. Graph showing the dependencies between the hypothesis and the requirements

Requirement A1 has two sub-requirements.

- A.1.a RKS expresses the identity of each record
- A.1.b RKS expresses the integrity of each record

These dependencies are represented by the graph shown in Figure 5.

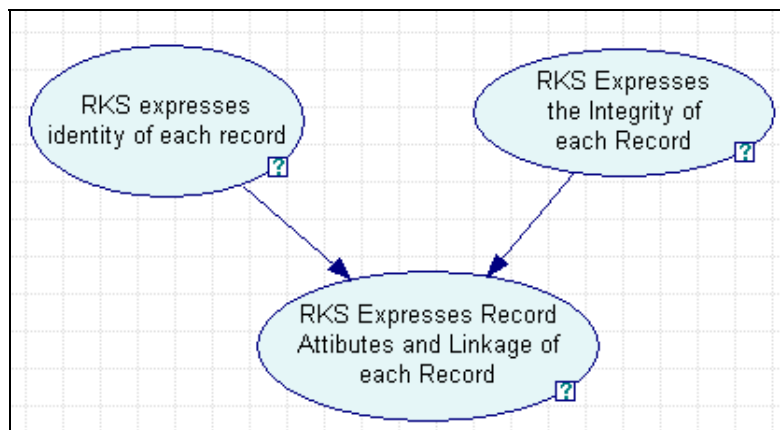


Figure 5. Dependence of Requirement A1 on Requirements A1a and A1b

The degree of belief in Requirement A.1.a, RKS expresses the identity attributes of each record, is dependent on the following requirements.

- A.1.a.i Record attributes include name of author and writer, if different from author
- A.1.a.ii Record attributes include name of action or matter
- A.1.a.iii Record attributes include chronological date, archival date, transmission date, and received date
- A.1.a.iv Record attributes include classification code
- A.1.a.v Record attributes include indication of attachments

These dependencies are at the left of the graph shown in Figure 6.

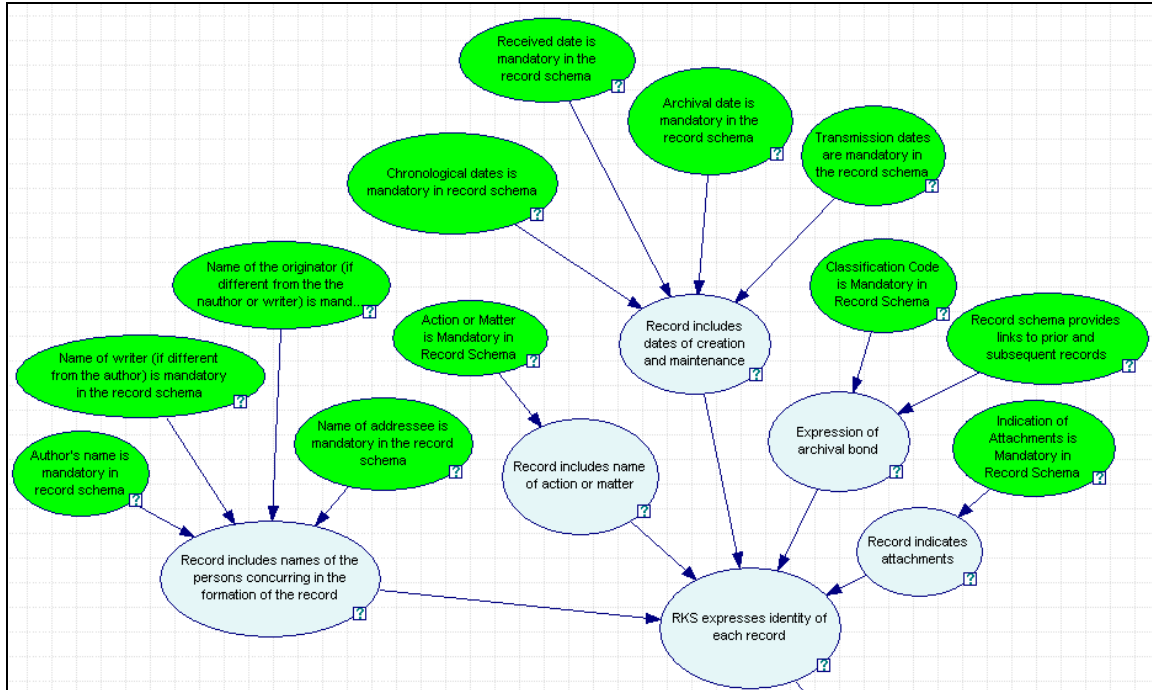


Figure 6. Dependence of Requirement A.1.a on other Requirements.

As shown in Fig. 6, the degree of belief in requirement A.1.a.i, Names of the persons concurring in the formation of the record, is dependent on belief in the statements

- Authors name is mandatory in the record schema
- Name of writer (if different from the author) is mandatory in the record schema
- Name of the originator (if different from the author or writer) is mandatory in the record schema
- Name of the addressee is mandatory in the record schema

As shown in Fig. 6, the degree of belief in requirement A.1.a.iii, Record includes dates of creation and maintenance, is dependent on

- Chronological date is mandatory in record schema
- Received date is mandatory in record schema
- Archival date is mandatory in record schema
- Transmission dates are mandatory in record schema

Also, as shown in Fig. 6, the degree of belief in requirement A.1.a.iv, expression of archival bond, is dependent on

- Classification code is mandatory in record schema
- Record schema provides links to prior and subsequent records

The degree of belief in requirement A.1.b, RKS expresses the integrity attributes of each record, is dependent on requirements

A.1.b.i Record attributes include name of handling office

A.1.b.ii Record attributes include name of office of primary responsibility

A.1.b.iii Record attributes include indication of types of annotations added to the record

A.1.b.iv Record attributes include indication of technical modifications

These dependences are represented in the graph shown in Figure 7.

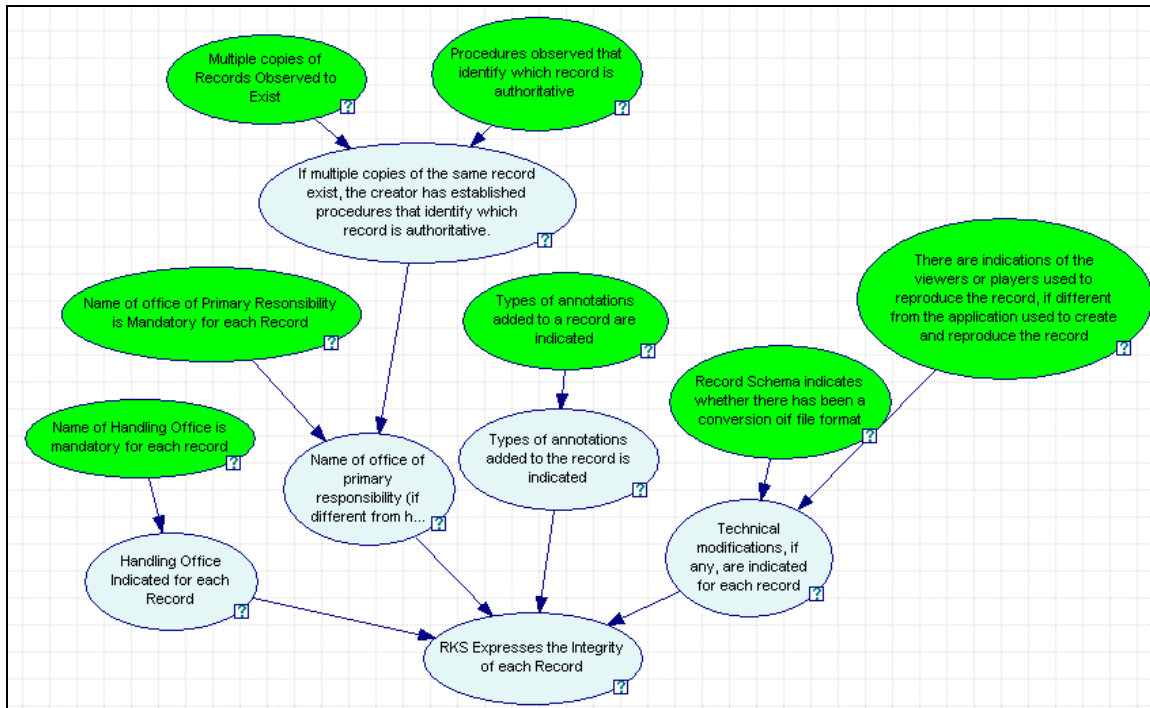


Figure 7. Dependence of Requirement A.1.b on other Requirements

Requirement A.7, titled Authoritative records, is “If multiple copies of the same record exist, the creator has established procedures that identify which record is authoritative.” As shown in Fig. 5, Requirement A.1.b.ii is also dependent on requirement A.7, “If multiple copies of the same record exist, the creator has established procedures that identify which record is authoritative.” When Requirement A.7 is discussed later in this section, it will be explained why this is the case.

There are two technological requirements that should be considered for inclusion in the Benchmark Requirements. The first is that a message digest should be computed for each record. The second is that when a record is accessed or transmitted, there should be check using the message digest of the record’s integrity. Requirement A.1.b would be dependent on these requirements

A third candidate requirement for inclusion in the Benchmark Requirements is that the record attributes should include the document type of the record. This would seem necessary because

requirement A.5, “The creator has established the documentary forms of records associated with each procedure,” is an essential resource for determining the authenticity of records.

Requirement A.2 concerning access privileges is that “the creator has defined and effectively implemented access privileges concerning the creation, modification, annotation, relocation, and destruction of records.” However, in the “Commentary on the Benchmark Requirements” it is pointed out that “Effective implementation involves the monitoring of access through an audit trail that records every interaction that an officer has with each record” [Duranti 2005, Appendix 2]. Separating the definition of access privileges from their effective implementation, the requirement for an audit trail is made an explicit requirement in the BBN as shown in Figure 8.

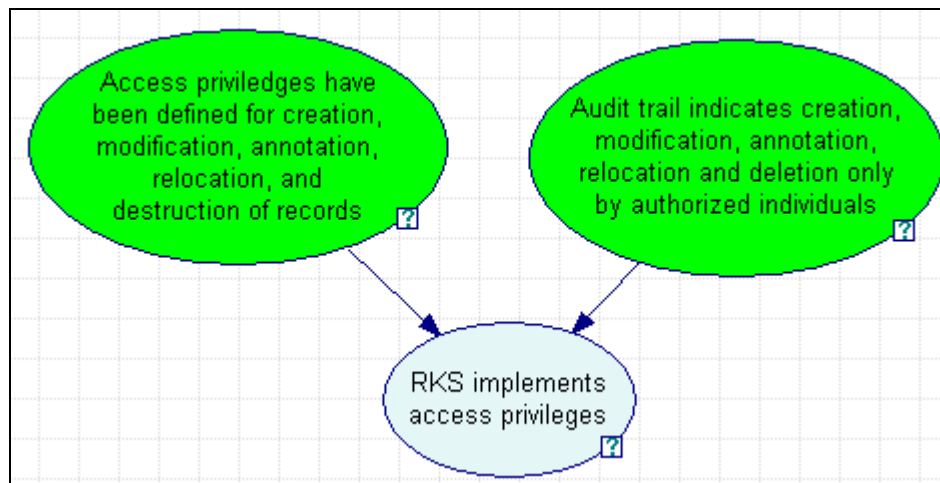


Figure 8. Dependence of Requirement A.2 on other Requirements

In addition to definition and effective implementation of access privileges, there are other requirements that are necessary to secure records and records metadata from unauthorized modification, deletion or insertion [Underwood 2003]. To name a few, the RKS or operating system in which the RKS is embedded should

- authenticate all users
- deny service to non-users
- prevent unauthorized modification of records

Dependence of Requirement A2 on these requirements should be considered for inclusion in the Benchmark Requirements.

Requirement A.3 is that the RKS prevents, discovers and corrects loss or corruption of records. In the “Commentary on the Benchmark Requirements”, it is pointed out that “Procedures to protect records against loss or corruption include: prescribing regular back-up copies of records and their attributes...” and that it is possible to recover these records and their attributes from the back-up copies. Consequently, requirement A.3 depends on the following requirements.

- RKS regularly backs up records and metadata
- RKS recovers records and records metadata from backup

This is reflected by the dependencies shown in Fig. 9.

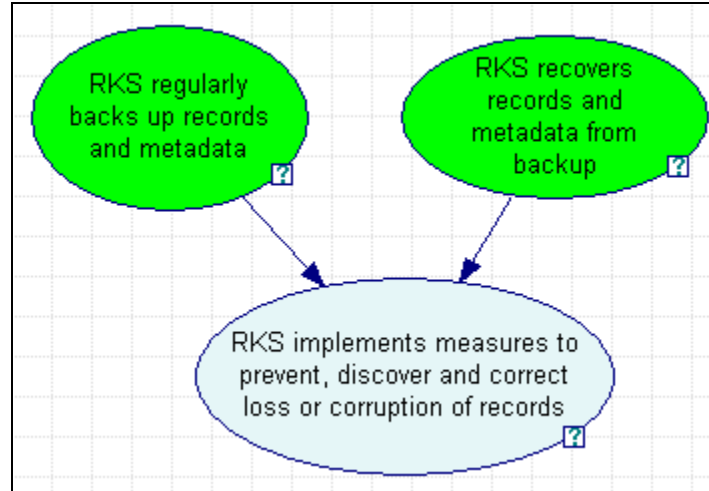


Figure 9. Dependence of Requirement A.3 on other Requirements

Other threats to loss or corruption of records include viruses, Trojans and worms. These threats can be prevented by use of Virus scanners. Dependence of Requirement A.3 on this requirement should be considered for inclusion in the BBN.

Requirement A.4 is that “the creator has established and effectively implemented procedures to protect the continued identity and integrity of records against media deterioration and across technological change.” The “Commentary on Benchmark Requirements” identifies media refreshment and migration of records from an obsolescent technology to new technologies as two of these procedures. This is reflected in the BBN by the dependencies shown in Figure 10.

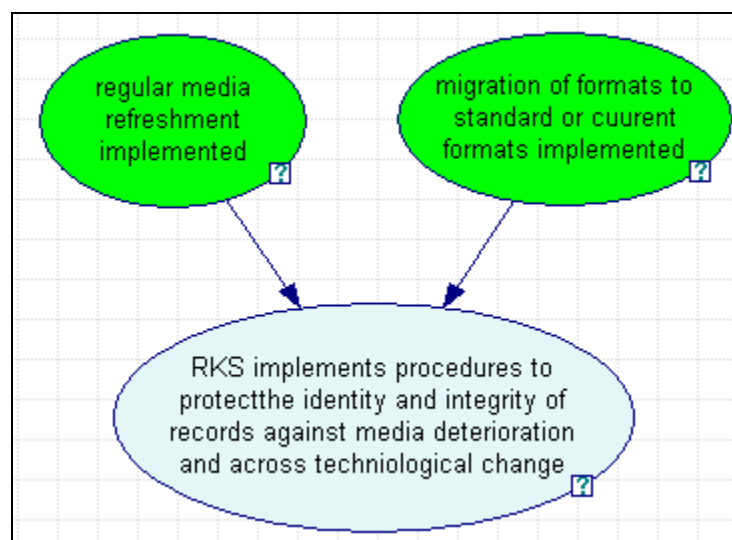


Figure 10. Dependence of Requirement A.4 on other Requirements

Requirement A.5 is that the creator has established the documentary forms of records associated with each procedure. As shown in Figure 11, this can be made more explicit by specifying not only that the documentary forms are associated with each procedure but that the description of the documentary forms includes the definition of both intrinsic and extrinsic elements.

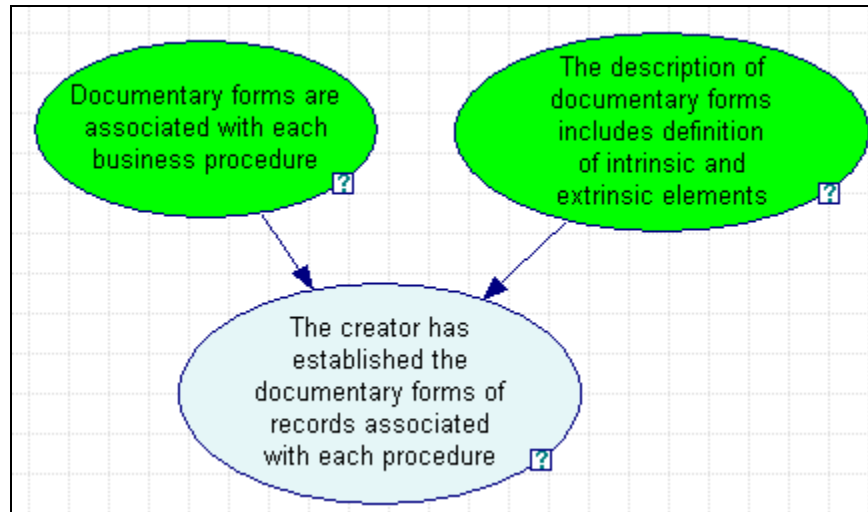


Figure 11. Dependence of Requirement A.5 on other Requirements

Requirement A.6, Authentication of Records, is that “if authentication is required by the juridical system or the needs of the organization, the creator has established specific rules regarding which records must be authenticated, by whom, and the means of authentication.”

This requirement is a conditional statement, that is, it is only necessary for the creator to have established specific rules regarding which records must be authenticated, by whom, and the means of authentication, if authentication is required. There are several ways to represent this in a BBN. The one adopted is to create two statements. The first is “Authentication is required of some records in the RKS.” The second is: “The creator has established rules regarding which records must be authenticated by whom and the means of authentication.” The requirement A.6 is dependent on both, and the dependency is defined by a table associated with the requirement statement A.6. This table is shown in Fig. 12.

Node properties: If authentication is required, the creator has established specific rules regarding which records must be authenticated

General Definition Format User properties

Add
 Insert
 X
 Copy
 Paste
 Undo
 Redo
 Find
 Help
 %
 t=0

Authentication is required of some records in the RKS		True		False	
The creator has established rules regarding which records must be authenticated		True	False	True	False
<input checked="" type="checkbox"/> Satisfied		1	0	1	1
<input type="checkbox"/> Unsatisfied		0	1	0	0

Figure 12. Truth Table showing Dependence of Requirement A.6 on its Component Statements

This table represents the conclusion that the Requirement A.6 is satisfied if its condition, “Authentication is required of some records in the RKS,” is true and its consequent is true. The requirement is also satisfied, if the condition is false. The requirement is only unsatisfied, if the condition is true and the consequent is false.

The graphical representation of these dependences is shown in Fig. 13.

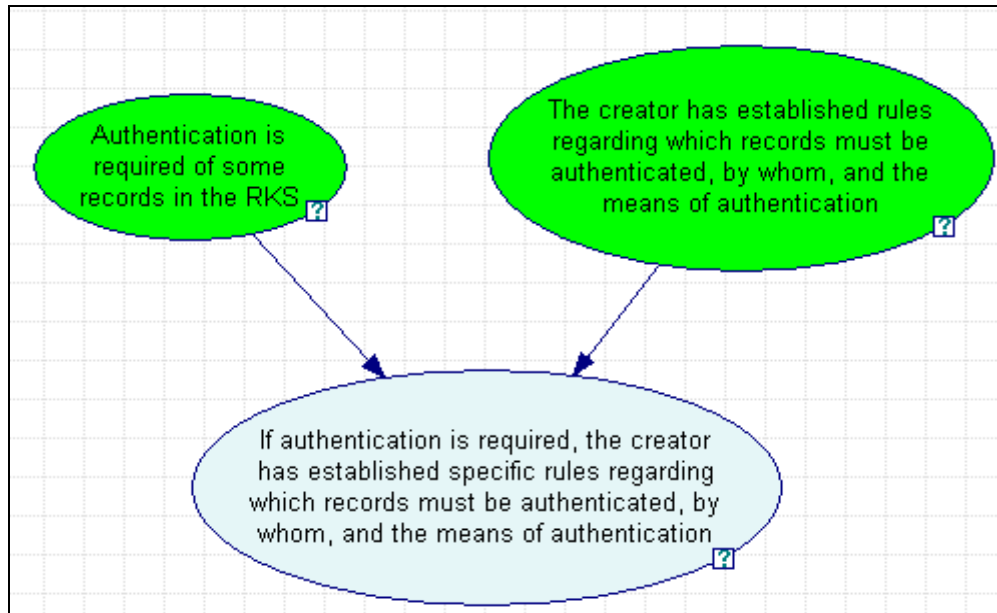


Figure 13. Conditional Requirement A.6

Requirement A.7, titled Authoritative records, is “If multiple copies of the same record exist, the creator has established procedures that identify which record is authoritative.” As was the case for requirement A.6, this requirement is a conditional requirement. The dependency of the requirement on its condition and its consequent are as represented in Fig. 14.

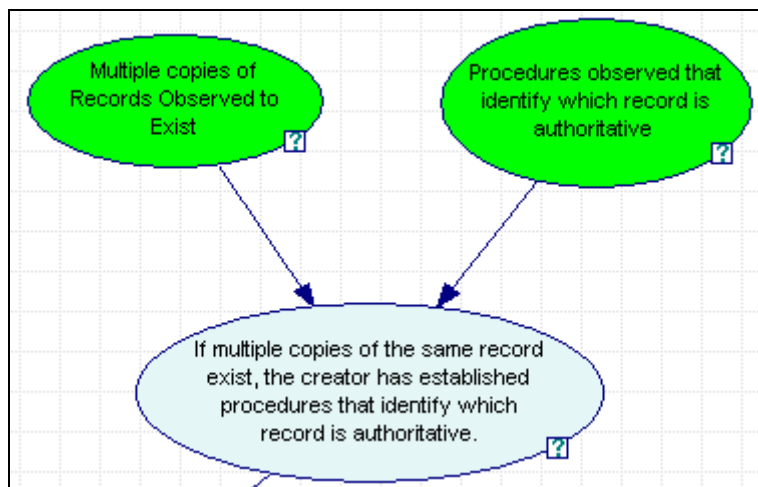


Figure 14. Conditional Requirement A.7

The dependency between the requirement and the statements of its condition and consequent are defined by a table similar to the one shown for requirement A.6

The primary hypothesis of the model should not be directly dependent on requirement A.7. Rather requirement A.1.b.ii, Record attributes include name of office of primary responsibility, is conditionally dependent on requirement A.7 as is shown in Fig.7. The rationale for this conclusion can be found in the “Commentary on the Benchmark Requirements [Duranti 2005, Appendix 2]. In commenting on requirement A.7, Identification of Authoritative Record, it is stated that:

The identification of authoritative records corresponds to the designation of an office of primary responsibility as one of the components of a record retention schedule. The Office of Primary Responsibility is the office given the formal competence for maintaining the authoritative (that is, official) records belonging to a given class within an integrated classification scheme and retention schedule. The purpose of designating an Office of Primary Responsibility for each class of record is to reduce duplication and to designate accountability for records.

If the record attributes include the Office of Primary Responsibility, then it must be because there are procedures that identify the authoritative record.

Requirement A.8 is that “if there is a transition of records from active status to semi-active and inactive status, which involves the removal of records from the electronic system, the creator has established and effectively implemented procedures determining what documentation has to be removed and transferred to the preserver along with the records.” This requirement is a conditional, so it will be dependent on the truth or falsity of its condition, as was the case for requirements A.6 and A.7.

The “Commentary on the Benchmark requirements” states for Requirement A.8 that

This requirement implies that the creator needs to carry forward with the removed records all the information that is necessary to establish the identity and demonstrate the integrity of those records, as well as the information necessary to place the records in their relevant contexts.

Thus, it is implied that the consequent of the requirement is dependent on two sub-requirements.

- A procedure that removes the information necessary to establish the identity of the records, and
- A procedure that removes the information necessary to demonstrate the integrity of the records.

These dependencies are reflected in the diagram shown in Fig. 15.

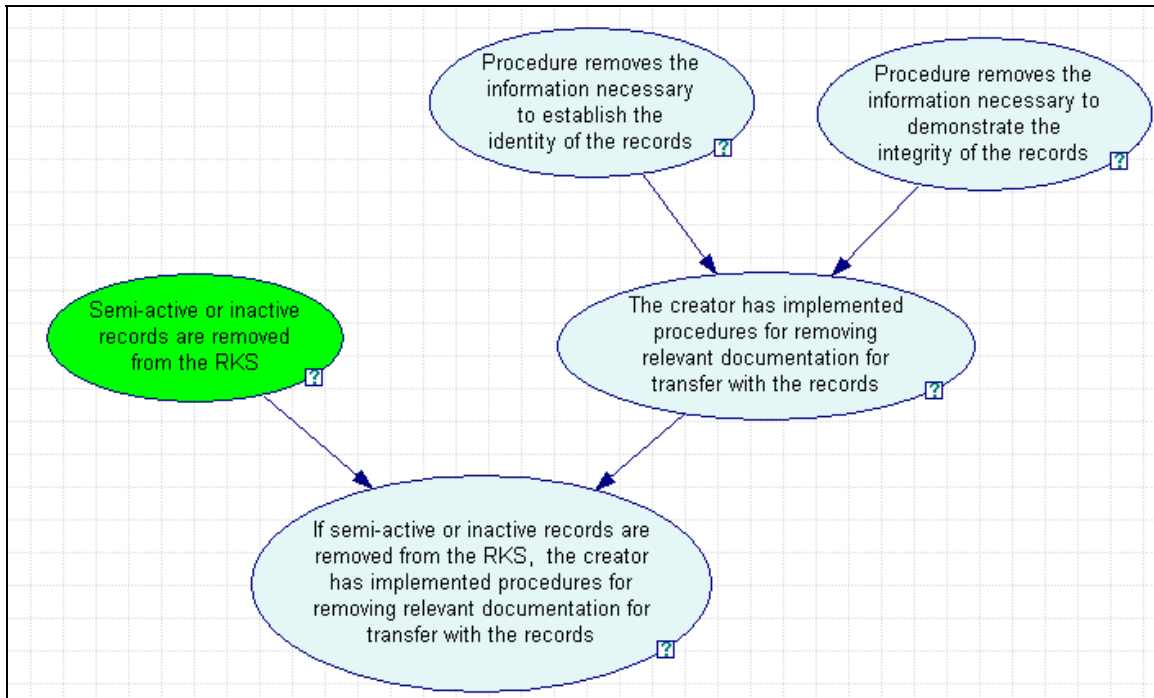


Figure 15. Conditional Requirement A.8

Requirement A.8 is conditionally dependent on requirements A.1.a and A.1.b. One cannot remove information necessary to identify records and demonstrate the integrity of records unless one has met the requirement of identifying and expressing the integrity of the records in the first place.

The graphical structure of the entire BBN for the Benchmark Requirements is shown in Fig. 16.

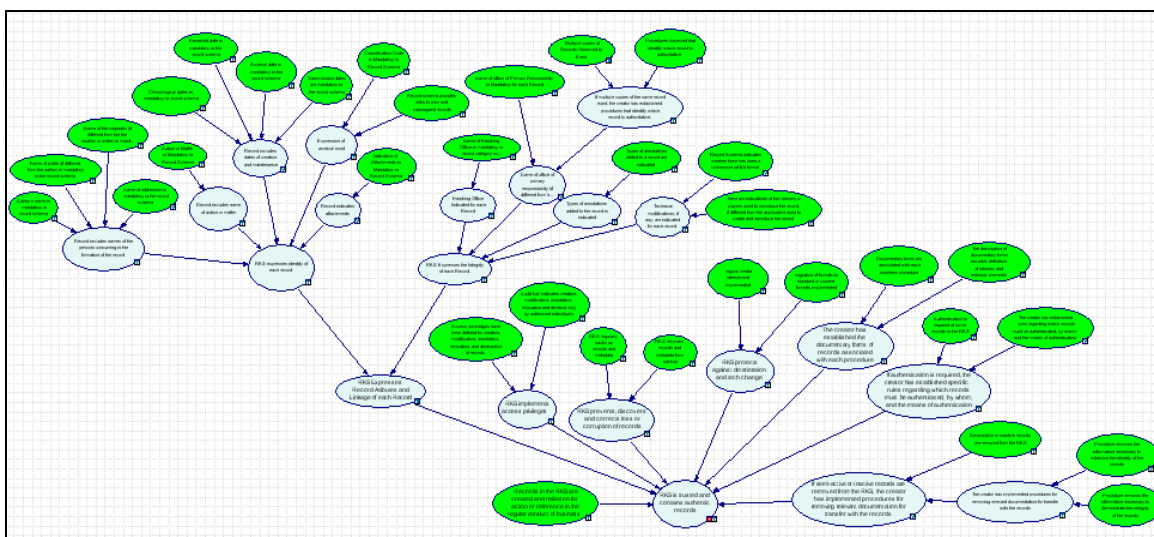


Figure 16. The BBN Model for the Benchmark Requirements

5. Using the BBN to Support Assessment

In this section, data collected in the Planetary Data System (PDS) case study will be used to demonstrate the use of the Bayesian Belief Network in supporting the assessment of the degree of belief in the trustworthiness of the PDS in maintaining authentic data records.

5.1 The Mars Global Explorer Data Records in the Planetary Data System

The Planetary Data System case study was designed to collect information to answer InterPARES 2 core research questions regarding the creation, management and preservation of e-science data records [Underwood 2007]. It was also designed to collect information to perform a walkthrough of the InterPARES 2 activity model of records creation, maintenance, appraisal, and preservation of authentic digital records [Underwood et al 2007]. It was also designed to collect data that could be used to test a BBN model of the Benchmark Requirements.

The focus of the case study is the data records of the Mars Global Surveyor (MGS) mission and the Planetary Data System (PDS). The MGS was the first mission of the National Aeronautics and Space Administration (NASA) Mars Surveyor Program. The MGS spacecraft was launched in 1996. Last contact with the spacecraft was in November 2006. The MGS spacecraft contains instruments that sent observational data back to the Jet Propulsion Laboratory Spacecraft Control Center where it was stored as experimental data records in a Project Database. Spacecraft Operations Planning Computers at the Principal Investigator's site, and at seven discipline nodes—Atmospheres, Geosciences, Imaging, Navigation Ancillary Information Facility, Plasma Physics, Rings and Small Bodies—were used to create reduced data records.

A data product is a labeled, reduced data record resulting from a scientific observation, usually stored in one file. Templates are used to create product labels that identify describe, and define the structure of the data. Examples of data products are a planetary image, a spectrum table, or a time series table. An archive volume is data products plus documentation and ancillary files. Volumes are transferred to the Science Data Validation Team for peer review. If judged complete, reliable and PDS-compliant, they are cataloged and stored in the PDS archive for access by planetary scientists.

The PDS is referred to as an “active archive,” whereas the National Space Science Data Center's (NSSDC's) repository is referred to as a “deep archive.” The PDS is the entrance for Planetary Science data into the NSSDC archives for long-term preservation. It was a conclusion of the walkthrough of the Manage Chain of Preservation model using the PDS Case Study data that the PDS is a record-making and recordkeeping system [Underwood et al 2007].

5.2 Assessment of the Trustworthiness of the Planetary Data System

In this section, it is demonstrated that the BBN model of the Benchmark Requirements can be used with GeNIe to assess the degree of belief that can be afforded PDS as a trusted recordkeeping system. For each requirement at the terminal nodes of the BBN, evidence will be provided when available. Then a degree of belief will be assigned to that requirement given the evidence. These degrees of belief will be entered into the BBN and GeNIe will compute the

degree of belief in the primary hypothesis that “the PDS is a trusted RKS that maintains authentic digital records.”

For each of the Requirements listed below,

- The evidence for the requirement will be described.
- Based on the evidence, the degree of belief that the authors have that the requirement is true will be indicated. The table shown in Figure 17 is used as a guide for expressing verbally the subjective belief. When the degrees of belief are entered into the BBN, the corresponding numerical expression is selected to express the subjective probability that the requirement is true.

Verbal Expression³	Numerical Expression
Certainly true	1.0
Almost certainly true	.95
Highly probable that true	.90
Very probably true	.85
Quite likely true	.80
Good chance that true	.75
Rather likely true	.70
Better than even chance that true	.60
Slightly more than half the time true	.55
Toss-up or No evidence	.50
Slightly more than half the time false	.45
Better than even chance that false	.40
Rather likely false	.33
Good chance that false	.25
Quite likely false	.20
Very likely false	.15
Very likely false	.10
Almost certainly false	.05
Certainly false	0.0

Figure 17. Correspondence of Verbal Expressions of Degrees of Belief to Numerical Expressions

1. The records in the RKS are relied on for action or reference in the regular conduct of business.

Evidence: There are hundreds of journal articles that are written by members of the MGS team and planetary scientists that reference the scientific data. For example,

M. C. Malin, K. S. Edgett, L. V. Posiolova, S. M. McColley and E. Z. N. Dobrea (2006), “Present-day impact cratering rate and contemporary gully activity on Mars,” *Science* 314(5805): 1573–1577, [doi:10.1126/science.1135156](https://doi.org/10.1126/science.1135156).

Degree of belief that statement is true: Certainly true

³ Table adapted from Hamm [1988].

- The name of the author is mandatory in the record schema.

Evidence: The PDS Data Preparation Workbook [JPL 1995] prescribes that PDS Templates be used in creating data products. The templates for data products all contain the PRODUCER_ID whose value is the author of the data product (see Fig. 18). Every data product is checked to be sure that all product labels in a volume contain this attribute and a value before it is stored in the PDS archive.

Degree of belief that statement is true: Certainly true

```

PDS_VERSION_ID = PDS3
FILE_NAME = "/tmp/S0902603.img"
RECORD_TYPE = FIXED_LENGTH
RECORD_BYTES = 2048
FILE_RECORDS = 6657
LABEL_RECORDS = 1
^IMAGE = 2
SPACECRAFT_NAME = MARS GLOBAL SURVEYOR
MISSION_PHASE_NAME = "SUPPORT"
TARGET_NAME = MARS
INSTRUMENT_ID = "MOC-NA"
PRODUCER_ID = MGS_MOC_TEAM
DATA_SET_ID = "MGS-M-MOC-NA/WA-2-SDP-I0-V1.0"
PRODUCT_CREATION_TIME = 2005-12-28T22:01:28
SOFTWARE_NAME = "makepds 1.11"
UPLOAD_ID = "UNK"
PRODUCT_ID = "S09/02603"
START_TIME = 2005-08-26T03:29:46.82
IMAGE_TIME = 2005-08-26T03:29:46.82
STOP_TIME = 2005-08-26T03:29:50.03
SPACECRAFT_CLOCK_START_COUNT = "809494240.80"
SPACECRAFT_CLOCK_STOP_COUNT = "N/A"
FOCAL_PLANE_TEMPERATURE = 275.7 <K>
GAIN_MODE_ID = "0A"
OFFSET_MODE_ID = "34"
LINE_EXPOSURE_DURATION = 0.482100 <MILLISECONDS>
DOWNTRACK_SUMMING = 1
CROSSTRACK_SUMMING = 1
EDIT_MODE_ID = "0"
RATIONALE_DESC = "Monitor gullies and bright features
in S05-01463 and E11-03412"
DATA_QUALITY_DESC = "OK"
ORBIT_NUMBER = 28896
OBJECT = IMAGE
LINES = 6656
LINE_SAMPLES = 2048
LINE_PREFIX_BYTES = 0
LINE_SUFFIX_BYTES = 0
SAMPLE_TYPE = UNSIGNED_INTEGER
SAMPLE_BITS = 8
SAMPLE_BIT_MASK = 2#11111111#
CHECKSUM = 16#FF208C26#
END_OBJECT = IMAGE
END

```

Figure 18. A PDS Label for a Mars Orbital Camera Image of the Martian Surface

- The name of the writer, if different from the author, is mandatory in the record schema.

Evidence: The writer is the same as the author in all cases.

Degree of belief that statement is true: Certainly true

4. The name of the originator (if different from the author or writer) is mandatory in the record schema.
Evidence: The PDS does not store records received from other organizations or copies of records sent to other organization. Hence, there are no originators of data records. The statement does not apply to PDS data products.⁴
Degree of belief that statement is true: Toss-up or no evidence.
5. The name of addressees is mandatory in the record schema.
Evidence: The PDS does not store copies of records sent to other organization. Hence, there are no addressees of data records. The statement does not apply to PDS data products
Degree of belief that statement is true: Toss-up or no evidence
6. Action or matter is mandatory in the record schema.
Evidence: The PDS Data Preparation Workbook [JPL 1995] prescribes that PDS Templates be used in creating data products. The templates for data products all contain the RATIONALE_DESCRIPTION whose value is the action or matter of the data product. Every data product is checked to be sure that all product labels in a volume contain this attribute and a value before it is stored in the PDS archive.
Degree of belief that statement is true: Certainly true
7. Chronological date is mandatory in the record schema
Evidence: The PDS Data Preparation Workbook [JPL 1995] prescribes that PDS Templates be used in creating data products. The templates for data products all contain the PRODUCT_CREATION_TIME whose value is the chronological date of the data product. Every data product is checked to be sure that all product labels in a volume contain this attribute and a value before it is stored in the PDS archive.
Degree of belief that statement is true: Certainly true
8. Archival date is mandatory in the record schema.
Evidence: Archival date is the date the volume of data products is stored in the PDS archive. Data was not collected during the case study on whether an archive date was or was not associated with a volume.
Degree of belief that statement is true: Toss-up or no evidence
9. Transmission date is mandatory in the record schema.
Evidence: The PDS does not store copies of records sent to other organizations. Hence, there is no transmission date for data records. There is an instrument observation date, but that is not what is meant by transmission date. The statement does not apply to PDS data products.⁵
Degree of belief that statement is true: Toss-up or no evidence

⁴ The statement of this requirement and the next should probably be rephrased. For the purpose of this demonstration, the degree of belief will be "Toss-up or no evidence."

⁵ The statement of this requirement and the next should probably be rephrased. For the purpose of this demonstration, the degree of belief will be "Toss-up or no evidence."

10. Received date is mandatory in the record schema.

Evidence: The PDS does not store records received from other organizations. Hence there is no received date for data records. The statement does not apply to PDS data products.

Degree of belief that statement is true: Toss-up or no evidence

11. Classification code is mandatory in the record schema.

Evidence: The following attributes in the templates for PDS labels of the data product are assigned values (classification codes). [JPL 1995]

DATA_SET_ID
SPACECRAFT_NAME
INSTRUMENT_ID
TARGET_NAME

DATA_SET_ID is similar to a record series id. SPACECRAFT_NAME, INSTRUMENT_ID, and TARGET_NAME are similar to subject classification codes. Classification of supplementary documents corresponds to assigning a value to the keyword DOCUMENT_TOPIC_TYPE.

Degree of belief that statement is true: Certainly true

12. Record Schema provides links to prior and subsequent records

Evidence: Sequence of data records is indicated by the relative time of observations.

Degree of belief that statement is true: Certainly true

13. Indication of attachments is mandatory in the record schema.

Evidence: The Data Preparation Workbook [JPL 1995] prescribes that data products be written to a volume. The files in a volume are organized starting at the root directory. Below the root directory is a directory tree containing, in addition to the data records, the documentation and index files. These seem to correspond to attachments. The volume verifier checks that the prescribed documentation and index files are on the volume.

Degree of belief that statement is true: Certainly true

14. Name of handling office is mandatory in the record schema.

Evidence: The PDS is a distributed archive. The catalog at the Central Node indicates the Node at which the volumes of data records are kept, which is the handling office and the office of primary responsibility.

Degree of belief that statement is true: Certainly true

15. Name of office of primary responsibility (if different from handing office) is mandatory in the record schema.

Evidence: The PDS is a distributed archive. The catalog at the Central Node indicates the Node at which the volumes of data records are kept, which is the handling office and the office of primary responsibility.

Degree of belief that statement is true: Certainly true

16. Multiple copies of records observed to exist.
Evidence: Since copies of records are not sent to other offices, multiple copies of data records do not exist.
Degree of belief that statement is true: Certainly false
17. Procedures observed that identify which copy of multiple records is authoritative.
Evidence: Since copies of records are not sent to other offices, multiple copies of data records do not exist, and procedures are not needed to identify which copy of records is authoritative.
Degree of belief that statement is true: Certainly false
18. Indication of types of annotations added to the record is mandatory in the record schema.
Evidence: Every data product is annotated. The possible annotations are listed in the Data Preparation Workbook [JPL 1995].
Degree of belief that statement is true: Certainly true
19. Record schema indicates whether there has been a conversion of file format.
Evidence: If data is converted to other formats, for example, for publication in a journal article or for publication on the web, the original source of the data is indicated. It is not necessary to convert the files to other formats for preservation purposes, since the PDS data records are self describing files that include a description of the file format.
Degree of belief that statement is true: Certainly true
20. There are indications of the viewers or players used to reproduce the record if different from the application used to create and reproduce the record.
Evidence: NASAView can reproduce all of the data records.⁶
Degree of belief that statement is true: Certainly true
21. Access privileges have been defined for creation, modification, annotation, relocation and destruction of records.
Evidence: Access to PDS resources is controlled through user authentication. Access to restricted areas of the PDS is determined by the user's assigned role. In the PDS, roles may include subscriber, Data Engineer, System Administrator, etc. This type of access control is called role-based access authentication. The PDS logs accesses to restricted areas of the system. User ID, date, time and operations are logged.
Degree of belief that statement is true: Certainly true
22. Audit trails indicate creation, modification, annotation, relocation and deletion only by authorized individuals.
Evidence: The PDS logs accesses to restricted areas of the system. User ID, date, time and operations are logged.
Degree of belief that statement is true: Certainly true
23. RKS regularly backs up records and metadata.
Evidence: The PDS Central Node provides system backup.
Degree of belief that statement is true: Certainly true

⁶ A copy of NASAView can be obtained at <http://pdsproto.jpl.nasa.gov/Distribution/license.html>.

24. RKS recovers records and metadata from backup.
Evidence: PDS recovers data records and metadata from backups.
Degree of belief that statement is true: Certainly true
25. Regular media refreshment implemented.
Evidence: In the PDS, ensuring data integrity and usability requires periodic data renewal cycles. Some such cycles involve media refreshment.
Degree of belief that statement is true: Certainly true
26. Migration of formats to standard or current formats implemented.
Evidence: Records are created in PDS standard formats. Thus, migration of formats is not required. Since data records are already in standard formats, it was concluded that this statement was certainly true.
Degree of belief that statement is true: Certainly true
27. Documentary forms are associated with each business procedure.
Evidence: PDS templates for labels are defined for all created objects [JPL 1995]
Degree of belief that statement is true: Certainly true
28. The description of documentary forms includes definition of intrinsic and extrinsic elements.
Evidence: The attributes of a template for labels comprise the intrinsic elements of documentary form.
Degree of belief that statement is true: Certainly true
29. Authentication is required of some records in the RKS.
Evidence: Although scientists and PDS managers do not use the term authentic to characterize the data products in the PDS, they are concerned that all the data records are complete, reliable, accurate, and that the integrity of the data record is assured.
Degree of belief that statement is true: Certainly true
30. The creator has established rules regarding which records must be authenticated, by whom, and the means of authentication.
Evidence: Peer review ensures that data records are complete, reliable and accurate.
Degree of belief that statement is true: Certainly true
31. Semi-active or inactive records are removed from the RKS.
Evidence: The PDS is called an active archive. It contains records of all NASA Planetary science spacecraft missions. The PDS contains semi-active and inactive records, but they have not been removed. However, copies of inactive and semi-active datasets are transferred to the National Space Science Data Center (NSSDC). So, it was decided that this statement was certainly true.
Degree of belief that statement is true: Certainly true

32. The creator has implemented procedures for removing relevant documentation for transfer with the records.

Evidence: When copies of the dataset volumes are transferred to the NSSDC, all relevant documentation is included in the volumes of a dataset.

Degree of belief that statement is true: Certainly true

The next step is to translate the verbal expressions for degrees of belief to numerical expressions of subjective probability and enter them into the BBN. Figure 19 shows the GeNIe interface for doing this.

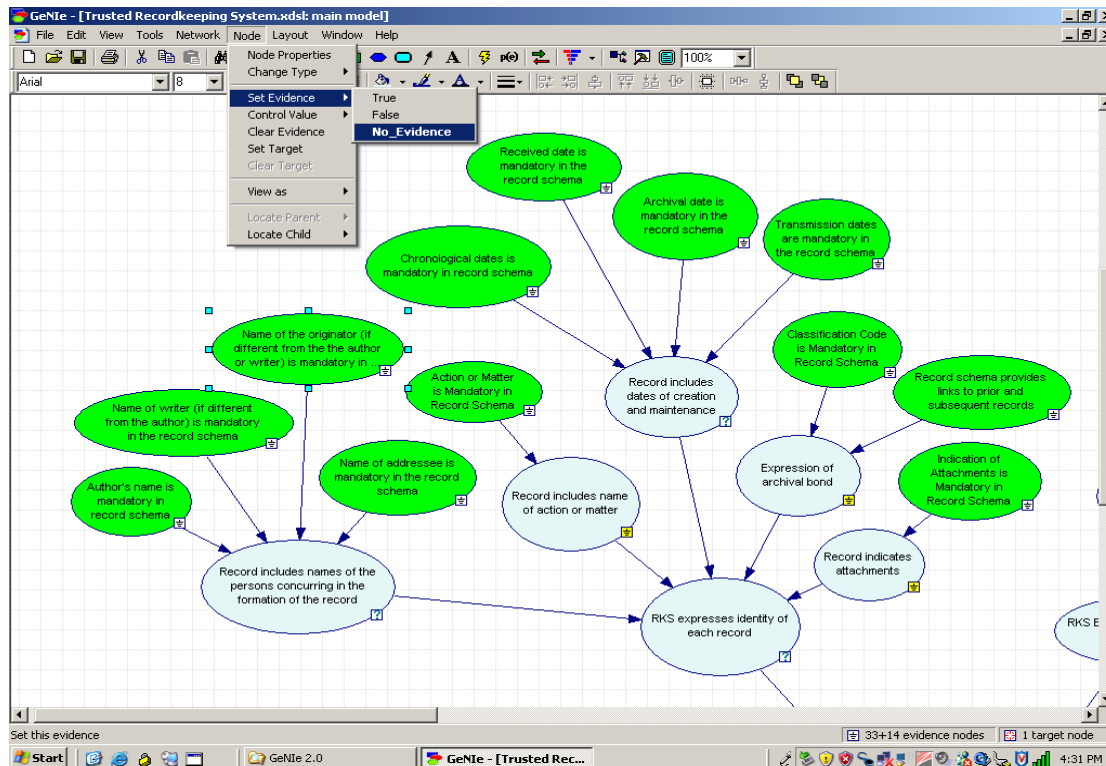


Figure 19. Using GeNIe to Set Evidence of Nodes in the BBN.

When all of the subjective probabilities have been entered, belief are updated for all of the requirements and the primary hypothesis that the “RKS is trusted and contains authentic records.” The results of the test will not be reported yet because not enough thought has been put into the a prior probabilities and conditional probabilities. This will require consultation with the archival scientists who created the Benchmark Requirements.

GeNIe is a tool for creating and refining BBNs. SMILE, the software package on which GeNIe is based, enables us to create user interfaces that resemble the list of requirements shown in this section, to allow the user to enter descriptions of the evidence they have found, and to enter verbal descriptions of their degrees of belief in the requirements warranted by the evidence. The user interface then could translate this to the numerical subjective probabilities and update the beliefs. The results of the assessment, including the evidence and degrees of belief, can be saved as a record of the assessment.

6. Conclusion and Areas of Future Research

After reviewing the Authenticity Task Force's Benchmark Requirements and method of assessment, the difficulties inherent in subjective probability assessment were discussed. It was suggested that the Bayesian approach to reasoning with degrees of belief combined with Bayesian Belief Networks could overcome some of the common psychological biases and fallacies in reasoning due to misunderstanding of probability.

The Baseline requirements and their conditional dependencies were represented as a Bayesian Belief Network using the GeNIe graphical user interface. A number of conditional dependencies in the Benchmark Requirements were exposed. Other factors influencing the Benchmark requirements such as user authentication and denial of service to non-users were identified. An approach to testing the BBN model was described using data from a recordkeeping system for scientific data records. The next step is to specify the a priori probabilities and conditional probabilities of the model through consultations with the developers of the Benchmark Requirements.

The correctness and accuracy of the BBN model can be tested by constructing additional test cases for which optimal decisions are known. The BBN model can be tested on these cases and further refined and improved. Domain experts should review the results generated by the model and their opinions should be incorporated into the model.

References

Decision Systems Laboratory. GeNIe Online Help

<http://genie.sis.pitt.edu/GeNIeHelp/index.html>

[Duranti 2003] L. Duranti, Trusted Recordkeeping System, Seminar entitled “Lost in Migration”, sponsored by Superintendent of Library and Documentary Resources of the IBC in collaboration with the Archival National Association of Italy, Dec 5, 2003, Bologna, Italy.

<http://ibc.regione.emilia-romagna.it/soprintendenza/htm/interpares.htm>

[Duranti 2005] L. Duranti (ed.) *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*. Archilab, San Miniato, Italy, 2005.

<http://www.interpares.org/book/index.cfm>

[Hamm 1988] R. M. Hamm. Evaluation of a Method of Verbally Expressing Degree of Belief by Selecting Phrases from a List. Publication No. 88-6, Institute of Cognitive Science, University of Colorado, April 1988.

[Howson and Urbach 1993] C. Howson and P. Urbach. *Scientific Reasoning: The Bayesian Approach*. Second Edition. Open Court, Chicago, 1993

[JPL 1995] Planetary Data System Data Preparation Workbook, Version 3.1, JPL D-7669, Part 1, Feb. 1995, Jet Propulsion Laboratory, Pasadena.

[Kahneman et al 1982] D. Kahneman, P. Slovic and A. Tversky (eds). *Judgement under Uncertainty: Heuristics and Biases*. Cambridge University Press, 1982.

[MacNeil 2000] H. MacNeil. Providing Grounds for Trust: Developing Conceptual Requirements for the Long-Term Preservation of Authentic Electronic Records. *Archivaria* 50 (Fall 2000): 56-67.

[Meyer and Booker 1991] M. Meyer and J. Booker. Eliciting and Analyzing Expert Judgment. A Practical Guide. *Knowledge Based Systems* Volume 5. Academic Press, 1991.

[Underwood 2003] B. Underwood. Informatica verso teoria archivistica a diplomatica: un’interazione da coltivare. Seminar entitled “Lost in Migration,” sponsored by Superintendent of Library and Documentary Resources of the IBC in collaboration with the Archival National Association of Italy, Dec 5, 2003, Bologna, Italy.

<http://ibc.regione.emilia-romagna.it/soprintendenza/htm/interpares.htm>

[Underwood 2005] W. Underwood (2005), “InterPARES 2 Project - Case Study 08 Final Report: Mars Global Surveyor Data Records in the Planetary Data System.”

http://www.interpares.org/display_file.cfm?doc=ip2_cs08_final_report.pdf

[Underwood et al 2007] W. Underwood, K. Glick and M. Wolfe (2007), “InterPARES 2 Project - General Study 12 Final Report: Validation of the InterPARES 2 Project Chain of Preservation

Model Using Case Study Data.”

http://www.interpares.org/display_file.cfm?doc=ip2_gs12_final_report.pdf

[Wright and Ayton 1994] G. Wright and P. Ayton (eds). *Subjective Probability*. John Wiley & Sons Ltd., 1994.