# US Department of Health and Human Services

## Privacy Impact Assessment

### Date Signed:

10/03/2017

### OPDIV:

NIH

### Name:

NLM Biomedical Terminology System

**PIA Unique Identifier:**

P-9082031-584336

**The subject of this PIA is which of the following?**

Major Application

**Identify the Enterprise Performance Lifecycle Phase of the system.**

Operations and Maintenance

**Is this a FISMA-Reportable system?**

Yes

**Does the system include a Website or online application available to and for the use of the general public?**

Yes

**Identify the operator.**

Agency

**Is this a new or existing system?**

New

**Does the system have Security Authorization (SA)?**

No

**Indicate the following reason(s) for updating this PIA.**

**Describe the purpose of the system.**

The National Library of Medicine (NLM) Biomedical Terminology System (NBTS) is a multi-purpose heterogeneous system consisting of and providing access to health and biomedical vocabularies and standards.

The purpose is to build and maintain vocabularies and standards.

NLM NBTS consists of a controlled vocabulary thesaurus, value sets (purpose-specific subsets of standard terminologies), which are used in the specification of quality measures for healthcare, Health Level 7 (HL7) messaging guides, patient reported outcomes measures, and software tools.

NBTS applications include the following and each maintain their own privacy assessment:

Unified Medical Language System (UMLS)
MeSH
RxNorm Editing System
Value Set Authority Center (VSAC)

International Health Terminology Standards Development Organization (IHTSDO)
U.S. SNOMED CT Content Request System (USCRS)

**Describe the type of information the system will collect, maintain (store), or share.**

The NLM NBTS system predominantly collects, processes, curates, and represents medical terminologies, and the co-involvement of these terminologies. However, access to the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) medical terminology is controlled under license agreement between the system users and the publishers of terminologies. Users must complete a license registration in order to access these controlled vocabularies. To obtain a license, users are asked to read about their responsibilities under the agreements, and provide some basic information about themselves or their organizations. Licenses are issued to individuals, but license holders can represent an organization-wide use of the licensed data. Individuals are asked to provide first and last name, title, phone number, email address, mailing address, organization name and organizational category.

The publishers of terminologies are a list of roughly 120 agencies, an international organization which are outlined in Appendix 1 of the UMLS Metathesaurus License agreement.

The NBTS applications require system administrators, (who are either federal government employees, or direct contractors), to login to perform maintenance on the system. There are different user access levels for different roles which dictate how they log in to the system. There are details for these roles and access levels in the System Management section of the Disaster Recovery and Business Continuity Plan under the respective application name.

NLM staff, including system administrators and direct contractors, access NBTS using NIH Active Directory, which maintains its own unique privacy impact assessment (PIA). The purpose of NIH Active Directory is to authenticate and authorize all users and computers in a Windows domain type network; assigning and enforcing information security policies for all computers and installing or updating software. NIH Active Directory collects unique user names and passwords (user credentials) and stores them in an encrypted format. NIH Active Directory is an essential service which facilitates and governs network access to various resources. There is no opt-out for system administrators. If system administrators don't want to enter their credentials then they aren't able to access the system to perform their duties.

**Provide an overview of the system and describe the information it will collect, maintain (store), or share, either permanently or temporarily.**

The National Library of Medicine (NLM) Biomedical Terminology System (NBTS) leverages the proprietary products of many medical vocabulary data providers to collect and assemble resources and make these available collectively, and in relationship, to one another. The purpose is to build and maintain vocabularies and standards.

The licensing module collects and retains basic contact and location information provided by licensors about themselves. This registration information is provided voluntarily.

NLM NBTS is a multi-purpose heterogeneous system consisting of and providing access to health and biomedical vocabularies and standards. It consists of a controlled vocabulary thesaurus, value sets (purpose-specific subsets of standard terminologies), which are used in the specification of quality measures for healthcare, Health Level 7 (HL7) messaging guides, patient reported outcomes measures, etc., and software tools.

NBTS applications include:

Unified Medical Language System (UMLS)
MeSH

RxNorm Editing System
Value Set Authority Center (VSAC)
International Health Terminology Standards Development Organization (IHTSDO)
U.S. SNOMED CT Content Request System (USCRS)

The NBTS applications do not collect any information from system users, the applications are all open to the general public for viewing. System Administrators access and log in credentials are linked to their role, please see the Disaster Recovery and Business Continuity Plan for more details.

**Does the system collect, maintain, use or share PII?**

Yes

**Indicate the type of PII that the system will collect or maintain.**

Name

E-Mail Address

Mailing Address

Phone Numbers

Organization

User credentials

**Indicate the categories of individuals about whom PII is collected, maintained or shared.**

Employees

Public Citizens

Business Partner/Contacts (Federal/state/local agencies)

Vendor/Suppliers/Contractors

**How many individuals' PII is in the system?**

10,000-49,999

**For what primary purpose is the PII used?**

Personally Identifiable Information (PII) is used for contacting licensees, if needed, and to validate that their license use is appropriate. Regional and site restrictions apply to qualified use of some of this content.  Where content is being used beyond the terms of the license agreement; fees may be charged for use of the content, which NLM is not involved in.

Basic information about licensees is collected. Personally Identifiable Information (PII) is used for contacting licensees, if needed, and to validate that their license use is appropriate.

**Describe the secondary uses for which the PII will be used.**

Not Applicable

**Identify legal authorities governing information used and disclosure specific to the system and program.**

The NLM was established by statute in order to assist the advancement of medical and related sciences, and to aid the dissemination and exchange of scientific and other information important to the progress of medicine and to the public health, (section 465 of the Public Health Service Act, as amended (42 U.S.C. section 286).

**Are records on the system retrieved by one or more PII data elements?**

Yes

**Identify the number and title of the Privacy Act System of Records Notice (SORN) that is being use to cover the system or identify if a SORN is being developed.**

SORN is In Progress

**Identify the sources of PII in the system.**

**Directly from an individual about whom the information pertains**

In-Person

Online

**Government Sources**

Within OpDiv

Other Federal Entities

**Identify the OMB information collection approval number and expiration date**

Not applicable

**Is the PII shared with other organizations?**

Yes

**Identify with whom the PII is shared or disclosed and for what purpose.**

**Within HHS**

Over 20,000 individuals license Unified Medical Language System (UMLS) data. Examples of applications that have integrated UMLS data and services include:

National Guidelines Clearinghouse, an AHRQ initiative, is a public resource for evidence-based clinical practice guidelines.

National Quality Measures Clearinghouse, sponsored by AHRQ, is a public repository for evidence-based quality measures and measure sets

NCI Thesaurus and Enterprise Vocabulary Services includes reference terminology and related vocabulary resources to help researchers, clinicians, and administrators identify types of cancer, treatments, and other cancer related topics.

Anyone from an HHS Operating Division can request a list of UMLS (Unified Medical Language System) licensees. NLM also grants read-only administrative account to the International Health Terminology Standards Development Organization (IHTSDO), who are the publishers of SNOMED CT, a widely used clinical ontology. Because all UMLS licensees are automatically SNOMED CT affiliates per the license agreement, IHTDSO queries the licensee database from time to time to make sure that any UMLS users outside of IHTSDO member territories have registered in the IHTDSO licensing system. From their end, they're trying to protect their investment into the intellectual property of SNOMED CT, which NLM distributes through the UMLS, free of charge to U.S. users. Individuals requesting a UMLS License agree to the following term as stated within the UMLS License Agreement: "The names and addresses of licensees authorized to use the UMLS products are public information."

**Other Federal Agencies**

Over 20,000 individuals license Unified Medical Language System (UMLS) data. Examples of applications that have integrated UMLS data and services include:

National Guidelines Clearinghouse, an AHRQ initiative, is a public resource for evidence-based clinical practice guidelines.

National Quality Measures Clearinghouse, sponsored by AHRQ, is a public repository for evidence-based quality measures and measure sets

NCI Thesaurus and Enterprise Vocabulary Services includes reference terminology and related vocabulary resources to help researchers, clinicians, and administrators identify types of cancer, treatments, and other cancer related topics.

Anyone from another federal agency can request a list of UMLS (Unified Medical Language System) licensees. NLM also grants read-only administrative account to the International Health Terminology Standards Development Organization (IHTSDO), who are the publishers of SNOMED CT, a widely used clinical ontology. Because all UMLS licensees are automatically SNOMED CT affiliates per the license agreement, IHTDSO queries the licensee database from time to time to make sure that any UMLS users outside of IHTSDO member territories have registered in the IHTDSO licensing system. From their end, they're trying to protect their investment into the intellectual property of SNOMED CT, which NLM distributes through the UMLS, free of charge to U.S. users. Individuals requesting a UMLS License agree to the following term as stated within the UMLS License Agreement: "The names and addresses of licensees authorized to use the UMLS products are public information."

**State or Local Agencies**

Anyone from a state, local, or tribal organization can request a list of UMLS (Unified Medical Language System) licensees. NLM also grants read-only administrative account to the International Health Terminology Standards Development Organization (IHTSDO), who are the publishers of SNOMED CT, a widely used clinical ontology. Because all UMLS licensees are automatically SNOMED CT affiliates per the license agreement, IHTDSO queries the licensee database from time to time to make sure that any UMLS users outside of IHTSDO member territories have registered in the IHTDSO licensing system. From their end, they're trying to protect their investment into the intellectual property of SNOMED CT, which NLM distributes through the UMLS, free of charge to U.S. users.

Individuals requesting a UMLS License agree to the following term as stated within the UMLS License Agreement: "The names and addresses of licensees authorized to use the UMLS products are public information."

**Private Sector**

Anyone from the public can request a list of UMLS (Unified Medical Language System) licensees. NLM also grants read-only administrative account to the International Health Terminology Standards Development Organisation (IHTSDO), who are the publishers of SNOMED CT, a widely used clinical ontology. Because all UMLS licensees are automatically SNOMED CT affiliates per the license agreement, IHTDSO queries the licensee database from time to time to make sure that any UMLS users outside of IHTSDO member territories have registered in the IHTDSO licensing system. From their end, they're trying to protect their investment into the intellectual property of SNOMED CT, which NLM distributes through the UMLS, free of charge to U.S. users.

Individuals requesting a UMLS License agree to the following term as stated within the UMLS License Agreement: "The names and addresses of licensees authorized to use the UMLS products are public information."

**Describe any agreements in place that authorizes the information sharing or disclosure.**

UMLS Metathesaurus License (https://uts.nlm.nih.gov//license.html)

Unified Medical Language System (UMLS) is a set of files and software that NLM BTS uses to brings together many health and biomedical vocabularies and standards to enable interoperability between computer systems. UMLS can be used to enhance or develop applications, such as electronic health records, classification tools, dictionaries and language translators, link health information, medical terms, drug names, and billing codes across different computer systems, search engine retrieval, data mining, public health statistics reporting, and terminology research.

UMLS has three tools (knowledge sources):
Metathesaurus: Terms and codes from many vocabularies, including CPT®, ICD-10-CM, LOINC®, MeSH®, RxNorm, and SNOMED CT®
Semantic Network: Broad categories (semantic types) and their relationships (semantic relations)
SPECIALIST Lexicon and Lexical Tools: Natural language processing tools

UMLS Terminology Services (UTS) provides many options to access the UMLS:
Web Services APIs to query UMLS data within a unique application
Metathesaurus (web) Browser - Retrieve UMLS concept information, including CUIs, semantic types, and synonymous terms
Semantic Network (web) Browser - View the names, definitions, and hierarchical structure of the Semantic Network
Local Installation; files are downloaded through the UTS

**Describe the procedures for accounting for disclosures.**

Procedures for accounting for disclosures follow OMB Circular No. A-108, Federal Agency Responsibilities for Review, Reporting, and Publication under the Privacy Act.

Records are retained and disposed of under the authority of the NIH Records Control Schedule contained in NIH Manual Chapter 1743, Appendix 1 - "Keeping and Destroying Records" (HHS Records Management Manual, Appendix B-361), item 4000-A-2, which allows records to be destroyed when no longer needed for administrative purposes. Refer to the NIH Manual Chapter for specific disposition instructions.

**Describe the process in place to notify individuals that their personal information will be collected. If no prior notice is given, explain the reason.**

Individuals requesting a UMLS License through the NLM browser-based system must read through the UMLS License Agreement, fill out a license profile, and click a radio button confirming that they read and agree to the terms of the UMLS License Agreement and has provided accurate information in their license profile.

System administrators, which are direct contractors, provide credentials only to perform maintenance on the system. Log in access is through NIH Active Directory, which maintains its own privacy impact assessment (PIA).

**Is the submission of PII by individuals voluntary or mandatory?**

Voluntary

**Describe the method for individuals to opt-out of the collection or use of their PII. If there is no option to object to the information collection, provide a reason.**

Users can decline to register as a licensee.

The system administrators are given the opportunity to opt-in to the collection hence no need to provide an opt-out.

**Process to notify and obtain consent from individuals whose PII is in the system when major changes occur to the system.**

No changes have occurred, but given that the licensees have provided their contact information, the means are available for the NLM NBTS system owner to reach those license holders via email or database queries if the need arises.

System Administrators complete the annual Security Awareness and Privacy training, which includes the NIH Information Technology General Rules of Behavior (ROB) training requiring acceptance electronically to complete. The ROB governs our process of obtaining consent. A role-based training course relevant to their role is also required for all System Administrators with significant IT responsibility.

**Describe the process in place to resolve an individual's concerns when they believe their PII has been inappropriately obtained, used, or disclosed, or that the PII is inaccurate.**

All licensee-provided information is collected from the user. They can ask to have their license status dropped, and information purged.

NLM staff, including system administrators and direct contractors, access NBTS using NIH Active Directory, which maintains its own unique privacy impact assessment (PIA). The purpose of NIH Active Directory is to authenticate and authorize all users and computers in a Windows domain type network; assigning and enforcing information security policies for all computers and installing or updating software. NIH Active Directory collects unique user names and passwords (user credentials) and stores them in an encrypted format. NIH Active Directory is an essential service which facilitates and governs network access to various resources. There is no opt-out for system administrators. If system administrators don’t want to enter their credentials then they aren’t able to access the system to perform their duties.

**Describe the process in place for periodic reviews of PII contained in the system to ensure the data's integrity, availability, accuracy and relevancy.**

There are annual reviews of all licenses, where users must in fact opt back in to continue to use the NLM NBTS resources. Old records are purged if they are not updated during the license renewal cycle.

**Identify who will have access to the PII in the system and the reason why they require access.**

**Users:**

Anyone can request a list of UMLS (Unified Medical Language System) licensees within the NLM NBTS.

**Administrators:**

Only those with administrative permissions can log in and search/retrieve licensee records.

**Developers:**

Where needed to fulfill NLM NBTS business processes.

**Contractors:**

Where needed to fulfill NLM NBTS business processes. These are direct contractors.

**Others:**

Publishers of constituent NBTS medical terminologies.

**Describe the procedures in place to determine which system users (administrators, developers, contractors, etc.) may access PII.**

Licensee data is restricted for use in administering the system to those individuals who have a need to work with licensee data.

**Describe the methods in place to allow those with access to PII to only access the minimum amount of information necessary to perform their job.**

Oversight of the licensee data is tightly managed by NBTS Administrators. Controls are in place to provide least privileged to the system, on a need-to-know basis, which is tied to their roles.

**Identify training and awareness provided to personnel (system owners, managers, operators, contractors and/or program managers) using the system to make them aware of their responsibilities for protecting the information being collected and maintained.**

All users with access complete NIH mandatory security and privacy training on a yearly basis. In addition an ongoing orientation is provided to staff who work with the NBTS system.

**Describe training system users receive (above and beyond general security and privacy awareness training).**

System users receive role-based security training, Application Development Security training, and Secure Software Development Lifecycle training, as appropriate.

**Do contracts include Federal Acquisition Regulation and other appropriate clauses ensuring adherence to privacy provisions and practices?**

Yes

**Describe the process and guidelines in place with regard to the retention and destruction of PII.**

Individuals' records are retained as long as they are active users of the system. An annual review of records ensures that irrelevant and inactive records are purged from the system.

General Records Schedule (GRS) 3.2. Item 010, Disposition Authority: DAA-GRS-2013-0006-0001. Destroy 1 year(s) after system is superseded by a new iteration or when no longer needed for agency/IT administrative purposes to ensure a continuity of security controls throughout the life of the system.

**Describe, briefly but with specificity, how the PII will be secured in the system using administrative, technical, and physical controls.**

The UMLS licensee database is a website that is accessed over a secure internet connection, as well as through role-based permissions in the system. Only those with administrative permissions can log in and search/retrieve licensee records.

Physical controls include 24x7 guards of mobile units used to collect data, Personal Identify Verification (PIV), key cards and closed circuit TV.

Technical controls include User identification (ID), passwords, network firewall, Virtual Private Network (VPN), Intrusion Detection System, Role Based Access Controls, System logs.

Administrative controls include system security and contingency plan. Files are backed up regularly and stored off-site. Contract clauses ensure adherence to privacy provisions and practices, least privilege through role-based access, and policies for retention and destruction of PII.

**Identify the publicly-available URL:**

https://uts.nlm.nih.gov/home.html

Note: web address is a hyperlink.

**Does the website have a posted privacy notice?**

Yes

**Is the privacy policy available in a machine-readable format?**

Yes

**Does the website use web measurement and customization technology?**

Yes

**Select the type of website measurement and customization technologies is in use and if it is used to collect PII.**

Session Cookies that collect PII.

Persistent Cookies that collect PII.

Other technologies that do not collect PII:

Not applicable

**Does the website have any information or pages directed at children under the age of thirteen?**

No

**Does the website contain links to non- federal government websites external to HHS?**

No

**Is a disclaimer notice provided to users that follow external links to websites not owned or operated by HHS?**

null