



<Medical Semantic Search Engine Program – 2016>

Medical Semantic Search Engine Program – 2016

</Medical Semantic Search Engine Program – 2016>

A Console Application

Copyright (C) 2016. All Rights Reserved. Zakia Salod

INFT8F2H2 - PROGRAMMING MEDICAL INFORMATICS SYSTEMS

ASSIGNMENT 03 : SOFTWARE : MEDICAL SEMANTIC SEARCH ENGINE PROGRAM - 2016

208501583 – Zakia Salod – zakia.salod@ukzn.ac.za

This document serves as documentation / an overview of the Software developed as part of Assignment 03 of this module. It includes an “About” of this program, the Technologies Used, Sample Run of Program, Source Code, & About the Programmer.

| Faculty of Health Sciences |
| School of Nursing & Public Health |
| University of Kwa-Zulu Natal, Durban, South Africa |



INFT8F2H2 - PROGRAMMING MEDICAL INFORMATICS SYSTEMS

ASSIGNMENT 03 : SOFTWARE : MEDICAL SEMANTIC SEARCH ENGINE PROGRAM - 2016

ABOUT : MEDICAL SEMANTIC SEARCH ENGINE PROGRAM – 2016 . . .

WHAT IS IT?

The Medical Semantic Search Engine Program - 2016 is a software used to tag keywords that are found in a text file, and to also search the database for files on the system with the medical term entered by the user. The program reads-in a text file, for example, a medical journal article, Doctor's notes and finds pre-defined keywords, and tags them. The program then writes this tagged information onto a

new .TXT file and stores them onto a "medicaloutputfiles" directory. The pre-defined keywords are read from an (.OWL)

Ontology file, which was created using Protege. Optionally, the user has a facility to also enter medical term/s to find

all files (papers) on the system that has this medical term/s. The program does this by consulting the Ontology file,

and then using the hierarchy from there, together with the search term entered by the user, and then searches the database

for all relevant files. If the term is found on the database, then the program displays all file (paper) information, in

the order of the most relevant - i.e. in terms of the date in which the file was loaded onto the system - in reverse

chronological order (latest entries first).

This program was created by the programmer, as part of a Medical Informatics Programming course at the University of KwaZulu-Natal Nelson R Mandela School of Medicine campus in South Africa in the year 2016.

NOTE

INPUT :

Sample Medical Text File/s & (.OWL) Ontology file can be found here :

- o medicalinputfiles\sampleinputtxtfile1.txt
- o medicalinputfiles\sampleinputtxtfile2.txt
- o medicalinputfiles\sampleinputtxtfile3.txt

Sample Medical Keywords Text File can be found here :

- o medicalinputfiles\medicalterm.owl

OUTPUT :

ABSTRACT

This document serves as documentation / an overview of the Software developed as part of Assignment 03 of this module. It includes an "About" of this program, the Technologies Used, Sample Run of Program, Source Code, & About the Programmer.

FROM TAGGING :-

DIRECTORY :

- o Output file/s are stored in the "medicaloutputfiles" directory.
- o If the "medicaloutputfiles" directory does not exist, the program will create this directory - for the first run of this program.
- o Subsequent runs of the program will store output files onto this directory.

FILE/S :

- o The user is prompted for the output file name. If the file name already exists in the "medicaloutputfiles" directory, the program continues to prompt the user for a unique file name.

FROM USER SEARCH TERM ENTERED :-

CONSOLE :

- o File (paper) information from the database, that satisfies the search term/s entered by the user.

CONTACT

o If you would like more info about the Medical Semantic Search Engine Program, or require free support for this program, please contact the programmer at <http://www.zakiasalod.weebly.com> or zakia.salod@gmail.com

Copyright (C) 2016. All Rights Reserved. Zakia Salod

TECHNOLOGIES USED . . .

- ✓ Programming Language : Java
- ✓ Integrated Development Environment (IDE) : Eclipse LUNA
- ✓ JRE : 1.8.0_45
- ✓ JDK : 1.8.0_45
- ✓ Webservice : USBWebserver v8.6
- ✓ Database : MySQL
- ✓ Mysql-connector : mysql-connector-java-5.1.40-bin.jar
- ✓ Protégé 5.1.0
- ✓ JENA API Library

SAMPLE RUN OF PROGRAM . . .

INPUT

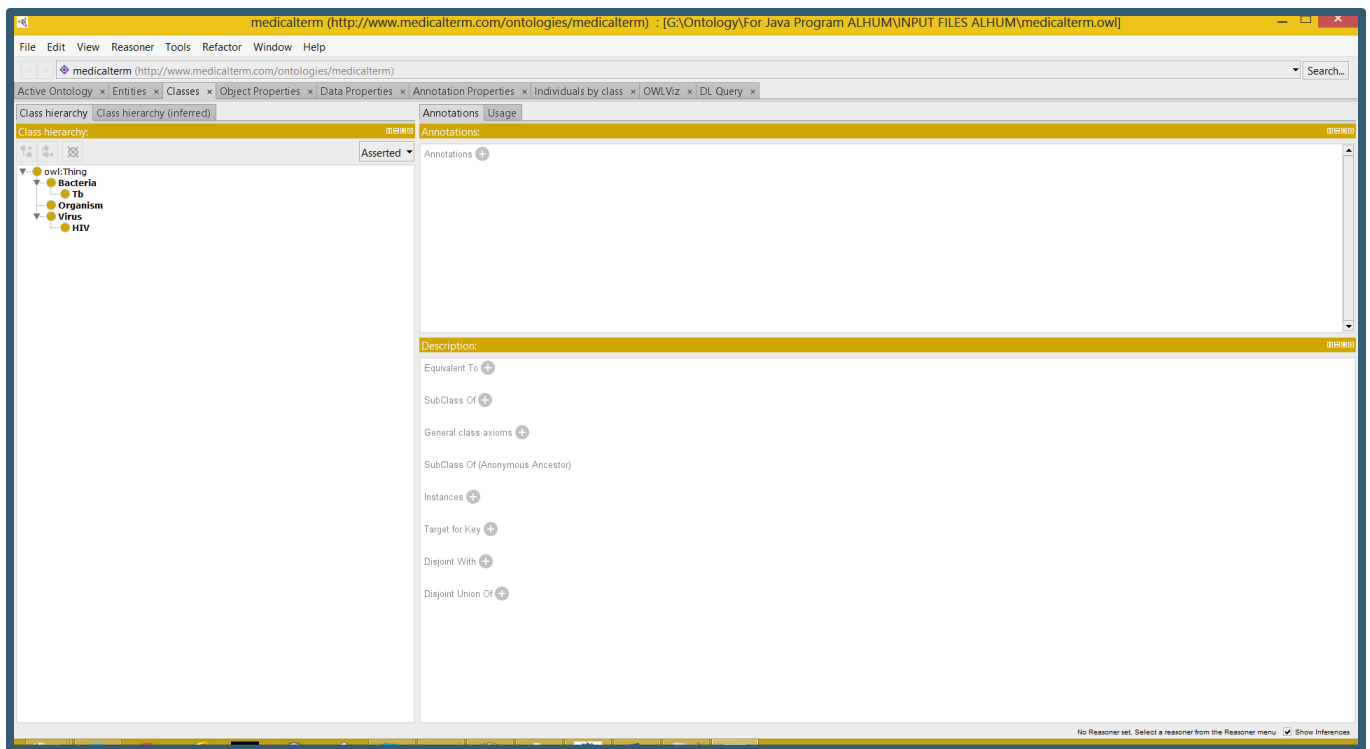
MEDICALINPUTFILES\MEDICALTERM.OWL

{Opened in Protégé}

The screenshot displays the Protégé ontology editor interface for the 'medicalterm' ontology. The main window title is 'medicalterm (http://www.medicalterm.com/ontologies/medicalterm) : [G:\Ontology\For Java Program ALHUM\INPUT FILES ALHUM\medicalterm.owl]'. The menu bar includes File, Edit, View, Reasoner, Tools, Refactor, Window, and Help. The toolbar contains icons for various actions. The 'Active Ontology' tab is selected, showing the 'medicalterm' ontology. The 'Ontology header' panel displays the IRI 'http://www.medicalterm.com/ontologies/medicalterm' and the version 'e.g. http://www.medicalterm.com/ontologies/medicalterm/1.0.0'. The 'Annotations' panel shows a comment 'A medicalterm ontology that stores various medical terms with different medical categories'. The 'Ontology metrics' panel shows a table of metrics:

Metrics	
Axiom	7
Logical axiom count	2
Declaration axioms count	5
Class count	5
Object property count	0
Data property count	0
Individual count	0
DL expressivity	AL
Class axioms	
SubClassOf	2
EquivalentClasses	0
DisjointClasses	0
GCI count	0
Hidden GCI Count	0
Object property axioms	
SubObjectPropertyOf	0
EquivalentObjectProperties	0
InverseObjectProperties	0

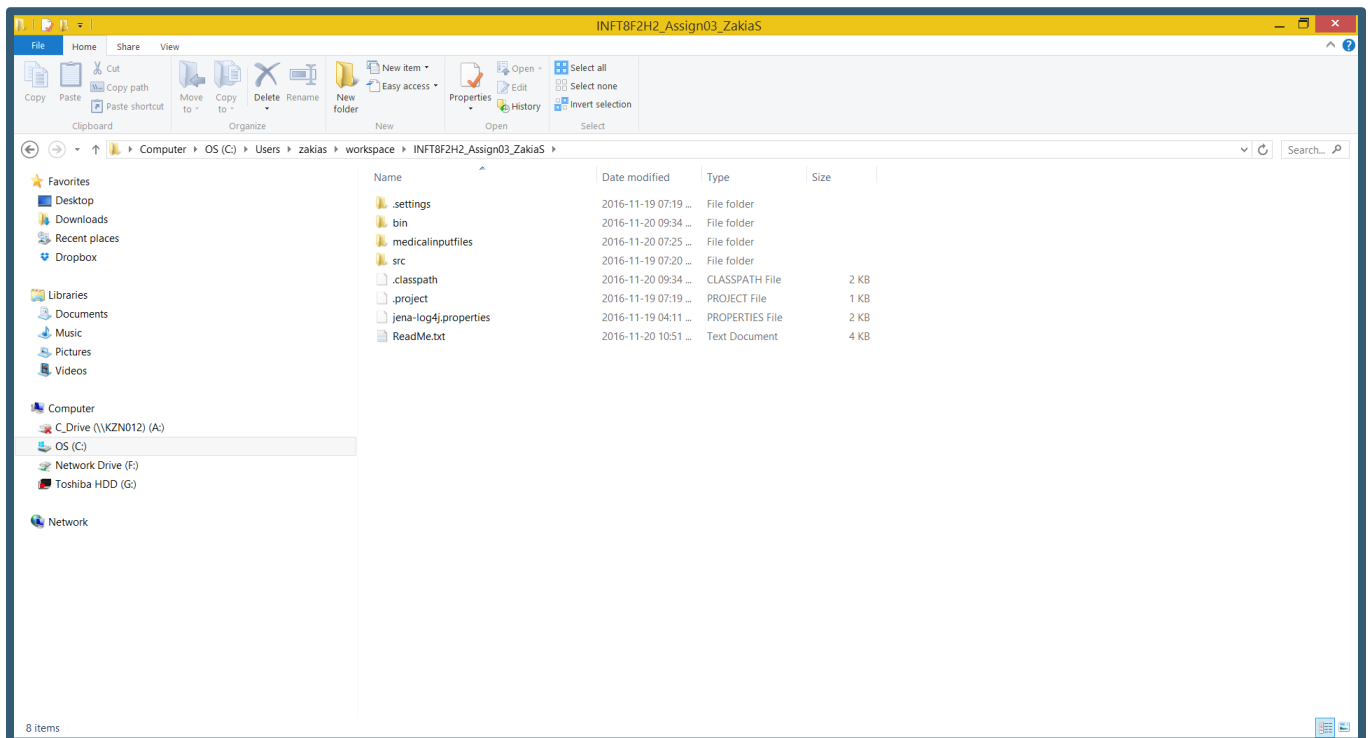
The 'Ontology imports' panel shows 'Imported ontologies' with 'Direct Imports' and 'Indirect Imports' sections. The status bar at the bottom indicates 'No Reasoner set. Select a reasoner from the Reasoner menu' and 'Show Inferences'.



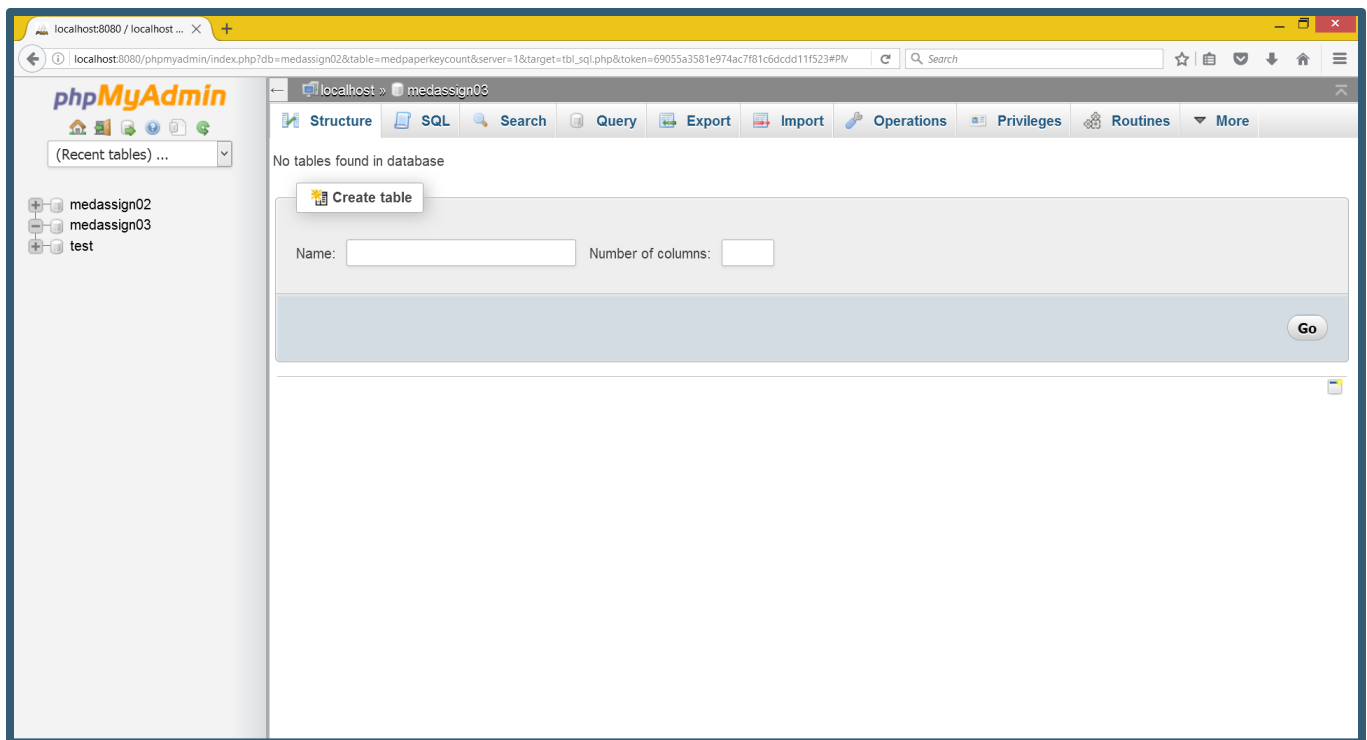
MEDICALINPUTFILES\SAMPLEINPUTTXTFILE1.TXT

Organism Bacteria Tb Virus HIV

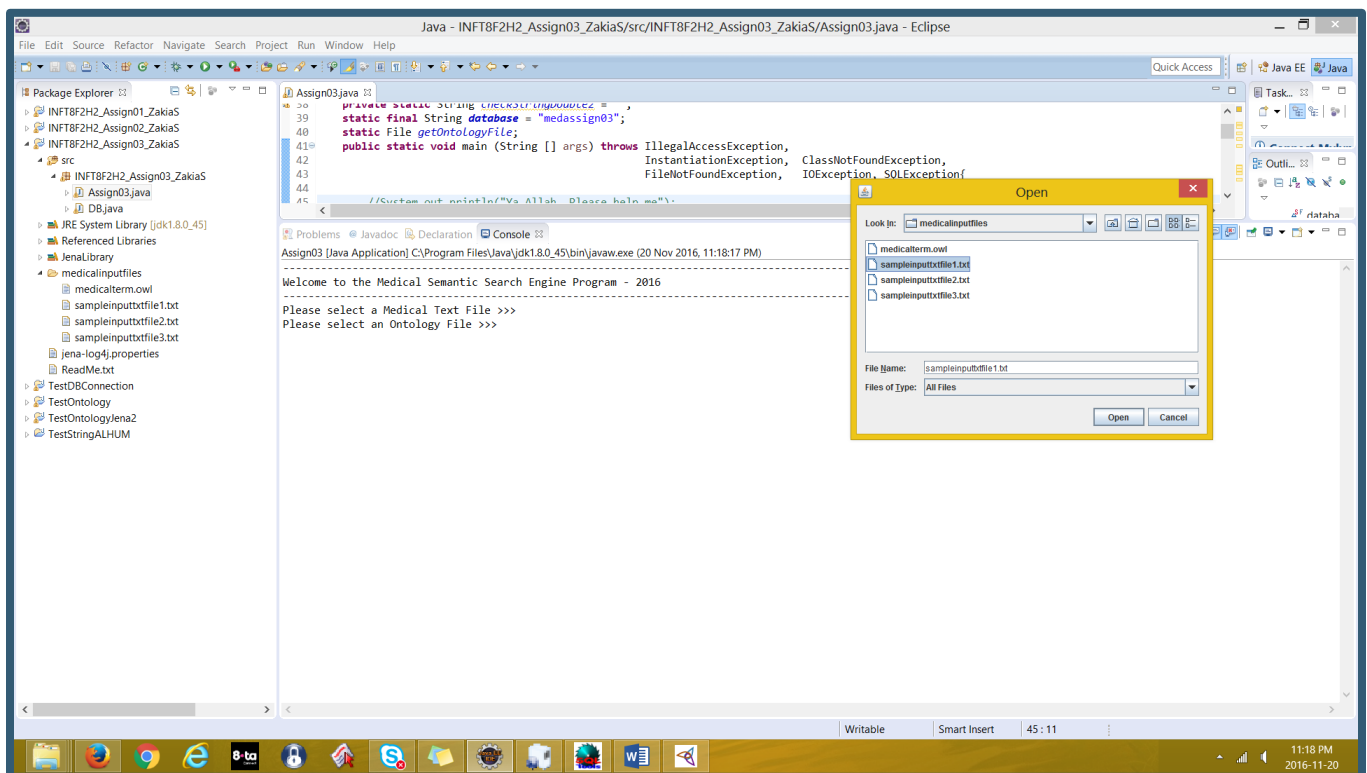
BEFORE



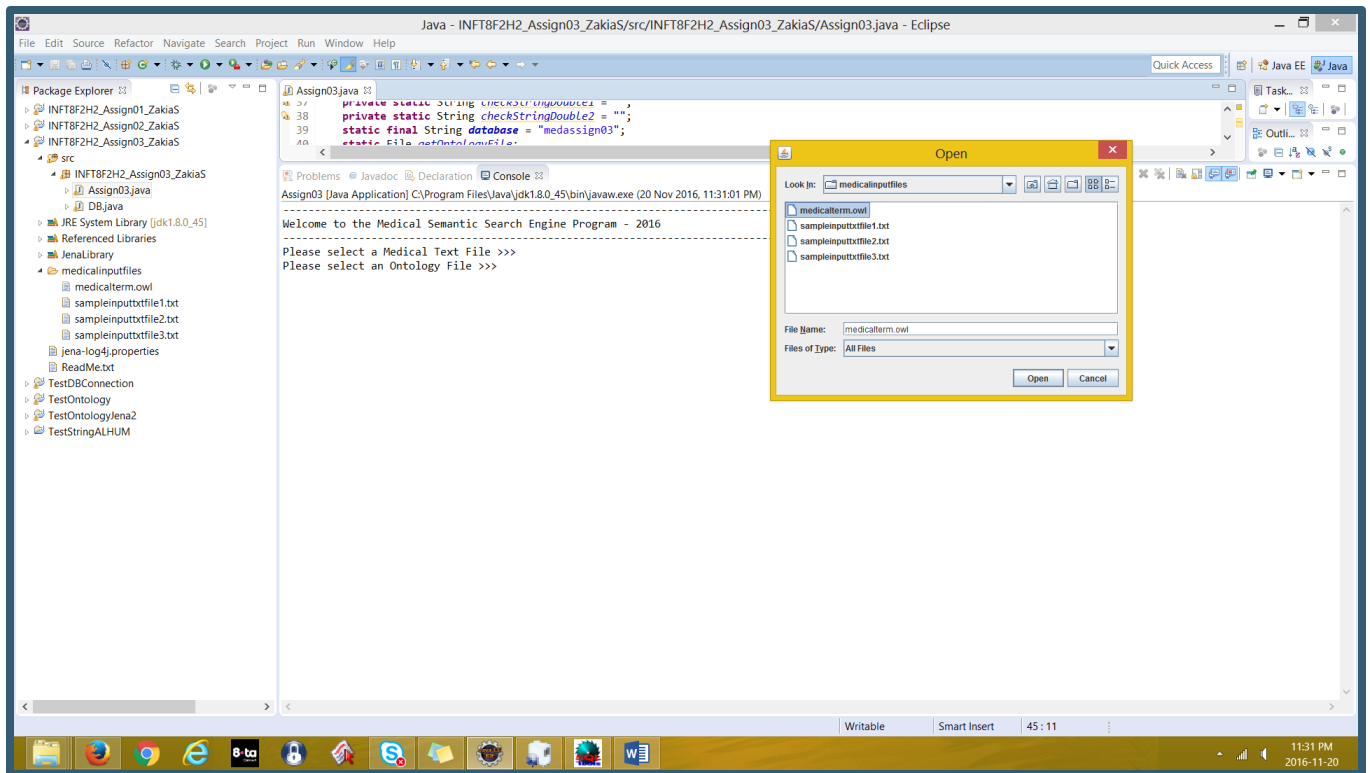
medontologylog Database Table Does NOT Exist in medassign03 Database (Note the medassign03 database node is expanded below, but it does not contain any tables currently)



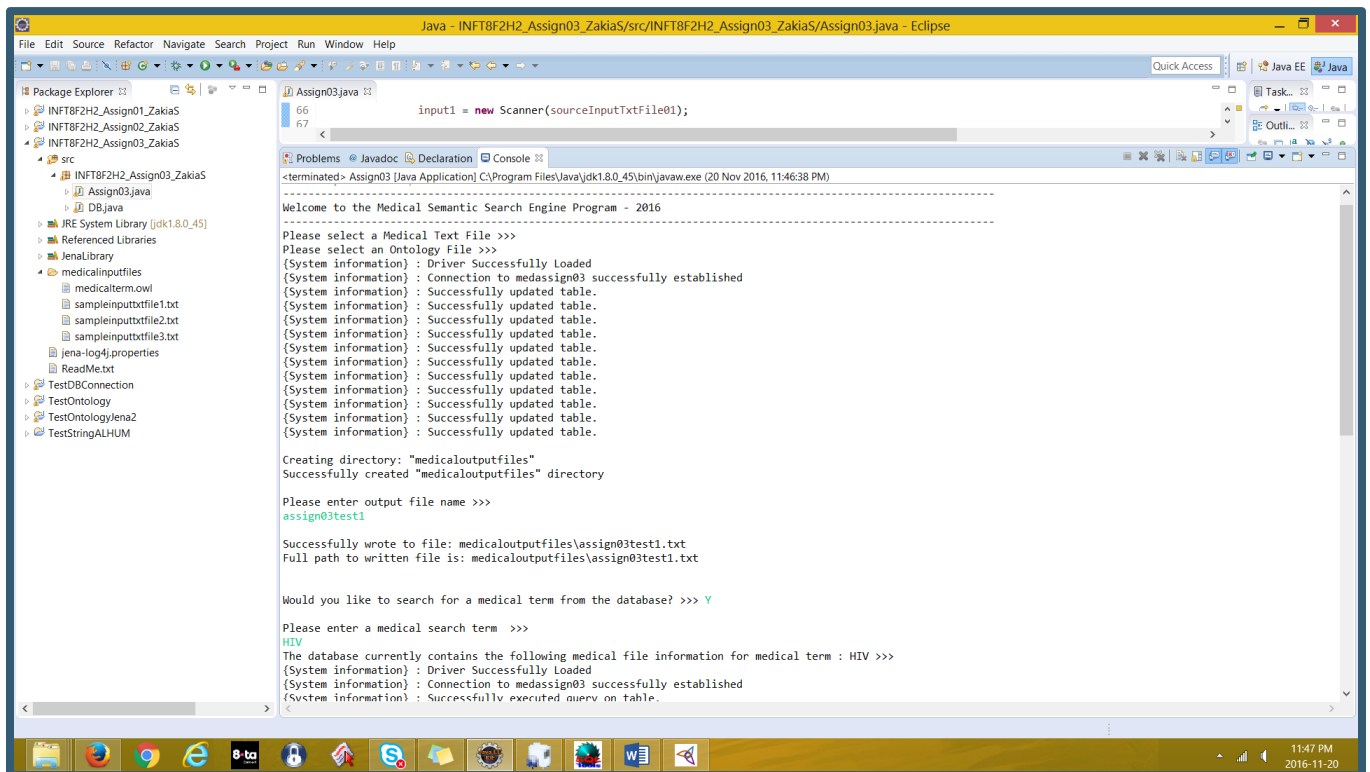
JOURNAL



Click 'Open' on the 'Open' modal dialog box



Click 'Open' on the 'Open' modal dialog box




```
Java - INF8F2H2_Assign03_ZakiaS/src/INF8F2H2_Assign03_ZakiaS/Assign03.java - Eclipse
File Edit Source Refactor Navigate Search Project Run Window Help
Package Explorer
INF8F2H2_Assign01_ZakiaS
INF8F2H2_Assign02_ZakiaS
INF8F2H2_Assign03_ZakiaS
src
INF8F2H2_Assign03_ZakiaS
Assign03.java
DB.java
JRE System Library [jdk1.8.0_45]
Referenced Libraries
JenaLibrary
medicalinputfiles
medicalterm.owl
sampleinputtxtfile1.txt
sampleinputtxtfile2.txt
sampleinputtxtfile3.txt
jena-log4j.properties
ReadMe.txt
TestDBConnection
TestOntology
TestOntologyJena2
TestStringALHUM
Assign03.java
66 input1 = new Scanner(sourceInputTxtFile01);
67
Problems Javadoc Declaration Console
<terminated> Assign03 [Java Application] C:\Program Files\Java\jdk1.8.0_45\bin\javaw.exe (20 Nov 2016, 11:46:38 PM)
Successfully wrote to file: medicaloutputfiles\assign03test1.txt
Full path to written file is: medicaloutputfiles\assign03test1.txt

Would you like to search for a medical term from the database? >>> Y
Please enter a medical search term >>>
HIV
The database currently contains the following medical file information for medical term : HIV >>>
(System information) : Driver Successfully Loaded
(System information) : Connection to medassign03 successfully established
(System information) : Successfully executed query on table.

-----
Paper : sampleinputtxtfile1.txt
Code : 2
Medical Term : HIV
Medical Category : Virus
Program : Assign3
Userstamp : ZS
Timestamp : 2016-11-20 23:47:09.0

-----
Would you like to continue searching for medical terms from the database? >>> Y
Please enter a medical search term >>>
bacteria
The database currently contains the following medical file information for medical term : bacteria >>>
(System information) : Driver Successfully Loaded
(System information) : Connection to medassign03 successfully established
(System information) : Successfully executed query on table.

-----
Paper : sampleinputtxtfile1.txt
Code : 4
Medical Term : Bacteria
```

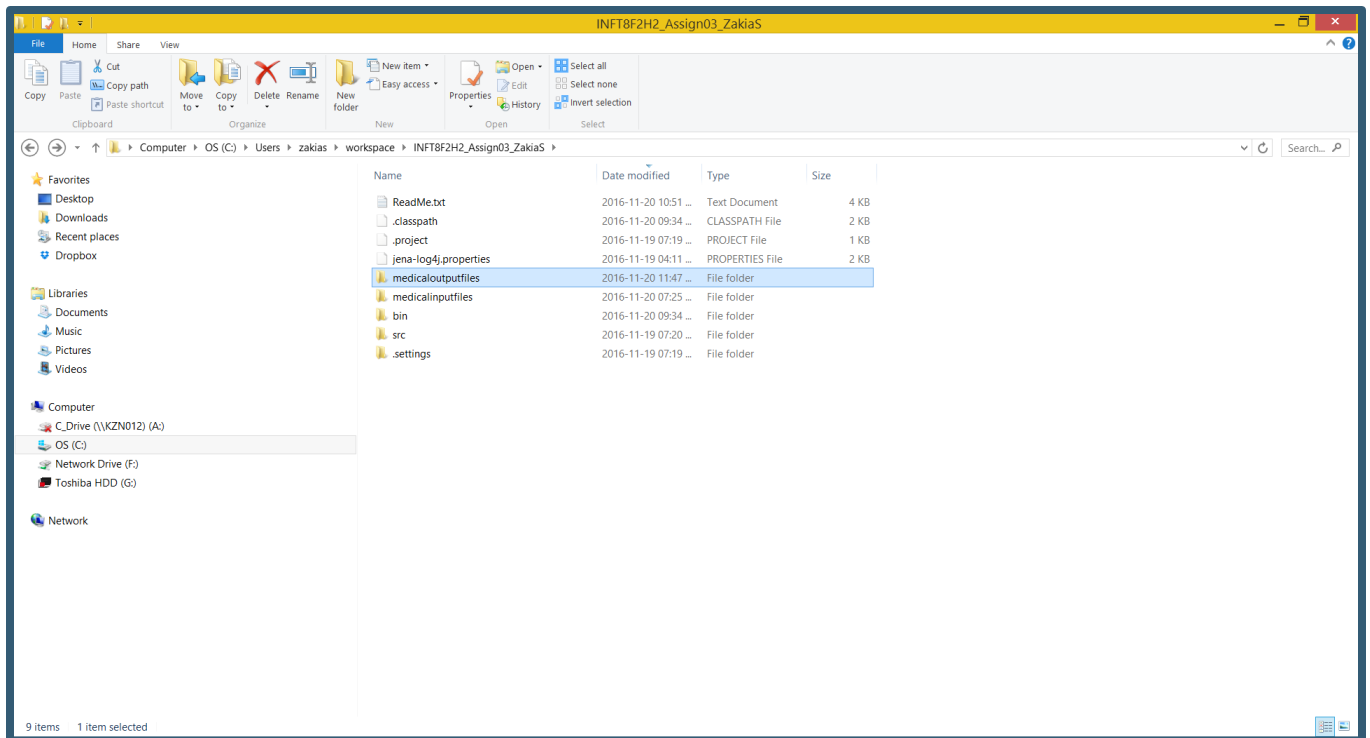
```
Java - INF8F2H2_Assign03_ZakiaS/src/INF8F2H2_Assign03_ZakiaS/Assign03.java - Eclipse
File Edit Source Refactor Navigate Search Project Run Window Help
Package Explorer
INF8F2H2_Assign01_ZakiaS
INF8F2H2_Assign02_ZakiaS
INF8F2H2_Assign03_ZakiaS
src
INF8F2H2_Assign03_ZakiaS
Assign03.java
DB.java
JRE System Library [jdk1.8.0_45]
Referenced Libraries
JenaLibrary
medicalinputfiles
medicalterm.owl
sampleinputtxtfile1.txt
sampleinputtxtfile2.txt
sampleinputtxtfile3.txt
jena-log4j.properties
ReadMe.txt
TestDBConnection
TestOntology
TestOntologyJena2
TestStringALHUM
Assign03.java
66 input1 = new Scanner(sourceInputTxtFile01);
67
Problems Javadoc Declaration Console
<terminated> Assign03 [Java Application] C:\Program Files\Java\jdk1.8.0_45\bin\javaw.exe (20 Nov 2016, 11:46:38 PM)
Code : 2
Medical Term : HIV
Medical Category : Virus
Program : Assign3
Userstamp : ZS
Timestamp : 2016-11-20 23:47:09.0

-----
Would you like to continue searching for medical terms from the database? >>> Y
Please enter a medical search term >>>
bacteria
The database currently contains the following medical file information for medical term : bacteria >>>
(System information) : Driver Successfully Loaded
(System information) : Connection to medassign03 successfully established
(System information) : Successfully executed query on table.

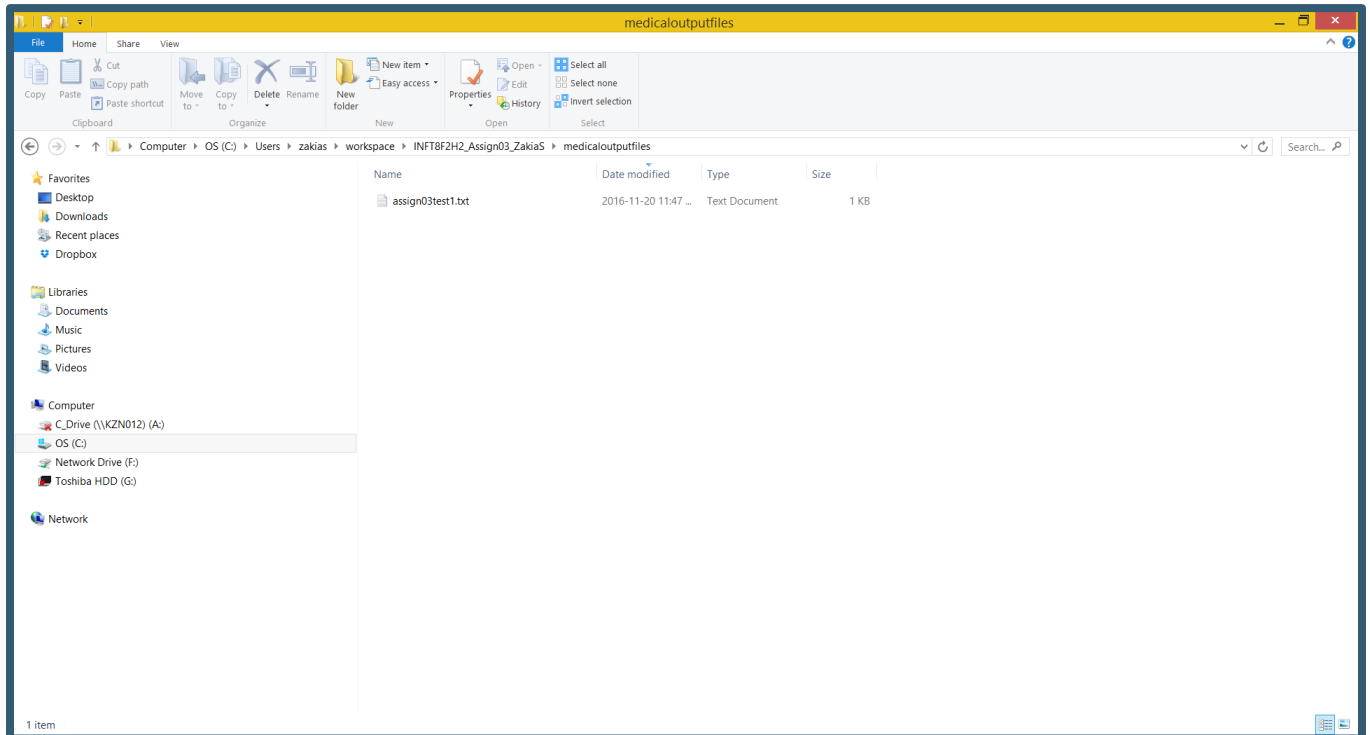
-----
Paper : sampleinputtxtfile1.txt
Code : 4
Medical Term : Bacteria
Medical Category : Bacteria
Program : Assign3
Userstamp : ZS
Timestamp : 2016-11-20 23:47:09.0

-----
Would you like to continue searching for medical terms from the database? >>> N
Exiting program . . . . .
End of Medical Semantic Search Engine Program - 2016
```

AFTER



Click 'medicaloutputfiles'



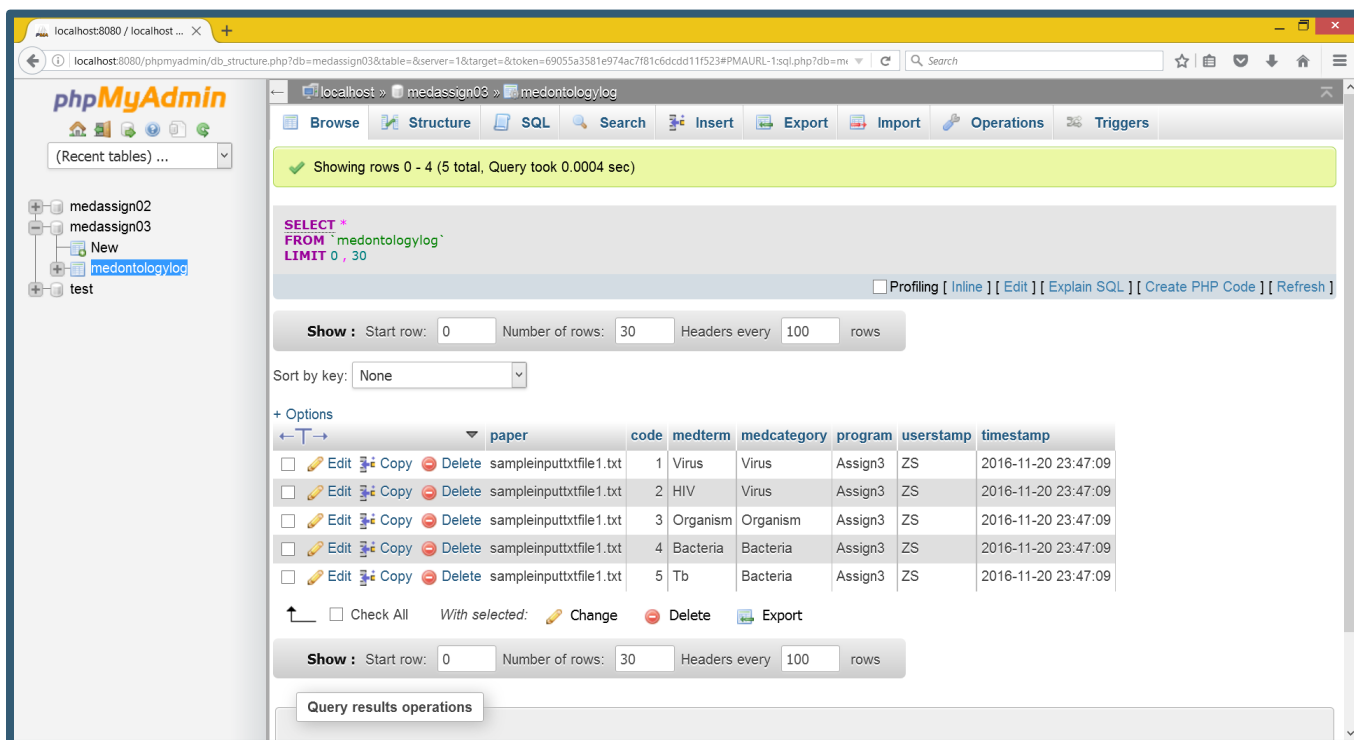
Open 'assign03test1.txt'

```
1 <Organism>Organism</Organism> <Bacteria>Bacteria</Bacteria> <Bacteria>Tb</Bacteria> <Virus>Virus</Virus> <Virus>HIV</Virus>
2
```

Full-text copy-pasted from above file :

```
<Organism>Organism</Organism> <Bacteria>Bacteria</Bacteria> <Bacteria>Tb</Bacteria> <Virus>Virus</Virus>
<Virus>HIV</Virus>
```

medontologylog Database Table DOES Exist in medassign03 Database



Showing rows 0 - 4 (5 total, Query took 0.0004 sec)

```
SELECT *
FROM `medontologylog`
LIMIT 0 , 30
```

Options: ☐ Check All With selected:

	paper	code	medterm	medcategory	program	userstamp	timestamp
<input type="checkbox"/>	sampleinputtxtfile1.txt	1	Virus	Virus	Assign3	ZS	2016-11-20 23:47:09
<input type="checkbox"/>	sampleinputtxtfile1.txt	2	HIV	Virus	Assign3	ZS	2016-11-20 23:47:09
<input type="checkbox"/>	sampleinputtxtfile1.txt	3	Organism	Organism	Assign3	ZS	2016-11-20 23:47:09
<input type="checkbox"/>	sampleinputtxtfile1.txt	4	Bacteria	Bacteria	Assign3	ZS	2016-11-20 23:47:09
<input type="checkbox"/>	sampleinputtxtfile1.txt	5	Tb	Bacteria	Assign3	ZS	2016-11-20 23:47:09

ASSIGN03.JAVA

```

/*****
*Bismillahir Rahmaanir Raheem
*Almadadh Ya Gause Radi Allahu Ta'alah Anh - Ameen
*Student Number : 208501583
*Name : Zakia Salod
*Course : INFT8F2H2
*Assignment : 03
*Masters of Medical Science - Medical Informatics
*Year : 2016
*****/

package INFT8F2H2_Assign03_ZakiaS;
import java.io.*;
import java.net.URI;
import javax.swing.JFileChooser;
import org.apache.commons.logging.LogFactory;
import org.apache.jena.atlas.logging.Log;
import org.apache.jena.atlas.logging.LogCtl;
import org.apache.jena.ontology.*;
import org.apache.jena.rdf.model.ModelFactory;
import org.apache.jena.util.*;
import org.apache.jena.util.iterator.ExtendedIterator;
import java.sql.ResultSet;
import java.sql.ResultSetMetaData;
import java.sql.SQLException;
import java.sql.Timestamp;
import java.nio.channels.FileChannel;
import java.nio.channels.FileLock;
import java.nio.channels.OverlappingFileLockException;
import javax.swing.JFileChooser;
import java.sql.*;
import java.util.*;

public class Assign03 {

    private static String strInput1 = "";
    private static String checkStringDouble1 = "";
    private static String checkStringDouble2 = "";
    static final String database = "medassign03";
    static File getOntologyFile;

    public static void main (String [] args) throws IllegalAccessException,
                                                                    InstantiationException,
ClassNotFoundException,
                                                                    FileNotFoundException,   IOException,
SQLException{

        LogCtl.setCmdLogging();

        System.out.println("-----");
        -----");
        System.out.println("Welcome to the Medical Semantic Search Engine Program -
2016");

```



```

while (classes.hasNext()) {
    OntClass ontologyClass =
        (OntClass) classes.next();
    String ontologyClassStr =
        ontologyClass.getLocalName().toString().trim();

    if
        (ontologyClass.hasSubClass()) {
            if
                (strInput1.toLowerCase().contains(ontologyClassStr.toLowerCase())){
                    //tag medical
                    input file
                    tagMedicalText(ontologyClassStr, ontologyClassStr);
                    //delete then
                    insert latest information into database table
                    code+=1;

                    deletemedontologyLogentry(connectToAssign03Db, "medontologylog",
                        sourceInputTxtFile01.getName(), code, ontologyClassStr, ontologyClassStr);

                    insertmedontologyLogentry(connectToAssign03Db, "medontologylog",
                        sourceInputTxtFile01.getName(), code, ontologyClassStr, ontologyClassStr);
                    }//end if
                    OntClass cla =
                        inf.getOntClass(URI + ontologyClassStr);

                    for (Iterator i =
                        cla.listSubClasses(); i.hasNext();) {
                            OntClass c = (OntClass)
                                i.next();
                                    if
                                        (strInput1.toLowerCase().contains(c.getLocalName().toString().toLowerCase())){
                                            tagMedicalText(c.getLocalName().toString(), ontologyClassStr);
                                            //delete then insert
                                            latest information into database table
                                            code+=1;

                                            deletemedontologyLogentry(connectToAssign03Db, "medontologylog",
                                                sourceInputTxtFile01.getName(), code, c.getLocalName().toString(), ontologyClassStr);

                                            insertmedontologyLogentry(connectToAssign03Db, "medontologylog",
                                                sourceInputTxtFile01.getName(), code, c.getLocalName().toString(), ontologyClassStr);
                                            }//end if
                                            }//end for
                                            }//end if
                                            else
                                                if(!ontologyClass.hasSubClass() && !ontologyClass.hasSuperClass()){
                                                    if
                                                        (strInput1.toLowerCase().contains(ontologyClassStr.toLowerCase())){
                                                            tagMedicalText(ontologyClassStr, ontologyClassStr);
                                                            //delete then insert
                                                            latest information into database table
                                                            code+=1;

```



```

        uniqueOutputFileName = true;
    }//end if
} //end while

if (uniqueOutputFileName == true) {
    PrintWriter output = null;

    try{
        output = new PrintWriter(targetOutputTxtFile);
        output.println(strInput1);

        System.out.println("\nSuccessfully wrote to file: "
+ targetOutputTxtFile);
        System.out.println("Full path to written file is: "
+ targetOutputTxtFile);

        System.out.println();
    } //end try
    catch (FileNotFoundException e){
        e.printStackTrace();
    } //end catch
    finally{
        //close the file
        if (output != null)
            output.close();

    } //end finally
} //end if
} //end try
catch (FileNotFoundException e){
    e.printStackTrace();
} //end catch1
catch (Exception e) {
    System.out.println(e.getMessage());
} //end catch2
    finally{
        //close the medical input file
        if (input1 != null)
            input1.close();
    } //end finally
} //end if

    //handles searching of medical term/s entered by the user, and searches and
displays files in the database with this medical term/s
    userSearchTerm();

    System.out.println("Exiting program . . . . .");

    System.out.println();
    System.out.println("-----");
    System.out.println("-----");
    System.out.println("End of Medical Semantic Search Engine Program - 2016");
    System.out.println("-----");
    System.out.println("-----");

} //end main()

    /**pass-in the same variable (superclass string) for both parameters - for Superclass
    Ontology term tagging

```



```

    * pass-in Subclass Ontology term for "ontologyTerm" and Superclass Ontology term for
    "ontologyTagTerm" for Subclass Ontology term tagging
    */
    private static void tagMedicalText(String ontologyTerm, String ontologyTagTerm){
        strInput1 = strInput1.replace(ontologyTerm, "<" + ontologyTagTerm + ">" +
ontologyTerm + "</" + ontologyTagTerm + ">");
    }

    /**creates database table medontologylog*/
    private static void createDbTable(DB connectToAssign03Db){

        try{
            String createMedOntologyLog = "CREATE TABLE medontologylog(paper
varchar(250) NOT NULL, "+"
                                "code int(11) NOT NULL, medterm
varchar(255) NOT NULL, medcategory varchar(255) NOT NULL, "+"
                                "program varchar(15) NOT NULL,
userstamp varchar(15) NOT NULL, timestamp datetime NOT NULL, "+"
                                "PRIMARY KEY(paper, code, medterm,
medcategory))";

            connectToAssign03Db.updateTbl(createMedOntologyLog);
        }
        catch(SQLException e){
            System.out.println("Table already exists in the " + database + "
database");
        }
    } //end createDbTable()

    /**deletes records from the medontologylog database table, according to the specified
    Primary Key values*/
    private static void deletemedontologylogentry(DB connectToAssign03Db, String
tableName, String paper, int code, String medterm, String medcategory){

        try{
            connectToAssign03Db.updateTbl("DELETE FROM " + tableName + " WHERE
paper='" + paper + "' AND code=" + code + " AND medterm='" + medterm + "' AND medcategory
='" + medcategory + "'");
        }
        catch(SQLException e){
            e.printStackTrace();
        }
    } //end deletemedontologylogentry()

    /**inserts records onto the medontologylog database table*/
    private static void insertmedontologylogentry(DB connectToAssign03Db, String
tableName, String paper, int code, String medterm, String medcategory){

        try{
            String program    = "Assign3";
            String userstamp  = "ZS";
            java.util.Date dt = new java.util.Date();
            java.text.SimpleDateFormat sdf = new java.text.SimpleDateFormat("yyyy-
MM-dd HH:mm:ss");
            String currentTime = sdf.format(dt);

```

```

        connectToAssign03Db.updateTbl("INSERT INTO " + tableName + "(paper, code,
medterm, medcategory, program, userstamp, timestamp) "+
                                "VALUES('"+paper+"',"+code+"','"+
medterm+"','"+medcategory+"','"+program+"','"+userstamp+"','"+currentTime+"
'" +")");
    }
    catch(SQLException e){
        e.printStackTrace();
    }
} //end insertmedontologylogentry()

/**prompts the user on whether or not he/she would like to search the database for a
medical term
 * to retrieve file information for this term/s from the database
 */
private static void userSearchTerm(){
    String promptUser = "";
    Scanner getpromptUser;
    System.out.print("\nWould you like to search for a medical term from the
database? >>> ");
    getpromptUser = new Scanner(System.in);
    promptUser = getpromptUser.nextLine();

    System.out.println();
    //user wants to proceed to search the database
    while (promptUser.equalsIgnoreCase("Y") || promptUser.equalsIgnoreCase("Yes")){
        String searchTerm = "";
        System.out.println("Please enter a medical search term >>> ");
        Scanner getUserMedicalSearchTerm = new Scanner(System.in);
        searchTerm = getUserMedicalSearchTerm.nextLine();
        System.out.println("The database currently contains the following
medical file information for medical term : " + searchTerm + " >>>");
        searchUserTermAndOntology(searchTerm);

        System.out.print("\nWould you like to continue searching for medical
terms from the database? >>> ");
        getpromptUser = new Scanner(System.in);
        promptUser = getpromptUser.nextLine();
    }
} //end userSearchTerm()

/**search the Ontology, by first looking at the user's search term - (searchTerm
passed-in as parameter)**/
private static void searchUserTermAndOntology(String searchTerm){
    OntModel inf = ModelFactory.createOntologyModel(OntModelSpec.OWL_MEM);

    inf.read(getOntologyFile.toString(), "");
    String URI = "http://www.organism.com/ontologies/organism.owl#";

    ExtendedIterator classes = inf.listClasses();
    boolean foundUserMedicalTermInOnt = false;

    //iterate through the Ontology file and look for the Medical term entered by
the user
    while (classes.hasNext()) {
        foundUserMedicalTermInOnt = true;

```


DB.JAVA

```

/*****
*Bismillahir Rahmaanir Raheem
*Almadadh Ya Gause Radi Allahu Ta'alah Anh - Ameen
*Student Number : 208501583
*Name : Zakia Salod
*Course : INFT8F2H2
*Assignment : 03
*Masters of Medical Science - Medical Informatics
*Year : 2016
*****/

package INFT8F2H2_Assign03_ZakiaS;

import java.io.*;
import java.sql.*;

public class DB {
    Connection conn;

    DB(String database){
        final String JDBC_DRIVER = "com.mysql.jdbc.Driver";
        final String DATABASE_URL = "jdbc:mysql://localhost:3307/"+database;

        //connect to mysql driver
        try{
            Class.forName(JDBC_DRIVER);
            System.out.println("{System information} : Driver Successfully
Loaded");
        }//end try
        catch(ClassNotFoundException e){
            System.out.println("{System information} : Unable to connect");
            System.exit(1);
        }//end catch

        try{
            conn = DriverManager.getConnection(DATABASE_URL, "root", "usbw");
            System.out.println("{System information} : Connection to " + database +
" successfully established" );
        }//end try
        catch(Exception e){
            System.out.println(e.getMessage());
        }//end catch
    }//end DB() Constructor

    //Method executes SQL queries, input as string argument
    ResultSet queryTbl(String sqlStmt) throws SQLException{
        Statement stmt = conn.createStatement();
        ResultSet rs = stmt.executeQuery(sqlStmt); // select * from a table
        System.out.println("{System information} : Successfully executed query on
table.");
        return rs;
    }
}
```

```
void updateTbl(String update) throws SQLException{
    Statement stmt = conn.createStatement();
    stmt.executeUpdate(update);
    System.out.println("{System information} : Successfully updated table.");
    stmt.close();
}

void closeDB() throws SQLException{
    conn.close();
    System.out.println("{System information} : Successfully closed table.");
}
} //end DB class
```

ABOUT THE PROGRAMMER . . .

Zakia Salod was born on December 24th 1989 in Durban. She is currently studying full-time towards her Masters in Medical Science Medical Informatics degree at the faculty of Health Sciences at UKZN. She is also working full-time as a Software Developer at a software company, 2Cana Solutions in La Lucia Ridge, Durban – on the Momentum medical aid system.

She graduated with a BSc in Computer Science and IS&T at UKZN in 2010. She had also graduated with a BCom IT Honours (Cum Laude) degree at UKZN in 2011, with first position in her degree from both Westville and Pietermaritzburg campuses.

END OF DOCUMENT