

Portal MeSHPol

XL Konferencja Problemowa Bibliotek Medycznych

11-12 września 2023 Kraków
Monika Fiedorowicz

sage

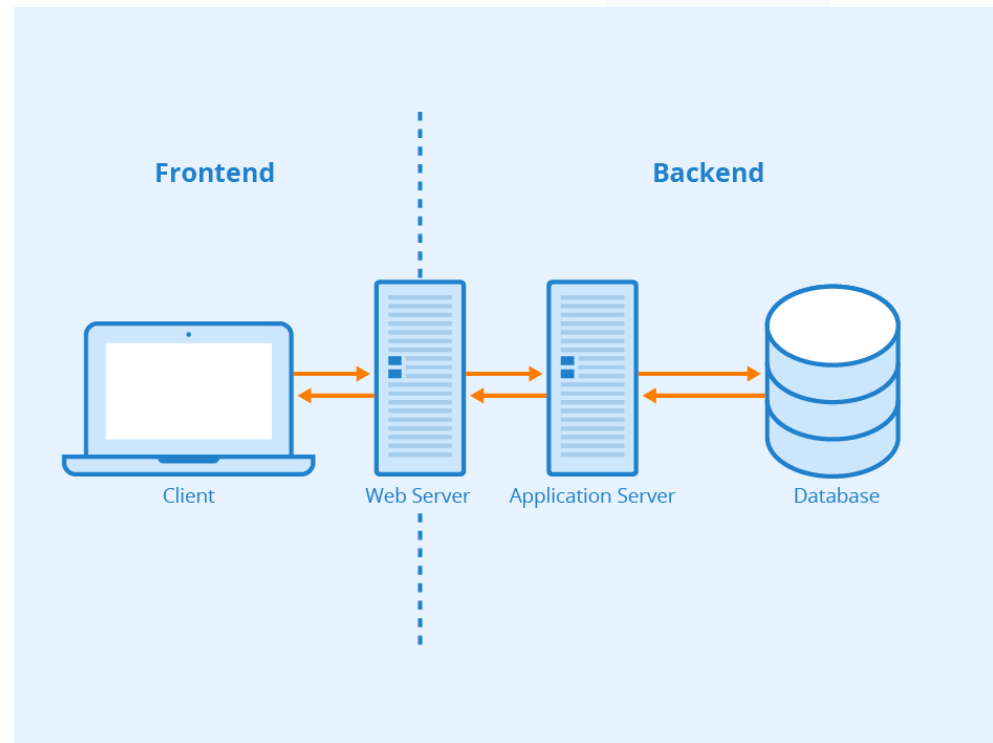
Agenda

1. Prezentacja techniczne/
administracyjne portalu
MeSHPol
2. Pobieranie plików do
katalogów
3. Integracja z Omegą PSiR
4. Proces aktualizowania
charakterystyk treściowych w
omedze



Strona techniczna

- część backendowa, czyli serwerowa
- oraz frontendowa, czyli aplikacja webowa
 - podsystem zarządzania Kartoteką MeSHPol
 - serwis WWW



Wykorzystane technologie

Arango DB
baza danych
klasy NoSQL

**Micronaut
Framework +
Java 20**



MeSHPol

**Apache
HTTP Server**

**Gatsby.js +
React**

Baza danych



Kolekcje m. in

- **deskryptory** w jednej kolekcji deskryptory EN i PL
- **kwalfikatory** w jednej kolekcji kwalifikatory EN i PL
- **relacje w drzewie** - relacje między pojęciami,
- **translacje** - wszystkie nowe i zmodyfikowane rekordy (podczas wgrywania kartoteki rocznej)
 - nowe tłumaczenia wprowadzone w danej wersji rocznej

Dane źródłowe

- dane źródłowe National Library of Medicine (USA)
- wersję polska MeSH w postaci plików xml



Hasła z poprzednich wersji rocznych

MeSH Descriptor Data 2022

Płucoserce

Zobacz rekord w języku angielskim

2022

- 2015
- 2016
- 2017
- 2018
- 2019
- 2020
- 2021
- 2022

Deskryptor	Płucoserce
Kod języka	PL
Numer(y) Drzew(a)	E07.858.082.458
Unikatowe ID	D006355
Unikalny Identyfikator RDF	http://id.nlm.nih.gov/
Adnotacja	external: do not confu
Inne formy	Serce mechaniczne
Historyczny przypis	98(63)
Data utworzenia	Fri Jan 01 1999
Data aktualizacji	Mon Nov 08 1999
Data wprowadzenia	
Nukat Number	m 2006169000

[Pokaż RDF](#)

[Wyświetl rekord w systemie MeSH NLM](#)

Interfejs wymiany danych - API REST

- umożliwia programistyczny dostęp do zasobów
- może działać z wieloma odbiorcami jednocześnie
- pobieranie danych na podstawie zdefiniowanych warunków

```
Descriptors
GET /rest/descriptors
GET /rest/descriptors/deleted
GET /rest/descriptors/modified
GET /rest/descriptors/{key}
GET /rest/descriptors/{key}/years

Export
GET /rest/nukat/startExport
```


REST API

Grupy endpointów

- komunikacja pomiędzy backendem i frontendem
- integracja z omegą
- integracja z NUKAT
- generowanie plików MARC deskryptorów i haseł rozwiniętych
 - formaty (MARC, xml, json, txt)

API - REST

SEARCH GENERATE PDF

- POST /rest/bulk/next
- Management**
 - GET /rest/beans
 - GET /rest/health
 - GET /rest/health/{selector}
 - GET /rest/loggers
 - GET /rest/loggers/{name}
 - GET /rest/routes
 - POST /rest/loggers/{name}
 - POST /rest/refresh
- Marc**
 - GET /rest/marc/descriptors/all
 - GET /rest/marc/descriptors/chunk
 - GET /rest/marc/descriptors/chunk/{descriptorUI}
 - GET /rest/marc/descriptors/{descriptorUI}
 - GET /rest/marc/expanded/all
 - GET /rest/marc/expanded/{descriptorUI}
 - GET /rest/marc/expanded/{descriptorUI}/{qualifierUI}

MARC

GET /rest/marc/descriptors/all

GET /api/rest/marc/descriptors/all

REQUEST

QUERY-STRING PARAMETERS

year	int32	<input type="text"/>
format	string null	<input type="text"/>

API Server <https://mesh.cm-uj.krakow.pl>
Authentication Required (None Applied)

RESPONSE

200

getAllDescriptors 200 response

EXAMPLE SCHEMA

Panel administratora

- zarządzanie użytkownikami
- zarządzanie roczną wersją kartoteki Mesh
- moduł audytu i integracji
- modyfikowanie informacji na stronie głównej, aktualności, pomoc
- Wysyłka plików do NUKAT

Panel administratora

[Lista użytkowników](#)

Dodaj nowego użytkownika

Zamrażanie kartoteki MESH

Moduł audytu integracji

Edytor strony głównej

Edytor strony pomoc

Edytor strony aktualności

Integracja z Nukat

Integracja z NUKAT



Panel administratora

[Lista użytkowników](#)

[Dodaj nowego użytkownika](#)

[Zamrażanie kartoteki MESH](#)

[Moduł audytu integracji](#)

[Edytor strony głównej](#)

[Edytor strony pomoc](#)

[Edytor strony aktualności](#)

[Integracja z Nukat](#)

Integracja z Nukat

przesłane zostaną wszystkie rekordy zmodyfikowane od daty ostatniego importu

Data ostatniego importu

Brak importów



Potwierdź i rozpocznij

Proces wgrywania kartoteki rocznej

Zamrożenie kartoteki

- tylko "Admina" może zamrozić roczną wersję kartoteki MeSHPol.
- Po zamrożeniu nie ma możliwości edycji rekordów w danym roku

Wgrywanie kolejnego roku

- pliki pobrane ze strony MeSH nlm

Rozwiązywanie konfliktów

hasła znajdujące się w zaktualizowanych w danym roku hasła (Spis tłumaczeń → Skonfliktowane → Zaktualizowane)

Zamrażanie kartoteki MeSH

Zamrożenie
kartoteki



[Strona główna](#) [Szukaj](#) [Aktualności](#) [Spis tłumaczeń](#) [Wgrywanie kartoteki MeSH](#) [Panel administratora](#) [Analityka](#) [Pomoc](#)

mfiedorowicz

MF

Panel administratora

[Lista użytkowników](#)

[Dodaj nowego użytkownika](#)

[Zamrażanie kartoteki MESH](#)

[Moduł audytu integracji](#)

[Edytor strony głównej](#)

[Edytor strony pomoc](#)

[Edytor strony aktualności](#)

Aktualna wersja kartoteki MeSH: 2022

Zamróz kartotekę



Wgrywanie kolejnego roku

- plik *.zip (pliki pobrane ze strony oryginalnego tezaurusa MeSH)
 - replace2016.txt - rekordy w których nastąpiły zmiany
 - meshnew2016.txt - nowe rekordy
 - qual2016.xml - kwalifikatory
 - desc2016.xml - deskryptory
- Przed załadowaniem kolejnego roku kartoteka powinna być zamrożona

Wgrywanie kolejnego roku

Wgrywanie
kolejnego roku

Spis tłumaczeń Wgrywanie kartoteki MeSH Panel administratora Analityka Pomoc

mfied

Rozpocznij proces wgrywania kartoteki MeSH

Aby rozpocząć proces wgrywania kartoteki MeSH dodaj plik zip i wybierz odpowiedni rok

Wybierz plik z dysku

Lub przeciągnij tutaj

Wybierz rok

Select year



Potwierdź i rozpocznij



Wgrywanie kolejnego roku

Sukces!



Wgrywanie kartoteki MeSH zakończono powodzeniem

[Zobacz listę tłumaczeń](#)

Podsumowanie

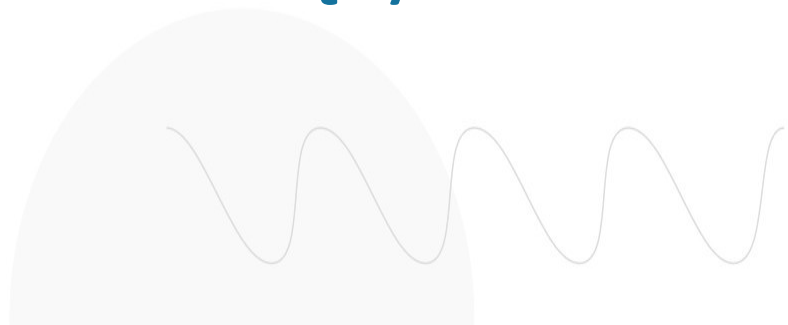
QUALIFIERS

Liczba rekordów w języku angielskim: 80
Liczba przeniesionych tłumaczeń: 83

DESCRIPTORS

Liczba rekordów w języku angielskim: 28472
Liczba usuniętych tłumaczeń: 42
Liczba przeniesionych tłumaczeń: 23880
Liczba nowych tłumaczeń: 629
Liczba zaktualizowanych tłumaczeń: 35

Rozwiązywanie konfliktów



Rozwiązywanie
konfliktów

[Spis tłumaczeń](#) [Wgrywanie kartoteki MeSH](#) [Panel administratora](#) [Analityka](#) [Pomoc](#)

mfiedorowicz

Spis tłumaczeń

Skonfliktowane

W trakcie

W recenzji

Zakceptowane

Nowe

Usunięte

Zaktualizowane

Szukaj



Pliki MARC - marc xml - pobieranie pojedynczych rekordów deskryptorów



[Strona główna](#) [Szukaj](#) [Aktualności](#) [Pomoc](#)

MeSH Descriptor Data 2022

Choroby płuc

Zobacz rekord w języku angielskim

2022

Pobierz w formacie MARC

[Szczegóły rekordu](#)

[Kwalifikatory](#)

[Struktury Drzewa MeSH](#)

[Pojęcia MeSH](#)

Deskryptor

Choroby płuc

Kod języka

PL

Łacińskie tłumaczenia

Pneumopathia

Pneumonopathia

Numer(y) Drzew(a)

C08.381

Unikatowe ID

D008171

Unikalny Identyfikator RDF

<http://id.nlm.nih.gov/mesh/D008171>

Adnotacja

general; prefer specifics; inflammatory

Inne formy

Choroba płuc

Choroby płuca

Choroba płuca

Data utworzenia

Fri Jan 01 1999

Data aktualizacji

Tue Jul 08 2008

Data wprowadzenia

Nukat Number

m 2007939000

[Pokaż RDF](#)

Pliki MARC - marcxml - przykład

```
<?xml version="1.0" encoding="UTF-8"?>
- <marc:collection xmlns:marc="http://www.loc.gov/MARC21/slim">
  - <marc:record>
    <marc:leader>00001nz a1100000n 4500</marc:leader>
    <marc:controlfield tag="008">230616 ||aacnnbaba |a ana |c</marc:controlfield>
  - <marc:datafield tag="040" ind2=" " ind1=" ">
    <marc:subfield code="a">MeSHPol</marc:subfield>
    <marc:subfield code="d">SZCZ010</marc:subfield>
  </marc:datafield>
  - <marc:datafield tag="010" ind2=" " ind1=" ">
    <marc:subfield code="a">m 2001203000</marc:subfield>
  </marc:datafield>
  - <marc:datafield tag="035" ind2=" " ind1=" ">
    <marc:subfield code="a">D001241</marc:subfield>
  </marc:datafield>
  - <marc:datafield tag="072" ind2=" " ind1=" ">
    <marc:subfield code="a">D02.</marc:subfield>
    <marc:subfield code="x">455.</marc:subfield>
    <marc:subfield code="x">426.</marc:subfield>
    <marc:subfield code="x">559.</marc:subfield>
    <marc:subfield code="x">389.</marc:subfield>
    <marc:subfield code="x">657.</marc:subfield>
    <marc:subfield code="x">410.</marc:subfield>
    <marc:subfield code="x">595.</marc:subfield>
    <marc:subfield code="x">176.</marc:subfield>
  </marc:datafield>
  - <marc:datafield tag="150" ind2=" " ind1=" ">
    <marc:subfield code="a">Aspiryna</marc:subfield>
  </marc:datafield>
  - <marc:datafield tag="450" ind2=" " ind1=" ">
    <marc:subfield code="a">Kwas acetylosalicylowy</marc:subfield>
  </marc:datafield>
  - <marc:datafield tag="450" ind2=" " ind1=" ">
    <marc:subfield code="a">Polopiryna</marc:subfield>
  </marc:datafield>
  - <marc:datafield tag="450" ind2=" " ind1=" ">
    <marc:subfield code="a">Aspirin</marc:subfield>
  </marc:datafield>
  - <marc:datafield tag="472" ind2=" " ind1=" ">
    <marc:subfield code="a">Aspirin</marc:subfield>
  </marc:datafield>
  - <marc:datafield tag="550" ind2=" " ind1=" ">
    <marc:subfield code="w">g</marc:subfield>
    <marc:subfield code="a">Salicylany</marc:subfield>
```

Pliki MARC - marc xml

035	Numer unikatowy rekordu MeSH, łącznik między oryginalnym rekordem MeSH a rekordem w Portalu MeSH-Pol oraz w bazach i katalogach.
072	pozycje w drzewie
150	Deskryptor (PL)
45X: 450, 451, 455	Inne formy pol.
472	Hasło główne w języku angielskim
55X \$ a	Odsyłacze typu „zob. też” w języku polskim.
55X \$wg \$a	Hasło nadrzędne w drzewie hierarchicznym
55X \$wh \$a	Hasło podrzędne w drzewie hierarchicznym
670	Informacja o źródle, wymóg licencyjny: podanie wersji np. MeSH 2016

Pliki MARC - marc xml - pobieranie pojedynczych rekordów deskryptorów



[Strona główna](#) [Szukaj](#) [Aktualności](#) [Pomoc](#)

MeSH Descriptor Data 2022

Choroby płuc

Zobacz rekord w języku angielskim

2022

Pobierz w formacie MARC

Szczegóły rekordu

Kwalifikatory

Struktury Drzewa MeSH

Pojęcia MeSH

Dozwolone zawężenia



- diagnostic imaging (DGm 2007939083)
- krew (m 2007939009)
- płyn mózgowo-rdzeniowy (m 2007939011)
- etiologia chemiczna (m 2007939013)
- klasyfikacja (m 2007939014)
- powikłania (m 2007939015)
- przypadki wrodzone (m 2007939016)
- diagnostyka (m 2007939019)
- leczenie dietą (m 2007939021)
- farmakoterapia (m 2007939023)
- ekonomika (m 2007939024)
- embriologia (m 2007939026)
- enzymologia (m 2007939027)
- etnologia (m 2007939028)
- etiopatogeneza (m 2007939029)
- genetyka (m 2007939030)
- historia (m 2007939032)
- immunologia (m 2007939033)

Pliki MARC - marc xml - przykład hasło rozwinięte

```
<?xml version="1.0" encoding="UTF-8"?>
- <marc:collection xmlns:marc="http://www.loc.gov/MARC21/slim">
  - <marc:record>
    <marc:leader>00001cz a1100000n 4500</marc:leader>
    <marc:controlfield tag="008">221215 ||a|cnnbabn |a ann |c</marc:controlfield>
    - <marc:datafield tag="040" ind2=" " ind1=" ">
      <marc:subfield code="a">MeSHPol</marc:subfield>
      <marc:subfield code="d">SZCZ010</marc:subfield>
    </marc:datafield>
    - <marc:datafield tag="010" ind2=" " ind1=" ">
      <marc:subfield code="a">m 2001203002</marc:subfield>
    </marc:datafield>
    - <marc:datafield tag="035" ind2=" " ind1=" ">
      <marc:subfield code="a">D001241Q000008</marc:subfield>
    </marc:datafield>
    - <marc:datafield tag="150" ind2=" " ind1=" ">
      <marc:subfield code="a">Aspiryna</marc:subfield>
      <marc:subfield code="x">dawkowanie</marc:subfield>
    </marc:datafield>
    - <marc:datafield tag="472" ind2=" " ind1=" ">
      <marc:subfield code="a">Aspirin</marc:subfield>
      <marc:subfield code="x">administration & dosage</marc:subfield>
    </marc:datafield>
    - <marc:datafield tag="670" ind2=" " ind1=" ">
      <marc:subfield code="a">MeSH 2022</marc:subfield>
    </marc:datafield>
  </marc:record>
</marc:collection>
```


Pliki MARC - marc xml

010	Numer kontrolny
035	Unikatowy numer ID powstający z połączenia numeru deskryptora z numerem kwalifikatora
150	Hasło rozwinięte w języku polskim (tylko deskryptor z kwalifikatorem)
472	Hasło rozwinięte w języku angielskim (tylko deskryptor z kwalifikatorem)
670	Informacja o źródle, wymóg licencyjny: podanie wersji np. MeSH 2016

Integracja z katalogami bibliotecznymi

System umożliwia

- przekazywanie informacji o nowych/zmodyfikowanych w nowej wersji rocznej hasłach (raz w roku)
- przekazywanie informacji o usuniętych rekordach (raz w roku)
- codzienne przekazywanie informacji o hasłach zmodyfikowanych
 - dodanie tłumaczenia
 - uzupełnione tłumaczeń

Integracja z katalogami bibliotecznymi

- API REST wykorzystywane do generowanie plików
 - marcxml
 - ISO_2709
 - json
 - txt
- jednoznaczny identyfikator każdego rekordu, którym jest numer Unique ID w oryginalnej bazie MeSH mającym swoje miejsce w polu 035 rekordu MARC21

Integracja z omegą



MeSHPol



Integracja z omegą

- moduł meshpol-integrator
 - aktualizacja słownika MeSH w omedze

The screenshot displays a software interface with a central dialog box titled "Properties of meshpol.integration". The dialog box contains a search bar and a table with the following data:

Name	Value	Description	Actions
apiUrl	https://meshdev.cm-uj.krakow.pl/api/rest		
url	https://meshdev.cm-uj.krakow.pl/		

At the bottom right of the dialog box, there is a "CLOSE" button. The background interface shows a "Features" list on the left and a "Properties" list on the right. The "Features" list includes items like "ID", "editorInterface.tooltip.labelE", "editorInterface.menu", "editorInterface.editorialLayo", "topmenu.journals", "topmenu.team", and "topmenu.otherActivities". The "Properties" list includes the number "1". A toggle switch is visible at the bottom center of the interface, labeled "topmenu".

Integracja z omegą

Deskryptory

Szukaj (z użyciem kryteriów formularza) Wyczyść zapytanie Wyszukiwanie przez przykład

Liczba pozycji: 30124 [Popraw zapytanie](#)

Pobierz 0 w formacie: Pełny CSV

- Achyrocline (Achyrocline) [i] [Q]
- Depresja naczyniowa (Vascular Depression) [i] [Q]
- Niedobór kompleksu dehydrogenazy pirogronianowej (Pyruvate Dehydrogenase Complex Deficiency Disease) [i] [Q]

Deskryptory

Szukaj (z użyciem kryteriów formularza) Wyczyść zapytanie Wyszukiwanie przez przykład

Wybierz rekord z listy do edycji; dla wprowadzenia nowego rekordu użyj przycisku

Liczba pozycji: 29714 [Popraw zapytanie](#)

Pobierz 0 w formacie: Pełny CSV

- (pro-3S)-liaza ATP cytrynianowa (ATP Citrate (pro-S)-Lyase) [i] [↺] [Q] [✎] [📄] [🗑️]
- (tiolo)-liaza D-acetyloserynowa (Cysteine Synthase) [i] [↺] [Q] [✎] [📄] [🗑️]
- 1-(5-izochinolinosulfonylo)-2-metylopiperazyna (1-(5-Isoquinolinesulfonyl)-2-Methylpiperazine) [i] [↺] [Q] [✎] [📄] [🗑️]

Integracja z omegą

- Zadanie cykliczne MeshPollImporter
 - pobiera nowe i zmodyfikowane rekordy
- Deskryptory domyślnie importują się po max 1000 rekordów. Można modyfikować - istotne przy pierwszym uruchomieniu integracji.
- możliwość zdefiniowania częstotliwości uruchomienia


Zadania cykliczne

Wybierz rekord z listy do edycji; dla wprowadzenia nowego rekordu użyj przycisku


Odśwież zadania w schedulerze

Aktualnie uruchomione zadania

Liczba pozycji: 1

Uporządkuj wg: aktywny/nazwa  

MeshPollImporter (59 59 23 31 12 ? 2099)

Uporządkuj wg: aktywny/nazwa 

Integracja z omegą - logi

Qualifiers:

```
ImportResult(params=ImportParams(modifiedAfter=Mon Jun 26 20:20:07 CEST 2023,  
year=null, limit=null), countByStatus={}, meshPolLastModified=Tue Jul 04 19:39:59 CEST  
2023, meshPolTotalModifiedCount=0, processedCount=0)  
new modified after: 2023-07-04T19:39:59.742
```

Descriptors:

```
ImportResult(params=ImportParams(modifiedAfter=Tue Aug 23 14:08:42 CEST 2022,  
year=null, limit=1000), countByStatus={UPDATED=973, NO_CHANGES=27},  
meshPolLastModified=Tue Aug 23 14:08:44 CEST 2022, meshPolTotalModifiedCount=56058,  
processedCount=1000)  
new modified after: 2022-08-23T14:08:44.089
```


Integracja z omegą - aktualizacja haseł w rekordach

- Zadanie cykliczne consistencyFixMesh
 - uspójnia hasła w opisach publikacji
 - po zmianie w słowniku poprawiane będą rekordy powiązane
- możliwość zdefiniowania częstotliwości uruchomienia
- zadanie działań na wszystkich rekordów (książki, artykuły, patenty, dane badawcze, doktoraty, inne materiały) które mają możliwość wprowadzania opisów za pomocą haseł MeSH


Odśwież zadania w schedulerze

Aktualnie uruchomione zadania

Liczba pozycji: 1

Uporządkuj wg: aktywny/nazwa  

consistencyFixMesh (59 59 23 31 12 ? 2099)

Uporządkuj wg: aktywny/nazwa 

Usuwanie deskryptorów i kwalifikatorów

Zadania cykliczne

- MeshPolDeletedDescriptorsHandler
 - pobiera z MeSHPol informacje o usuniętych w danym roku deskryptorach
 - usuwa pobrane deskryptory ze słownika MeSH w omedze z opisów rekordów publikacji
- MeshPolDeletedQualifiersHandler
 - pobiera z MeSHPol informacje o usuniętych w danym roku kwalifikatorach
 - usuwa pobrane kwalifikatory ze słownika MeSH i hasła rozwinięte z opisów rekordów publikacji

Usunięte deskryptory i kwalifikatory

- dodaje pole systemowe w rekordzie publikacji
- obsługiwane typy
 - książki,
 - artykuły, rozdziały,
 - patenty,
 - dane badawcze
 - doktoraty
- raport zawierający wszystkie rekordy w których zostały usunięte deskryptory i hasła rozwinięte

*Nazwa pola

mesh_removed_descriptor

Wartość

D051103

Przypadki obsłużone

01	dodanie/aktualizacja/ usunięcie deskryptora	<ul style="list-style-type: none">● podmiana w słowniku MeSH w omedze haseł deskryptorów● usunięcie ze słownika MeSH haseł
02	dodanie/aktualizacja/ usunięcie kwalifikatora	<ul style="list-style-type: none">● podmiana w słowniku MeSH w omedze haseł kwalifikatorów● usunięcie ze słownika MeSH haseł
03	podmiana zmodyfikowanych haseł w opisach publikacji	<ul style="list-style-type: none">● automatyczna podmiana (uspójnienie) opisów w artykułach, książkach, danych badawczych, patentach, doktoratach
04	usunięcie haseł z opisów publikacji	<ul style="list-style-type: none">● raz w roku usunięcie z opisów publikacji haseł deskryptorów i kwalifikatorów● uzupełnienie pól systemowych w celu weryfikacji rekordów
05	raport zawierający rekordy w których zostały usunięte deskryptory lub kwalifikatory	<ul style="list-style-type: none">● raport generowany na podstawie pól systemowych wypełnionych podczas usuwania haseł z publikacji

Przypadki do obsłużenia

Pomysły automatyzacji ale jednak tych przypadków może być na tyle dużo że trudno było to jednoznacznie określić

01

automatyczne zastąpienie usuniętego deskryptora

- automatyczne zastąpienie usuniętego deskryptora wskazanym
- zagrożenie - zbyt ogólne hasła

02

automatyczne zastąpienie usuniętego kwalifikatora

- automatyczne zastąpienie usuniętego kwalifikatora wskazanym
- przypadek z 2017 - kilka kwalifikatorów zastąpione jednym

PubMed

- we wszystkich publikacjach, które posiadają PubMedID hasła MeSH pobierane są automatycznie z PubMed
- możliwe zablokowanie kontroli haseł z PubMed
 - flaga “Synchronizacja: blokuj aktualizację”

Artykuły i rozdziały (w tym materiały konferencyjne) ?

Ostania modyfikacja: ,

Dane podstawowe Dane publikacji Opisy Punktacja systemowa Powiązania **Pola operacyjne**

Identyfikatory w systemach zewnętrznych [Dodaj nowy](#)

Udostępnianie Prywatne Publiczne

Synchronizacja: blokuj propagację

Synchronizacja: blokuj aktualizację

[Sprawdź poprawność](#)

MeSHPol a Polska Platforma Medyczna

- na wszystkich środowiskach lokalnych uruchomiona może być integracja z MeSHPol w pełnym zakresie
- na platformie centralnej uruchomiona zostanie aktualizacja słownika MeSH
- Platforma Centralna będzie agregowała już zmodyfikowane opisy ze środowisk lokalnych
 - nie ma potrzeby uruchamiania zadań uspoźniających



MeSHPol



Polska
Platforma
Medyczna



Dziękuję

Kontakt

m.fiedorowicz@sages.com.pl

sages