

Using the UMLS Knowledge Source Server (UMLSKS) via the Internet

[7]

The UMLS Knowledge Source Server (UMLSKS) is a computer application that provides Internet access to the information stored in the UMLS Knowledge Sources. The purpose of the Knowledge Source Server is to make the UMLS data more accessible to users, and in particular to systems developers. The system architecture is based on the client server model, allowing remote site users (individuals as well as computer programs) to send requests to a centrally managed server at the U.S. National Library of Medicine. Access to the system is provided through a command line interface, through the World Wide Web, an Extensible Markup Language (XML)-based socket programming interface, and through an Application Programmer Interface (API).

Users are encouraged to consult the UMLSKS Web site for the most current UMLSKS documentation, including the User's Guide, Developer's Guide, and information on downloading the UMLS release files.

7.1 Downloading UMLS Knowledge Sources

UMLS licensees may access the UMLSKS and create an account with a LOGINID and password of their choosing. They then use that LOGINID and password for subsequent accesses. Licensees can download the current UMLS Knowledge Sources from the UMLS Web site. Archives of UMLS releases are kept and made available for several previous years on the Web site. For detailed technical specifications and installation instructions refer to the README.TXT file available on the Knowledge Sources downloads page.

7.2 System Architecture

The UMLSKS, made available in March 2002, was a redesign of the original "C" programming language system with new features added, including access to the UMLS Knowledge Sources through a public Web interface, incorporation of XML support for programmers both in requesting and returning data, inclusion of a Java-based Object Model of the UMLS Metathesaurus data, and incorporation of a TCP socket-based interface for non-Java programs. Subsequent releases of the software have augmented the available API functions and refined system operations. The system was designed with following design tenets in mind:

- Extensibility for ease of new feature incorporation
- Scalability in handling ever increasing user loads and increasing numbers of UMLS vocabularies
- Performance considerations permitting faster access to UMLS data
- Flexibility in access modes including a rich API set with access to all of the UMLS data
- Ease of administration by NLM staff and contractors
- Limited system interruptions during system software upgrades

7.3 Querying the Knowledge Source Server

7.3.1 Metathesaurus

The UMLS[®]SKS allows the user to request information about particular Metathesaurus concepts, including attributes such as the concept's definition, its semantic types, the concepts that are related to it, etc. It also allows the user to request information about the attributes themselves; for example, by asking for all the concepts that have been assigned to a particular semantic type.

Basic concept information includes the Metathesaurus unique identifier of the concept, the preferred name for the concept, and the names and sources of all terms that comprise that concept. Additional concept information often includes a definition and the source of that definition. Semantic type information is also included. Information about the hierarchical contexts of Metathesaurus concepts is readily available in the system. Related concepts are easily found. If a user were interested in information about a particular term within a concept, then the results could be limited in that way.

An important perspective on the Metathesaurus is source specific data. It is possible to query the server by limiting the query to a particular vocabulary. The user may wish to see the ancestors or descendants for a term in just a particular vocabulary, or the user may wish to see just the synonyms for a particular term in a particular vocabulary.

Attributes may be queried in the system. Thus, all concepts with a particular semantic type, all terms with a particular syntactic category, and all terms from a particular source vocabulary may be found. Searching for all concepts with a particular semantic type will, for example, give the user a good idea of the coverage of the Metathesaurus in a subject domain.

7.3.2 Semantic Network

The Semantic Network contains information about semantic types and their relationships. The implementation of the network module computes the relationships between semantic types using the inheritance property of the network type hierarchy. Information in the Semantic Network can be queried for semantic types and the relationships between them. Individual queries are specified by providing the known types or relations and leaving out the unknowns. The system then retrieves the corresponding values for the unknowns. For example, if the user wished to know what semantic types are related by a particular relation, then the user would indicate only the relationship name and all the semantic type pairs linked by that relationship would be retrieved. The user might also wish to know if a particular relationship holds between a pair of types.

It is possible to retrieve all the relations between a pair of types. For example, "treats", "prevents", and "complicates" would be listed, among others, as potential relationships between drugs and diseases. It is also possible to retrieve an exhaustive list of all related types in the network. Queries can be made about the definition, unique identifier, tree number, ancestors, parents, children, descendants, and siblings of a semantic type or relation.

7.3.3 SPECIALIST Lexicon

The UMLS[®]SKS provides access to lexical records in the SPECIALIST Lexicon. The SPECIALIST Lexicon is an English language lexicon containing many biomedical terms. The lexicon entry for each word or term records syntactic, morphological, and orthographic information. Lexical entries may be single or multiword terms. Lexical information includes syntactic category, inflectional variation (e.g., singular and plural for nouns, the conjugations of verbs, the positive, comparative, and superlative for adjectives and adverbs), and allowable

complementation patterns (i.e., the objects and other arguments that verbs, nouns, and adjectives can take).

7.4 Gaining Access to the UMLS Knowledge Source Server

Access to the UMLSKS is available to anyone who has signed the UMLS License Agreement and received a license code from NLM. First time users should establish a LOGINID and a password through the online registration at the Web site. Any questions or problems should be addressed via e-mail to NLM Customer Support.

7.5 UMLS Knowledge Source Server Documentation

UMLSKS users should always consult the documentation on the UMLSKS Web site for the most current information.

The following are publicly available on the site, under Documentation:

- Overview.
- FAQ.
- Developer's API Guide -- documentation generated using the javadoc facility that includes the object model, interfaces and some examples.
- A link back to the UMLS Reference Manual.
- A link to RxNorm documentation.