Data-PASS Metadata Requirements (version 1.2, approved 12/15/2007)

Introduction

Metadata, "data about data", is structured information that describes a resource, object, or data. A common example is the library catalog entry, which contains information about holdings in a library.

In traditional archives and libraries, metadata has always played a significant role in enabling users to discover and locate resources. In a digital environment, metadata takes on additional roles: It is used to support the federation or interoperation with other libraries and archives. And it is often needed to enable many additional dissemination services.

Moreover, metadata is vital for digital preservation. Unlike traditional preservation, which often strives to protect holdings from change, change is inevitable in digital preservation. Information that is stored on the same physical media will inevitably degrade, and information that is kept in its original format will predictably become inaccessible – to preserve digital objects requires *managed change*, migrating the object to new media and new formats. Managing this process requires appropriate metadata.

1. What the Data-PASS Metadata Does

The requirements described in this document are designed particularly to support two activities: First, these requirements provide the information necessary for the partners in the alliance to validate and assist with the preservation of data that is held within each member archive. Second, these requirements provide a basis for services provided to users across these archives, through such mechanisms as the Data-PASS shared catalog.

More specifically, the metadata requirements are designed to support the following categories of services and functions across federated archives (the last three of which are optional):

- resource discovery
- resource identification and citation
- resource location
- resource administration
- data replication: integrity checking, provenance
- public data dissemination
- access control
- layered services: dissemination reformatting, generating subset of variables, data analysis

A number of considerations influenced the design of these requirements.

First, these requirements are intended to be used for interoperation among heterogeneous data archives. These archives may differ in their use of software and protocols; the manner in which they store their holdings; their cataloging procedures; and in their internal metadata. These metadata requirements are meant to enable interoperation in the absence of such uniformity. In particular, since internal procedures and metadata may differ, these requirements describe a core of common elements that are easily convertible among multiple different schemas, and thus can be used for interchange between archives using differing metadata formats and procedures.

Second, following the same principle, these requirements are intended to be lightweight, since requiring metadata in excess of the minimum needed to ensure preservation of the data it describes would increase the cost of preparing data for preservation. At the same time, additional effort devoted to metadata does have distinct benefits in terms of usability, discovery, and the provision of other services. So, we identify optional metadata elements that support these additional services.

Third, to the greatest extent possible, we have made use of existing standards, schemas and protocols. Rather than invent new elements, we have sought to identify a core of common elements, and practices for using them in conformance with the OAIS reference model, that can be leveraged to provide the services described above. Under this approach, we believe, it is likely that we will be able to build upon the work of the existing archiving and library communities, and take advantage of existing tools.

Finally, we have attempted to use schemas and protocols for which multiple open implementations exist. We have done this for three reasons: First, preservation is more likely to be successful where the infrastructure and procedures for managing and preserving resources are transparent. Second, the partnership is intended to be open to others to join, and open implementations minimize the barriers to participate. Third, a standard supported by multiple independent implementations is likely to be more durable

2. Metadata Types

Metadata may be supplied at multiple levels of granularity. While there is no single set of levels that is universal, it is common in the field of data archiving to describe metadata at the collection, series, study, file, variable group, and variable levels. These metadata requirements focus on the four most fundamental levels:

1. A *study* is meant to indicate a data collection, study, or compilation collected/brought together by at a single time, or for a single purpose, or by a single (although perhaps corporate) investigator. A study can be comprised of multiple files.
2. A *file* (or more technically, a *bitstream*) is a sequence of bits representing data, extended metadata (documentation), or other resources comprising the study.

3. A *variable group* is a logical set of variables which share common attributes, most commonly, that of jointly forming a set of coherent observations (a logical table)
4. A *variable* is a set of observations, conducted with a study, using a single measure. E.g. the income of a family, the consumer price index, the age of a person.

We follow standard practice in distinguishing among three types of metadata:

1. *Descriptive metadata* describes and identifies a resource. Its primary use is to enable users to discover resources of interest. For example, a descriptive title can be used to support a catalog search. (Accompanying documentation, such as 'codebooks', data collection instruments, usage guides, and frequency tables, can be considered extended descriptive metadata intended to support human use and understanding of the data object.)
2. *Structural metadata* facilitates retrieval, navigation, and presentation of a resource. For example structural metadata may be used to bind together a set of multiple files that comprise a study.
3. *Administrative metadata* facilitates management of digital objects. This may include technical information regarding format, creation, and quality control; rights management information regarding usage restrictions, and access control; and preservation action information restriction including provenance and migration activities.

3. Metadata Requirements

3.1. Metadata Exchange

Repository metadata and data exchange is based on a small number of, simple standard, protocols:

1. The OAI-PMH protocol is used to expose metadata for harvesting.
2. The DDI [1] (DDI-lite) [2] metadata format is used as the format for metadata exchange.
3. HTTP/HTTPS is used for data interchange
4. (Strongly recommend) persistent identifiers: CNRI Handles, DOI's , ARK's, Purl's, or URN's, with Handles and DOI's preferred.

Use of these standards for interchange has a number of significant advantages. Used together, in the ways described below, these standards can satisfy the need for federated search, preservation and other layered services. Each of these is well-known within the archiving community, and used in production at multiple major institutions. Furthermore, multiple independent implementations exist to support each technology, with at least one open-source implementation (such as the VDC) available. [3]

3.2. Required and Recommended Fields

The tables below describe the required and recommended metadata fields to be provided in DDI [2] and HTTP. Required fields are denoted with a "*" character. The tables also describe the recommended Dublin Core (DC) mappings for these fields. Fewer than a dozen fields are required to support basic resource discovery, access, and preservation. Recommended fields are used to support more robust format migration, more flexible access control, on-line services (such as analysis), and enhanced discovery (e.g. via geospatial queries). DDI fields other than those listed in the tables are welcome, and will be used for informational/display purposes.

Field Type	*Role*	*Description*	*Use in Standard Schemas*	*Comments*
Title*	Descriptive	Title of study	DDI: <titl> DC: Title	
Author*	Descriptive	Authoring Entity	DDI: <AuthEnty> DC: Creator	
Description*	Descriptive	Study Abstract	DDI: <abstract> DC: Description	
Identifier*	Descriptive	Identifier for collection	DDI: <IDNo agency=””> DC: Identifier	- A persistent globally unique identifier is required. - CNRI handles or DOI's are recommended. VDC will automatically supply on ingest. - Additional local-archive identifiers may be included. - Should explicitly specify the identify agency/authority
Publication Date*	Descriptive	Publication/Production date of study	DDI: <prodDate> DC: Date	
Subject	Descriptive	Subject Keywords	DDI: <keyword vocab=””> DC: Subject	- Studies collected specifically using NDIIPP funding should contain: <keyword vocab="DATAPASS">DATAPASS:NDIIPP</keyword> - Data-PASS studies can have additional keywords. - It is recommended that keyword elements specify vocabulary from which topic is drawn
Publisher[4]	Descriptive/ Administrative - Provenance	Producer of Data Collection	DDI: <producer> DC: Publisher	- No DC equivalent - Embedded <extLink> can be used to provide publisher logo for display in catalog
Distributor[4]	Descriptive/ Administrative - Provenance	Archive distributing data	DDI: <distStmt>	- No DC equivalent - Embedded <extLink> can be used to provide Archive logo for display in catalog
Data Sources	Descriptive	Data sources	DDI: <sources> DC: Source	
Time Period	Descriptive	Time Period Covered by Data	DDI: <timePrd> DC: Coverage/Temporal	
Collection Date	Descriptive	When data was collected	DDI: collDate DC: Coverage/Temporal	
Geographic Coverage	Descriptive	Geographic coverage	DDI: geogCover DC: Coverage/Spatial	- Should specify textually - May specify with latitude/longitude, to support geospatial queries
Kind of Data	Descriptive	Kind of Study	DDI: <dataKind> DC: Type	- In DC, generic type is always "DATASET" - dataKind indicates *subtype* which could be mapped to qualified DC Type with additional qualifier of "datatype"
Notes	Descriptive	Notes field	DDI:<stdyDscr>/<notes>	- Used for additional information - Studies deposited using the Data-PASS deposit agreement should include a documenting note, as per section 4.3, below.

Table 1 Study Level Metadata Fields: Part 1 ***(* indicates required)***

Field Type (indicates required)*	*Role*	*Description*	*Use in Standard Schemas*	*Comments*
Copyright	Administrative: Rights Management	Copyright Information	DDI: Copyright, DC: Rights/License	- Copyright on data - Will be displayed to user
Terms of Use*	Administrative: Rights Management	Confidentiality or other permissions	DDI: useStmt DC: Rights/AccessRights	- Required only if access to data is restricted in some way - General terms displayed - <useStmt/confDec, useStmt/specPerm> *will be used to construct clickthrough agreement before allowing access*
Location*	Structural	URL of study	DDI: <holdings URI> DC:Identifier/URI	- Resolves to study itself at canonical location, - DC qualifier is ambiguous
Version History	Administrative: Preservation		DDI: verStmt DC: provenance	- Machine readable version name, date - Details of reformatting/changes in human-readable form - Strongly recommended - VerStmt and docSrc in <docDscr> can be used to track changes to and provenance of metadata

Table 2: Study Level Metadata (Part 2)

File-level metadata is required if the study is going to be preserved by the partnership. With the exception of *access rights*, all of these fields can be automatically generated when the file is ingested by the VDC.

Field Type (indicates required)*	*Role*	*Description*	*Use in Standard Schemas*	*Comments*
File Local Identifier*	Descriptive	File identifier within study	DDI: <FileDescr ID><OtherMat ID>	
File Name	Descriptive	Human readable file name	HTTP: Content-disposition	
File Location*	Structural	URI for file	DDI: <FileDescr URI;OtherMat URI> DC: Relation	
File Description	Descriptive	File description	DDI: <FileDescr labl><OtherMat labl>	
File Format	Administrative: Preservation/Structural	File format	HTTP/MIME: Content-type DDI: <fileType>	- standard content/types can interoperate with extended mime content/types - extended types (both Mime and VDC) support preservation activities - in future will be replaced by format registry
File Fingerprint*[5]	Administrative: Preservation	Cryptographic Hash of File	HTTP: MD-5 VDC: UNF	- Eventually will be supplanted by replaced by stronger hash type - For datasets, UNF is format-independent method of checking content : http://thedata.org/index.php/Main/UNF - - Stored in typed <note> in DDI
Modification Date	Administrative: Preservation	Modification date of file	DDI: < VerStmt> HTTP: Last-Modified	File level <verStmt>, not study-level
File Role	Structural	Role of file in Study	DDI: Implied by use of <FileDescr> vs. <OtherMat>	
Access[6]	Administrative: Rights	Determines access to resource	HTTP: 403/401 Status Code VDC: Access Classes	VDC access control classes allow federated access control, fine-grained control

Table 3: File Level Metadata

Variable-level metadata is not required. If this metadata is supplied, variable identifiers are required, and other fields are recommended. Table 4 below, shows these fields. Normally, all of these metadata fields, with the exception of *concept*, and *question text* (if it is not already incorporated in the description), could be automatically generated and/or extracted at ingest from the original statistical application file (e.g. SPSS portable file).

Field Type (indicates required)*	*Role*	*Description*	*Use in Standard Schemas*	*Comments*
Variable Identifier*	Structural/Citation	Unique identifier for variable with study	<var ID>	- Unique within Study, usually automatically generated - Must be maintained consistently across versions of study if variable-level citation is supported - Mandatory if supplying variables
Variable Name	Structural/Descriptive	Variable name (could be same as above)	<var name>	Usually automatically imported from statistical file
Variable Measurement Type	Structural	Measurement level (ordinal, nominal, continuous, ratio)	<var intrvl>	Usually automatically imported from statistical file
Var location	Structural	Location of data for variable	<var/location>	Location in bitstream resource – resolves to file ID or URI Variables within same location assumed to be comparable
Var Description	Descriptive	Description of variable	<var/labl>	Usually automatically imported from statistical file
Question Text	Descriptive	Question Text	<var/qustn>	
Missing values	Descriptive/Structural	Indicates missing Values	<invalrng>	Usually automatically imported from statistical file
Sumstat	Descriptive	Simple summary statistics	<sumStat>,<catStat>	Usually automatically generated on ingest
Category Values	Descriptive/Structural	Values for each category of variable	<catValu>, <cagry/labl>	Usually automatically imported from statistical file
Concept[7]	Grouping concept	Grouping concept	<concept>	

Table 4: Variable (and variable group) Level Metadata

3.3. Terms of Use and Access

The particular terms of use for a study may be arbitrarily complex. Rather than develop controlled vocabularies for all conceivable access, we will define metadata standards based on the following four broad categories.

In addition, Data-PASS partners have developed a standard deposit agreement, which applies to cases in which data is shared among the partners in some way (categories 1-3). In these cases, the depositing partner *should* include either:

(1) As a linked file resource, a digitized copy of either the Data-PASS data deposit agreement pertaining to that study, *or* the Archive's grant of permission statement to Data-PASS. (Both of these documents are available on the Data-PASS web site).

Along with this agreement/permission statement, for ease in automated identification, the depositing partner should include the following *notes*:

> <notes level=”study” source=”producer” type=”DATAPASS:TERMS:STANDARD:1.0” subject=”STANDARD DEPOSIT TERMS 1.0”>This study was deposited under the terms of the Data-PASS standard deposit terms. A copy of the usage agreement is included in the file section of this study.</notes>

The note should be placed at the end of the <stdyDscr> section. A similar *notes* element with type “DATAPASS:DEPOSIT:TERMS:STANDARD:1.0:FILE” should be included with the usage terms file. And a *keyword* as in table 1 (above) should be added, if the data was collected using NDIIPP funding. The type and subject of the *notes*, and the vocabulary and content of the *keyword, should* be copied exactly. Additional information may be supplied in the text of the *notes*, but will not affect automated processing. See Appendix 3 for an example.

(2) Or, the archive should include the permission statement directly in meta-data, as follows:

> <notes level=”study” source=”producer” type=”DATAPASS:TERMS:STANDARD:1.0” subject=”STANDARD DEPOSIT TERMS 1.0”>[NAME OF ARCHIVE] (“the Archive”) gives permission and any required licenses to Data-PASS to make the Content available for archiving, preservation and access, within the Data Preservation Alliance for the Social Sciences ("Data-PASS") in accordance with the Data-PASS terms (the "Terms") of use (available from: http://www.icpsr.org/Data-PASS). Including permission to: (a) redisseminate copies of the Data Collection in a variety of distribution formats only according to the standard terms of use of the Archive (See [STANDARD TERMS URL]); (b) promote and advertise the Data Collection in any publicity (in any form) for Data-PASS and the Archive; (c) describe, catalog, validate and document the Content; (d) store, translate, copy or re-format the Data Collection in any way to ensure its future preservation and accessibility, improve usability and/or protect respondent confidentiality; (e) incorporate metadata (cataloging information) or documentation regarding this study into public access catalogues. The Archive represents and warrants that the Content conforms to all Terms, and that the Archive is lawfully entitled and has full authority to license Data-Pass to use the Materials in the ways described in the Terms.</notes>

Additional information *should* be provided based on the general usage category:

Category 1:
The study may be used, redistributed (by the partners/LOC), and transferred (by the partners/LOC) to preservation storage without restriction.

In this case, no usage metadata is required.

Category 2:
The study may be redistributed, and transferred to preservation storage without restriction. And, there are restrictions on use that do not require human mediation.

In this case, copyright and/or terms of use metadata should be supplied as described in Tables 1 & 2. Such usage and copyright terms will be displayed to the user prior to allowing the user before the distributor allows access any resource associated with the study, and the user will be required **accept a click-wrap agreement to the usage terms** before the distributor allows the user to access any resource associated with the study.

Category 3:
The study may be transferred by LOC/partners to preservation storage without restriction. And, there are restrictions on redistribution and/or use of all or part of the study that require human mediation (e.g. verifying a signature, establishing membership, review committee).

In this case, the terms of use should describe the criteria for authorization, in whole, or by reference to documentation included with the study, or externally. These terms of use will be displayed to the user, who will be redirected to the holdings page at the source archive for more information.

The source archive *may* choose to permit such access by assigning them a controlled VDC access "class" and making use of VDC federated access control mechanisms to permit selective redistribution. A set of standard Data-PASS access classes will be documented and published on the Data-PASS web site.

Sites not using VDC access classes *should* return a HTTP 401 to partner sites attempting to access these files, requiring explicit authorization.

Category 4:
There are restrictions on transferring the study to preservation storage.

Case 4 is exceptional. In this case, the terms of use *should* describe the criteria for authorized access to the study, in whole, or by reference to documentation included with the study, or externally. These terms of use will be displayed to the user, who will be redirected to the holdings page at the source archive for more information. A

"STANDARD DEPOSIT TERMS" *notes* field *should not* be included in this document.

In addition, files that are restricted in this manner *should not* provide URI attributes. This indicates the resource is not available on-line to the partnership.

3.4. Resource Formats

Until a global machine-accessible format registry is in place, and content-types are supplemented with controlled vocabulary and Jhove plugins for validation, we will document preferred formats, their corresponding extended content-types, and related documentation used by the partnership. This information will be available from the main Data-PASS site, as well as from the library of congress format registry. The VDC system currently provides some format validation tools, and will work to incorporate the Jhove tools for recommended formats, as the global registry standards become available.

3.5. Implementation Notes:

1. It is *not* required to use DDI internally. DDI is only required as a *metadata exchange mechanism.* It is possible to use any other format, including a METS set (although none currently exists specifically for datasets), MARC (with some uncontrolled/localized fields) or Dublin Core (with qualifiers).

   For study-level descriptive data, MARC and Dublin Core are adequate. However, there is no standardized way to use these to support the administrative and structural requirements for files and variables.

   We rely primarily on the DDI-lite recommended subset of tags from DDI version 2.0. DDI 3.0 will provide more flexible mechanisms to manage provenance of metadata, richer semantics, and more modularity. The VDC team is working with the DDI/SRG standards group, and the recommendations above incorporate current understandings of best practice using the current version of DDI. In additions, automated mappings will be provided so that metadata records created in DDI2.0 are transparently compatible with systems using DDI 3.0

2. With one exception (holdings) the recommended study-level tags above are a subset of the DDI-lite recommended set, to ensure greater interoperability.

3. The VDC provides an open source implementation of these interoperation standards bundled together to provide a complete set of digital library services. Other implementations are available for the standards separately.

4. Also, metadata can document changes to itself through use of <docDescr/verstmt> (analogous to verstmt for studies, above) and provenance through <docSrc>,

<docDescr/distStmt>> and <docDescr/Publisher>. The VDC catalog supports the management and display of chains of provenance information.

5. A UNF is created by rounding data values (or truncating strings) to a known number of digits (characters), representing those values in standard form (as 32bit unicode-formatted strings), and applying a fingerprinting method (such as cryptographic hashing function) to this representation. UNF's are computed from data values provided by the statistical package, so they directly reflect the internal representation of the data - the data as the statistical package interprets it.

   A UNF differs from an ordinary file checksum in several important ways: UNF's are format independent. UNF's are robust to insignificant rounding error. UNF's detect misinterpretation of the data by the statistical software.

   UNF libraries are available for standalone use, for use in C++, and for use with other packages. More information is available here: http://thedata.org/index.php/Main/UNF

6. For many resources, access will be defined to be public (after restrictions have been displayed, and user has clicked through documented access permission forms). These resources can also be made available through the VDC network and the DataWeb network run by the U.S. Census.

   Other data will be available only at the 'home' archive, or otherwise restricted. For restricted data, access can be permitted to a trusted server at a harvested data via IP recognition (or VPN) for the sole purpose of preservation, or could be disseminated to specified groups of authenticated users, based on VDC distributed access control metadata.

7. Ncubes are defined along with concepts in DDI as a mechanism to group variables. Unfortunately, the two current implementations of systems using Ncubes are incompatible. We expect this to be resolved in DDI 3.0

References

[1] The Data Documentation Initiative Specification Website:
<http://www.icpsr.umich.edu/DDI/>

[2] The Virtual Data Center System: <http://theData.org>

[3] RFC2616: The Hypertext Transfer Protocol [HTTP]
< http://www.faqs.org/rfcs/rfc2616.html>

[4] Dublin Core Standard Site:
< http://dublincore.org/ >

[5] RFC 1049: Content-Type Header for Internet messages
< http://www.faqs.org/rfcs/rfc1049.html >

[6] OAI-PMH Harvesting Protocol
< http://www.openarchives.org/>

[7] CNRI Handle System
< http://www.handle.net >

[8] UNF Distribution for the R Statistical Language
<http://cran.r-project.org/src/contrib/Descriptions/UNF.html>

Appendix 1: Examples of Extended Content-Types

Tab separated values
media type: text/tab-separated-values
default file extension: .tsv
reference: http://www.iana.org/assignments/media-types/

Comma separated values
text/csv
default extension:.csv
reference: http://www.iana.org/assignments/media-types/

Spss portable format
application/x-spss-portable; version="6"
default extension:.por
reference: VDC, Dspace

Fixed field data
text/plain; charset="us-ascii"; type="fixed-field"
reference: VDC, Dspace

Appendix 2: Relation to OAIS Reference Model

The OAIS reference model requires that an archival system support the model of information described in it (section 2.2) and be capable of providing support for archival responsibilities (section 3.1).

Section 2.2 describes information transferred to and from an OAIS conforming system:

Content Information. *In the Data-PASS system the Knowledge Base of the Designated Community is embodied in web browsers. The Content Information consists of bit streams with associated HTTP header information including MIME types sufficient for browsers to render the bit stream.*

Preservation Description Information, consisting of Provenance (the source), Context (links to other objects), Reference (identifiers for retrieval) and Fixity. *In the Data-PASS system Provenance is provided by the DDI metadata accompanying the content, Context is provided by the accompanying DDI metadata and by extended metadata, contained in bit-streams linked to by the accompanying metdata, Reference is provided by the CNRI handle or other persistent identifier and by the availability of the text and the metadata it includes to search engines, and Fixity is provided by the mutual mirroring protocol, using file fingerprint metadata, which supplies regular assurance that the content agrees with other replicas.*

Packaging Information. *In the Data-PASS system Packaging Information encoded in the DDI metadata.*

The system is required to support three types of Information Packages:

Submission Information Package (SIP). *In the Data-PASS system SIPs are created by the publisher, who makes available metadata via OAI-PMH and publishes the URL of their OAI-PMH repository. Individual Data-PASS system administrators direct their systems to preserve this page and the content it describes. Their LOCKSS system collects the page and the content it describes.*

Archival Information Package (AIP). *Internally, the Data-Pass system preserves content in a local repository. The AIP consists of a set of bitstreams, representing the content itself, DDI metadata and the HTTP header.*

Dissemination Information Package (DIP). *The LOCKSS system disseminates information by acting as an HTTP server). The DIP is a possibly reformatted version of the SIP, optionally using UNF's to ensure the semantic integrity of reformatting.*

ISO 14721:2003 also requires that Information Packages be associated with Descriptive Information sufficient to locate them. *The Descriptive Information in the Data-PASS system consists of the descriptive DDI metadata fields describe above.*

The mandatory requirements of Section 3.1 apply to the organization operating the OAIS archive, requiring the OAIS conforming system to enable the organization to:

Negotiate for and accept appropriate information from information Producers. The *organization's Data-PASS system will, as directed by the authorized administrator, enter content and information received from Producers in the form of an appropriate SIP*.

Obtain sufficient control of the information provided to the level needed to ensure Long-Term Preservation. *An organization's LOCKSS system will as directed by the authorized administrator, collect via HTTP the entire AIP containing the content and the of partner archives. This information is sufficient at the time of collection for a browser to render the content.*

Determine, either by itself or in conjunction with other parties, which communities should become the Designated Community and, therefore, should be able to understand the information provided. *The Data-PASS SIP provides descriptive data describing types of data sufficient to identify the designated community.*.

Ensure that the information to be preserved is Independently Understandable to the Designated Community. *The DATA-Pass system's AIP provides enough information to support dynamic reformatting to a format that is interpretable by the user's web browser.*

Follow documented policies and procedures which ensure that the information is preserved against all reasonable contingencies, and which enable the information to be disseminated as authenticated copies of the original, or as traceable to the original. *Data-PASS systems preserving the same AIP cooperate to audit and repair it, ensuring that the information is preserved against all reasonable contingencies.*

Make the preserved information available to the Designated Community. *An organization's LOCKSS system's DIP can replicate the AIP exactly.*

Note: The draft above was adapted from the LOCKSS Technical Specification-OAIS conformance document.

Appendix 3: Example Metadata

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
- <codeBook xmlns="http://www.icpsr.umich.edu/DDI"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.icpsr.umich.edu/DDI
    http://www.icpsr.umich.edu/DDI/Version1-3.xsd">
- <stdyDscr>
- <citation>
- <titlStmt>
  <titl>Business Week #3/Harris Poll 2000 National Issues Survey, Study No.
    11767</titl>
  <IDNo agency="odum">H-11767</IDNo>
  <IDNo agency="handle">1902.0/H-11767</IDNo>
    </titlStmt>
- <rspStmt>
    >
  <AuthEnty>Louis Harris and Associates</AuthEnty>
    </rspStmt>
    </citation>
- <stdyInfo>
- <subject>
  <keyword source="archive" vocab="ODUM:MAIN.HEADING">Omnibus and Public
    Opinion Polls</keyword>
    </subject>
  <abstract>This survey focuses on issues and ratings, Y2K, confidence in
    institutions, health behaviors, gay and lesbian rights. Standard demographic
    variables included are: number of adults, age, education, health problem,
    party affiliation, political philosophy, household income 1998, Hispanic
    origin, race, number of telephone lines in household, sex.</abstract>
- <sumDscr>
  <timePrd>January 6-10, 2000</timePrd>
  <dataKind>Numeric (Survey)</dataKind>
    </sumDscr>
    </stdyInfo>
   <notes level ="study" source="producer" type="DATAPASS" subject="STANDARD
    DEPOSIT TERMS 1.0">This study was deposited under the terms of the
    Data-PASS standard deposit terms. A copy of the usage agreement is
    included in the file section of this study</notes>

    </stdyDscr>
- <fileDscr ID="file1" URI="http://vdc-
    demo.hmdc.harvard.edu/VDC/Repository/0.1/Access/hdl:1902.0/H-
    11767/harris_s11767_spss.tab">
- <fileTxt>
  <fileType charset="ISO-8859-1">application/x-spss-por</fileType>
    </fileTxt>
```

```
  <notes subject="Universal Numeric Fingerprint" level="file" source="archive"
    type="VDC:UNF">UNF:3:32:liVW0q7OLlzDgX7b+7CfXg==</notes>
  </fileDscr>
- <dataDscr>
- <var ID="v1.1" name="ID" intrvl="discrete">
  <location fileid="file1" />
  <labl level="variable">id</labl>
  <varFormat type="numeric" />
  <notes subject="Universal Numeric Fingerprint" level="variable"
    source="archive"
    type="VDC:UNF">UNF:3:10:ZJWj+XqVmSu96GDXIjL3rA==</notes>
  </var>
- <var ID="v1.2" name="WT" intrvl="contin">
  <location fileid="file1" />
  <labl level="variable">wt</labl>
  <sumStat type="mean">1</sumStat>
  <sumStat type="medn">0.8226</sumStat>
  <sumStat type="mode">0.74171</sumStat>
  <sumStat type="vald">1010</sumStat>
  <sumStat type="invd">0</sumStat>
  <sumStat type="min">0.216</sumStat>
  <sumStat type="max">5.067</sumStat>
  <sumStat type="stdev">0.646043709881769</sumStat>
  <varFormat type="numeric" />
  <notes subject="Universal Numeric Fingerprint" level="variable"
    source="archive"
    type="VDC:UNF">UNF:3:10:hkVPP/vHhnDYuJPmsQRSqA==</notes>
  </var>
- <var ID="v1.3" name="REGION" intrvl="discrete">
  <location fileid="file1" />
  <labl level="variable">Region.</labl>
- <invalrng>
  <item VALUE="-99.99" />
  </invalrng>
- <catgry missing="Y">
  <catValu>-99.99</catValu>
  <labl level="CATEGORY">NA</labl>
  <catStat>0</catStat>
  </catgry>
- <catgry>
  <catValu>-9</catValu>
  <labl level="CATEGORY">Don't know</labl>
  <catStat>0</catStat>
  </catgry>
- <catgry>
  <catValu>-8</catValu>
  <labl level="CATEGORY">Refused</labl>
  <catStat>0</catStat>
  </catgry>
```

```
- <catgry>
  <catValu>1</catValu>
  <labl level="CATEGORY">East 1 (CT, ME, MA,</labl>
  <catStat>54</catStat>
    </catgry>
- <catgry>
  <catValu>2</catValu>
  <labl level="CATEGORY">East 2 (MD, NJ, NY,</labl>
  <catStat>182</catStat>
    </catgry>
- <catgry>
  <catValu>3</catValu>
  <labl level="CATEGORY">South 3 (AL, FL, GA,</labl>
  <catStat>223</catStat>
    </catgry>
- <catgry>
  <catValu>4</catValu>
  <labl level="CATEGORY">South 4 (AR, LA, OK,</labl>
  <catStat>97</catStat>
    </catgry>
- <catgry>
  <catValu>5</catValu>
  <labl level="CATEGORY">Midwest 5 (IL, IN, M</labl>
  <catStat>170</catStat>
    </catgry>
- <catgry>
  <catValu>6</catValu>
  <labl level="CATEGORY">Midwest 6 (IA, KS, M</labl>
  <catStat>76</catStat>
    </catgry>
- <catgry>
  <catValu>7</catValu>
  <labl level="CATEGORY">West 7 (AZ, CO, ID,</labl>
  <catStat>61</catStat>
    </catgry>
- <catgry>
  <catValu>8</catValu>
  <labl level="CATEGORY">West 8 (CA, OR, WA,</labl>
  <catStat>147</catStat>
    </catgry>
  <varFormat type="numeric" />
  <notes subject="Universal Numeric Fingerprint" level="variable"
    source="archive"
    type="VDC:UNF">UNF:3:10:n+4EVZGwvYtS2UX/+CYbZg==</notes>
    </var>
    </dataDscr>
- <otherMat URI="http://vdc-
    demo.hmdc.harvard.edu/VDC/Repository/0.1/Access/hdl:1902.0/H-
    11767/harris_s11767_quest.pdf" level="study" type="other">
```

```
  <labl>harris_s11767_quest.pdf</labl>
  <txt />
  <notes source="producer" subject="description"
     type="icpsr:category">Codebook, PDF File</notes>
    </otherMat>
<otherMat URI="http://vdc-
     demo.hmdc.harvard.edu/VDC/Repository/0.1/Access/hdl:1902.0/H-
     11767/agreement.pdf" level="study" type="other">
  <labl>Deposit_Agreement.pdf</labl>
  <txt />
  <notes source="producer" subject="DATA-PASS Deposit Agreement Text"
     type="DATAPASS:TERMS:STANDARD:1.0:FILE" level="file " >Deposit
     Agreement Text</notes>
    </otherMat>

    </codeBook>
```