

**PHASE II REPORT OF THE WORKING GROUP ON PERMANENCE  
OF NLM ELECTRONIC PUBLICATIONS  
Revised October 13, 2000**

**Introduction**

In its Phase I Report of December 15, 1999, the NLM Working Group on Permanence (WGP) identified three core categories of permanence (identifier validity, resource availability, content invariance) and proposed a system for rating NLM's electronic resources (Appendix A). During Phase II, the WGP focused on the following implementation issues:

**Condensed Ratings**

The WGP looked at the feasibility of using a condensed form of the proposed rating system that would be simpler for resource creators to apply and easier for users to understand. The subgroup developed the following sample ratings:

Permanent: Unchanging Content

Example: Image of correspondence in the Profiles in Science collection

Permanent: Stable Content

Example: a MEDLINE record

Permanent: Dynamic Content

Example: NLM Home Page

Permanence Not Guaranteed

Examples: Conference calendars, preliminary agendas

A rating of "Permanent" means that identifier validity and resource availability are guaranteed. A rating of "Permanence Not Guaranteed" means that identifier validity and/or resource availability could change.

Aggregate resources (e.g., databases and digital library collections) that are rated "Permanent" may be assigned the following subratings as appropriate:

Growing: Additional objects may be added to the resource.

Closed: Objects are no longer being added to a resource that previously was subject to growth.

For resources rated “Permanence Not Guaranteed”, optional subratings may be used for to indicate which of the elements is subject to change: (i.e., the identifier is transient and/or the resource could disappear).

It is recognized that condensed ratings may not be applicable to all resources. In some cases, the more descriptive elements in the originally proposed rating system (Appendix A) would be more appropriate.

### **Testing the Permanence Rating System**

A subgroup of the WGP reviewed a list of FY 1999 Top 100 [NLM] Web Files and selected a group of ten files that included a variety of resource types. Members of the subgroup independently assigned permanence ratings and then collectively reviewed their ratings to determine how much agreement existed and assess how easily the rating system might be used. After discussion, consensus was reached concerning most of the ratings. The exercise pointed out the need for written guidelines for assigning ratings if consistency across the Library is to be achieved.

The subgroup also looked at the NLM Web site and the Web sites of several major NLM program areas (e.g., NCBI, LHC, NNLM, SIS) and attempted to assign permanence ratings and resource categories. This exercise led to an expanded version of the Phase I resource category list. It also led to the recognition that ratings must be assigned by staff from the program areas in which the resources were created since they are most familiar with the content and best qualified to judge its long term importance to the Library.

Subgroup members rated the expanded list of resources categories to determine whether default ratings could be assigned. It was determined that many resources would not need to be rated individually. Program areas could use ratings that had been pre-assigned on the basis of resource category (see Appendix B). A template could be used that contains the list of resource categories and their default ratings. When needed, program areas could override pre-assigned ratings.

The heterogeneous nature of some resource categories (e.g., reports) made them not amenable to default ratings. In these cases, program areas would need to assign ratings to resources individually. For some resources rated “Permanence Not Guaranteed” (e.g., fact sheets, visitor information), only a selection needs to be retained to document minor changes in NLM programs or services over time. These categories could be flagged so that snapshots could be taken at regular intervals.

The subgroup currently is drafting guidelines for assigning permanence ratings. The guidelines will discuss the implications of assigning specific ratings, list considerations for selecting electronic resources to be retained for the NLM archives, and explain the rationale for the default ratings that have been assigned to resource categories.

The subgroup's work was an attempt to determine whether the proposed condensed ratings, list of resource categories, and default ratings would be workable. Once the system has been implemented, refinements and additions may be needed.

### **Coordination of the Permanence Rating System**

A subgroup of the Working Group on Permanence (representing the NLM Archives and all major programs areas that produce NLM Web resources) should be charged with ensuring that the rating system is implemented and used consistently throughout the Library. Members of the subgroup would serve as permanence coordinators for their respective areas, oversee the assignment of ratings, handle questions from staff, and conduct periodic reviews of ratings to ensure that the rating system is well understood and working as intended. Permanence coordinators would meet as needed to discuss and refine permanence guidelines and policies and make recommendations for changes in the rating system. Because of the significant long-term commitment involved, the group would also review ratings of "Permanent" for major resources such as large digital collections and refer them as needed to NLM higher management for approval.

If a permanence coordinator cannot review the permanence ratings of a new resource in a timely manner, the resource may be mounted on the Web with a provisional rating. A system could be developed whereby an email message to permanence coordinators could be generated automatically whenever a Web Document Management team promotes to the Web a resource which has been assigned a provisional rating. Authorized ratings should be marked in some manner to distinguish them from provisional ratings.

### **Metadata Elements**

NLM's Cataloging Section, working with staff from other Divisions, has developed a proposed set of metadata elements for NLM wide application. The proposed elements will include permanence metadata. When a resource has been assigned a rating of "Permanent", its metadata would be downloaded to NLM's Voyager system and upgraded to meet cataloging requirements. The permanence rating would be displayed to users in NLM's LocatorPlus and distributed to the bibliographic utilities as part of the record. Because users may access a resource through the Web rather than through Voyager, the natural language (i.e., condensed) rating should appear with the Web version. Mechanisms will need to be developed to keep Voyager bibliographic data and the metadata for electronic resources synchronized. Examples of permanence metadata and a corresponding Voyager record are given in Appendix C.

It should be noted that a considerable amount of work on preservation metadata already has been done by:

National Library of Australia ([www.nla.gov.au/preserve/pmeta.html](http://www.nla.gov.au/preserve/pmeta.html))

CEDARS Project in the U.K. ([www.leeds.ac.uk/cedars/MD-STR~5.pdf](http://www.leeds.ac.uk/cedars/MD-STR~5.pdf))

Research Libraries Group ([www.rlg.org/preserv/presmeta.html](http://www.rlg.org/preserv/presmeta.html)).

To our knowledge, only NLM has developed permanence ratings. We have begun to publicize our efforts and are participating in national and international efforts to develop metadata standards for electronic publications.

### **Resource Identifiers**

All NLM Web resources that have been rated “Permanent” should be assigned unique identifiers. A unique identifier (UI) should appear on or be attached to the rated resource. A Web-based system should be developed that, given the resource’s UI, would always link the user to that resource, regardless of where that resource is actually located. For a large resource such as MEDLINE, where it may not be practical to link directly to each item, the UI may point to an index page/application (i.e., PUBMED). Given the UI, the index application could automatically link the user to the item. A mechanism will be needed to prevent the assignment of the same UI to resources in different collections.

### **Dynamically Rendered Resources**

A particular challenge in defining permanence for Web pages comes from the increasing practice of generating information views dynamically. Pages are set up to be produced not from one stored resource but from a combination of resources and database fields that are aggregated by software just prior to leaving the server on their way to the user’s browser. The delivered resource may differ, sometimes significantly, from the stored resources from which it was created. The decision whether to rate a dynamically rendered resource will be left to the discretion of the program area responsible for creating it. All stored resources that have been produced by NLM and used to create composite resources will be assigned individual permanence ratings. The unique identifiers for each stored resource will be included in the metadata for the composite resource. This will allow the user to determine the permanence ratings of each of the stored resources and thereby assemble a more complete picture of the permanence of a composite resource than that provided by the rating for the composite resource alone.

### **Management of Servers**

Maintaining all of NLM’s public electronic resources on the same server is not practical. Resources with high and low permanence ratings could reside on the same server. However, if high and low priority resources are kept on separate servers, different rules could be applied to their management. Any server that contains resources with permanently available ratings should be maintained at the level required for the highest level of permanence.

It is important that common practices govern the management of NLM servers on which resources that have been assigned permanence ratings reside. Guidelines should be developed for management of all NLM servers that store resources that have been assigned ratings of Permanent.

Sample guidelines include:

1. The server is backed up regularly on to reliable media. At least one complete backup copy, which is less than three months old, must be stored off-site in a secure location.

2. Resources stored on the server can only be modified by following an established resource management process.
3. The server must be configured and administered in accordance with current best security practices to ensure content integrity. The server must have a designated administrator who will keep the server at current security patch levels. Access to the server will be highly restricted (e.g., the server will not maintain general user accounts or insecure applications).
4. All resources that have a guaranteed Identifier Validity must be registered either individually or collectively in the NLM UI database. Any action that would change a document's URL must be coordinated with and immediately reflected in the UI database.

The responsibility to ensure the long term availability of resources that have been rated permanent belongs to NLM rather than to the program area that produced the resource. Some strategies for ensuring long term availability include: 1) establishing a central repository for resources that are of diminished interest to their producers; or 2) implementing an NLM policy that resources residing on any NLM server that have been assigned ratings of "Permanent" will not be deleted without higher level review.

### **Recommended Phase III Activities:**

#### **A. Systems Related Work**

1. NLM-wide specifications for the format and location of permanence ratings should be developed. Tasks include:
  - a. defining the format for specifying/displaying a resource's rating (i.e., the text string used to describe the rating);
  - b. defining a standard practice for where the rating will be displayed. This standard practice should encompass all known electronic resource types, including: HTML and/or XML resources, text formatted resources (e.g., ASCII, Word, WordPerfect), still images, video, and databases (e.g., MEDLINE, PubMed, HSTAT). The standard practice will describe how to embed the rating into text items and/or where to place the ratings for other individual and aggregate resources.

The two specifications (a and b above) should be detailed enough to allow a simple Web spider (an automated Web indexing application) to visit each NLM Web resource, discover its permanence rating, and decode it sufficiently to trigger any actions that might be needed. The specifications

should indicate where a spider will find the rating (e.g., by looking in a specific metadata element embedded in an HTML or XML file) and enough about its structure to distinguish the separate components of the rating. The specifications will describe how to find the rating of an object that cannot contain an embedded rating (e.g., a graphics image). In this case, the rating may be associated with the link to that object. Finally, the specifications should indicate what should be done when a rating is not found.

2. NLM-wide specifications for the format and location of unique identifiers should be developed. Tasks include:
  - a. defining a UI format that will be compatible with existing NLM UIs and providing a means to identify uniquely every electronic resource that has been assigned a permanence rating;
  - b. defining a standard practice for where to display UIs. This practice will vary depending upon the nature of the resources (e.g., HTML and/or XML, text, image).

The two specifications (a and b above) should be detailed enough to allow a simple Web spider (an automated Web indexing application) to visit each NLM resource and discover its unique identifier. The specifications should indicate where a spider will find the identifier (e.g., in embedded metadata or in an adjunct file) and what should be done if no identifier is found.

3. Once standard practices for format and display of permanence ratings and UIs have been established, a set of applications should be developed that will:
  - a. assist in recording and maintaining permanence ratings for NLM digital resources;
  - b. link to an NLM electronic resource given the resource's UI.
4. A prototype system should be developed.
5. The WGP's suggested guidelines for management of the servers on which NLM's permanent electronic reside should be refined and expanded.

## **B. Program Coordination and Review:**

A subgroup of the Working Group on Permanence (representing the NLM Archives and all major program areas that produce NLM Web resources) should be assigned to serve as permanence coordinators. Their duties would include:

1. overseeing the assignment of permanence ratings to existing NLM electronic resources that were created in their respective program areas;
2. reviewing ratings for new resources to ensure that staff in their respective program areas understand the rating system and apply it consistently;
3. keeping staff in their respective program areas informed of changes in NLM permanence policy or procedures;
4. working with other Permanence Coordinators as needed to:
  - a) identify and resolve issues and problems related to permanence ratings;
  - b) recommend revised policies or guidelines for assigning permanence ratings to NLM Web resources;
  - c) prepare reports for NLM management on how well the permanence rating system is working and whether changes are needed.

This report was prepared by the NLM Working Group on Permanence:

Duane Arenales, Technical Services Division, LO  
Margaret Byrnes, Public Services Division, LO (Chair)  
Judith Eannarino, Technical Services Division, LO  
Jeanne Goshorn, Biomedical Services Branch, SIS  
Christa Hoffmann, Technical Services Division, LO  
John A. Kunze, IPA, Consultant  
Jennifer Marill, Public Services Division, LO  
Liz Pope, Information Engineering Branch, NCBI  
Maureen Prettyman, Computer Sciences Branch, LHNCBC  
Mark Silverman, Systems Services Branch, OCCS  
Paul Theerman, History of Medicine Division, LO

The group would like to acknowledge the contributions made by Cynthia Love, Public Services Division, LO, who assisted with testing the rating system and revising the list of resource categories and contributed to the final deliberations of the WGP's Phase II work.

## **Appendix A.**

### **Proposed Permanence Rating System from the December 15, 1999 Phase I Report of the NLM Working Group on Permanence (Revised 10-13-200)**

The Working Group on Permanence identified three core categories of permanence for electronic resources: identifier validity, resource availability, and content invariance. A fourth category, link stability, has been incorporated into content invariance.

The permanence rating system proposed for NLM is as follows:

#### **IV: Identifier Validity**

1. Transient
2. Guaranteed

#### **RA: Resource Availability**

1. No guarantee
2. Permanently available

#### **CI: Content Invariance**

1. Dynamic
  - a. Growing
  - b. Closed
2. Stable
  - a. Growing
  - b. Closed
3. Unchanging

## Explanation of the Ratings:

**IV: Identifier Validity:** the extent to which a user can be assured that a given name or other identifier will always retrieve the same resource

**1. Transient:** the identifier is subject to change at any time (e.g., a preliminary meeting agenda)

**2. Guaranteed:** identifier will always retrieve the same resource (e.g., a MEDLINE record, the NLM Home Page)

**RA: Resource Availability:** the extent to which a given resource is guaranteed to remain available in electronic form

**1. No Guarantee:** resource may become unavailable at any time (e.g., announcements)

**2. Permanently Available:** guaranteed to remain available. This rating implies NLM's commitment to archive the resource (e.g., minutes of NLM Board of Regents meetings)

**CI. Content Invariance:** the extent to which the content of a given resource and the links it contains are guaranteed to remain unchanged

**1. Dynamic:** Resource is subject to replacement, correction, revision, migration, recapture. Internal and external links could change. This includes changing the version of a resource to which a link points (e.g., MedlinePlus). While a rating of dynamic indicates that a resource could be replaced or updated, it is not a commitment to do so. The variables Growing and Closed are optional variables used to indicate whether a dynamic resource will grow in regular increments or is no longer growing. A dynamic rating without the growing/closed variable not subject to incremental growth.

**2. Stable:** Resource is subject only to correction, minor additions, migration, recapture. It cannot be replaced or revised. Internal links will be updated. External links should be updated to ensure that they point to the original content of an external resource. The variables Growing and Closed are optional variables used to indicate whether a stable resource will grow in regular increments or is no longer growing. A stable rating without the growing/closed variable means that the resource is not subject to incremental growth.

**3. Unchanging:** Content is static (e.g., individual slides from the Visible Human). Resource is subject only to migration. Internal links will be updated as part of migration.

## Appendix B

### CATEGORIES OF NLM ELECTRONIC RESOURCES WITH PRELIMINARY DEFAULT RATINGS

<b><u>Resource Category</u></b>	<b><u>Proposed Default Rating</u></b>
Announcements/News	Permanence Not Guaranteed
Applications/forms/registrations	Permanence Not Guaranteed
Bibliographies	[No Default Rating]
Calendars/schedules	Permanence Not Guaranteed
Clinical Alerts	Permanent: Unchanging Content
Databases	[No Default Rating]
Database records	Permanent: Stable Content
Digital Images in Digital Library Collections	Permanent: Unchanging Content
Digital library collections	[No Default Rating]
Exhibitions	Permanent: Stable Content
Fact Sheets	Permanence Not Guaranteed
FAQs, help files, pocket cards	Permanence Not Guaranteed
Grants and awards	[No Default Rating]
Home pages	Permanence Not Guaranteed
Lists of links	Permanence Not Guaranteed
Minutes – officially released by NLM	Permanent: Unchanging Content
Newsletters	Permanent: Stable Content
Organizational charts, organizational directories, staff directories	Permanence Not Guaranteed
Photos of staff, programs, activities, buildings & grounds	[No Default Rating]
Policies – officially released by NLM	Permanent: Stable Content
Press releases	Permanent: Stable Content

Procedures	Permanence Not Guaranteed
Product, program, and project descriptions	[No Default Rating]
Reports– officially released by NLM	Permanent: Stable Content
Requests for Proposals, Requests for Applications, Program Announcements, Requests for Quotations, Statements of Work or their summaries	Permanence Not Guaranteed
Search pages	Permanence Not Guaranteed
Software	Permanence Not Guaranteed
Staff biographical sketches	[No Default Rating]
Staff/contractor/consultant papers, presentations, slide shows [not including course or training materials]	[No Default Rating]
Training materials/manuals	Permanence Not Guaranteed
Visitor Information	Permanence Not Guaranteed

**Revised 10-3-00**

## Appendix C

### Example of Encoded Metadata Record

```
<html>
<head>
<title>      Breath of Life </title>
<link rel =  "schema.DC"
  href =     "http://purl.org/DC/elements/1.0/">
<meta name = "DC.Title"
  content =  "Breath of Life">
<meta name = "DC.Creator"
  content =  "National Library of Medicine"(U.S.)>
<meta name = "DC.Date"
  content =  "1999-03">
<meta name = "DC.Type"
  content =  "Exhibitions">
<meta name = "DC.Format"
  content =  "HTML/Java script">
<meta name = "DC.Identifier"
  content =  URL: http://www.nlm.nih.gov/hmd/breath/breathhome.html>
<meta name = "DC.Identifier"
  content =  "NLM 906148">
<meta name = "PermanenceRating"
  content =  "NLM: IV=2, RA=1, CI=2b">
<meta name = "CondensedPermanenceRating"
  content =  "Permanent:Stable Content /Closed">
<meta name = "PermanenceRatingStatus"
  content =  "NLM Approved">
<meta name = "MetadataCreator"
  content =  "National Library of Medicine (U.S.)">
<meta name = "PermanenceGuarantor"
  content =  "National Library of Medicine (U.S.)">
```

## Appendix C

### Example of a brief metadata record--Required fields only

**Title:** Breath of Life

**Subject. Keyword:** Asthma

**Publisher:** National Library of Medicine

**Date.Issued:** 1999-03

**Identifier.URL:** <http://www.nlm.nih.gov/hmd/breath/breathhome.html>

**Identifier.NLMID:** 906148

**Permanence.Level:** **Permanent. Stable content/ Closed**

**Permanence.Status:** **Provisional**

**Permanence.Guarantor:** **National Library of Medicine (U.S.)**

**Contact.E-mail:** custserv@nlm.nih.gov

**Contact.Affiliation:** History of Medicine Division

**Language:** EN

**Type:** Exhibitions

**Rights:** Public domain

## Appendix C

### Example of an enhanced metadata record

**Title:** Breath of life

**Subject.Keyword:** Asthma

**Subject.MeSH:** Asthma -- history

**Subject.MeSH:** Asthma -- prevention & control

**Subject.MeSH:** Famous Persons

**Subject.NLMClass:** WF 553

**Publisher:** National Library of Medicine

**Date.Created:** 1999-03

**Identifier.URL:** <http://www.nlm.nih.gov/hmd/breath/breathhome.html>

**Identifier.NLMID:** 906148

**Permanence.Level:** **Permanent: Stable content/ Closed**

**Permanence.Status:** **Approved 2001-03-15**

**Permanence.Guarantor:** **National Library of Medicine (U.S.)**

**Language:** EN

**Type:** Exhibitions  
Personal Narratives

**Rights:** Public domain

**Creator.CorporateName** National Library of Medicine (U.S.)

**Creator.CorporateName** National Heart, Lung, and Blood Institute

**Creator.CorporateName** National Institute of Allergy and Infectious Diseases  
(U.S.)

**Creator.CorporateName** National Institute of Environmental Health Sciences

**Contact.E-mail** custserv@nlm.nih.gov

**Contact.Section:** History of Medicine Division

**Description.Contents:** Exhibition -- The faces of asthma -- Research -- Resources -- Interactives -- Acknowledgements.

**Description.Abstract:** Examines the history of asthma, the experiences of people with asthma, and contemporary efforts to control the disease.

**Description.Audience:** General

**Relation:** Describes the exhibition held at the National Library of Medicine, Mar. 23, 1999-June 30, 2000.

**Format:** HTML and Java script

**Appendix C**  
**Example of an OPAC Record--Details View**

**Author(s):** National Library of Medicine (U.S.)  
National Heart, Lung, and Blood Institute  
National Institute of Allergy and Infectious Diseases (U.S.)  
National Institute of Environmental Health Sciences

**Title:** Breath of life [computer file] / created by the National Library of Medicine in collaboration with the National Heart, Lung, and Blood Institute, the National Institute of Allergy and Infectious Diseases, and the National Institute of Environmental Health Sciences.

**Publisher:** Bethesda, Md. : National Library of Medicine, 1999.

**Electronic Links:** <http://www.nlm.nih.gov/hmd/breath/breathhome.html>

**MeSH Subjects:** Asthma--history  
Asthma--prevention & control  
Famous Persons  
Exhibitions  
Personal Narratives

**Contents:** Exhibition -- The faces of asthma -- Research -- Resources -- Interactives -- Acknowledgements.

**Notes:** **Permanent: Stable content/Closed. Guarantor: National Library of Medicine.**  
Describes the exhibition held at the National Library of Medicine, Mar. 23, 1999-June 30, 2000.  
Public domain.  
Abstract: Examines the history of asthma, the experiences of people with asthma, and contemporary efforts to control the disease.  
Audience: General.  
HTML and Java script.

**NLM ID:** 906148

