Software User Manual
for the
Semantic Web Crawler
26 May 2012
Prepared for:
National University
Prepared by:
Michael Bastos
Davendra Patel
Manuel Covarrubias Jr.

Approvals

Title	Printed Name	Signature	Date
Client Representative	Dr. Appel		
Project Manager	Manual Covarrubias Jr.		
Project Leader	Davendra Patel		
Project Engineer	Michael Bastos		

Revision Change Record

Revision Number	Revision Date	ECO Number	Description of Change	Author
0	29-Mar-2012		Original	Manager/Leaders

Restrictions.

Table of Contents

Table of Figures

1.0 Introduction

This is the the Semantic Web Crawler (SWC) program User Manuel. The objective of this document is to provide a comprehensive description of the software objectives, define unfamiliar terms and make reference to additional documents which may provide further background where needed.

1.1 Audience Description

The Users Manual provides detailed information on the operation of the SWC program. The client customer and their authorized guests will use the program.

The experience level of the anticipated users will very from essentially no experience to the highly competent level. It is anticipated that the user will be competent in the use of web browsers and the most basic of computer functions. Therefore the system has been designed to support all levels of user's capability.

The SWC program is intended for use by personal and business users who are trying to crawl a specific web page or group of sites for specific types of data that they have a need for. No dedicated training is required to use the SWC program.

1.2 Applicability Statement

The software being provided shall encompass crawler instances made available upon login. The instances shall provide users with the capability to monitor, control, and program automatic operation of various sites and what data shall be crawled as well as when. These applications shall make maximum use of Commercial Off-The-Shelf (COTS) products.

The hardware requirements are:

- 1 Gigahertz (GHz) processor, Pentium or equivalent.
- Linux Operating System
- 2 Gigabytes (GB) of Random Access Memory (RAM)
- 5 GB of hard-disk space for storage
- Monitor
- Keyboard
- Pointing device

The Internet requirements are:

Access to the SWC program requires a full-time Internet connection (i.e. Digital Subscriber Line (DSL) or Cable service), to provide a gateway between the home and the Internet. To accomplish this, any one of the following connection configurations is required:

- Static Internet Protocol (IP) Address. A static IP address is a dedicated IP address assigned to the Internet account by the Internet Service Provider (ISP).

- Dynamic Domain Name Service (DNS) Service - Dynamic DNS is a service that allows the user to alias the assigned Dynamic IP address to a permanent hostname, allowing the computer to be more easily accessed from various locations on the Internet. This allows the user to access the SWC program via the web server over the Internet by entering the static host name instead of the currently assigned Dynamic IP address.

1.3 Purpose Statement

The purpose of this User Manuel is to explain the SWC both internally and externally, give proper use cases as well as determine the direction of the project for the client. The SWC Product scans or crawls through Internet pages to create an index of the data that it is looking for. The SWC will be programmed for long term usage so that the core can require very little maintenance down the road and all the real changes should be made to the User Interface (UI) which is built separately. It will provide the client relevant data from websites the client predefines. The primary purpose of the web crawler is to collect data so that when the client points it at a website, it can provide the client relevant and immediate information in either Excel or Comma Separated Value (CSV) formats pulled directly from that site.

1.4 Document Usage Description

The SWC User Manual provides an overall description of the software product consisting of a high-level perspective of its functionality, users, and constraints, including the assumptions implicit in the design, and the specific description of the system's operation.

1.5 Related Documents List or Information

Bass, Clements, Kazman, Software Architecture in Practice, 2nd Ed., April 2003

Bastos, M., Covarrubias, M., Patel, D., Semantic Web Crawler Operational Concept Document March 2012

Bastos, M., Covarrubias, M., Patel, D., Semantic Web Crawler Software Project Management Project Document March 2012

Bastos, M., Covarrubias, M., Patel, D., Semantic Web Crawler Software Requirements Specification
Document April 2012

Pressman, Roger S., Project Engineering: A Practitioner's Approach, Seventh Edition, McGraw Hill, 2010.

Used for definitions. www.webopedia.com accessed 20 May 2012

1.6 Conventions Description

Definitions, Abbreviations, and Acronyms are provided in the Appendix. No particular symbols, stylistic conventions, and command syntax conventions are used in the document.

1.7 Problem Reporting Instructions

Problems encountered during the operation of the SWC should be reported directly to the SWC Development Team as listed on page 3 of this User Manual.

2.0 User Operations

The User Operations section describes how the user interfaces with the SWC program Graphical User Interfaces (GUI)s to accomplish the required activities. The section is divided into two sections one consisting of an information-oriented instructional document to provide an overall system description and the other consisting of a task-oriented section to provide the user with the information necessary to carry out each of the specific tasks or achieve specific goals.

2.1 The Semantic Web Crawler

2.1.1 Theory

The Semantic Web Crawler shall allow the client to give the application a web address and/or some specific variables to quickly and easily crawl hundreds of pages of a website in a matter of minutes parsing and saving only what data is valuable to the client saving it in either a database or Comma Separated Values (CSV) and Excel spreadsheet format. With proper caching this process can be repeated hundreds of thousands of times with hundreds of websites in various ways allowing the client to better narrow down the type of data they wish to pull from the website without necessarily impacting the crawled site's overall performance or bandwidth.

2.1.2 SWC Program Software Features

The application programs will operate on Linux Operating System (OS) and capabilities.
The user interface shall provide a programming environment that is intuitive to the user and does not depend on the filling in of data other than the users name and password. Additional personal information of the user may be entered but is not considered essential to the system operation and are thus not checked for errors other than field boundaries. The specific programming functions to be provided include the following:

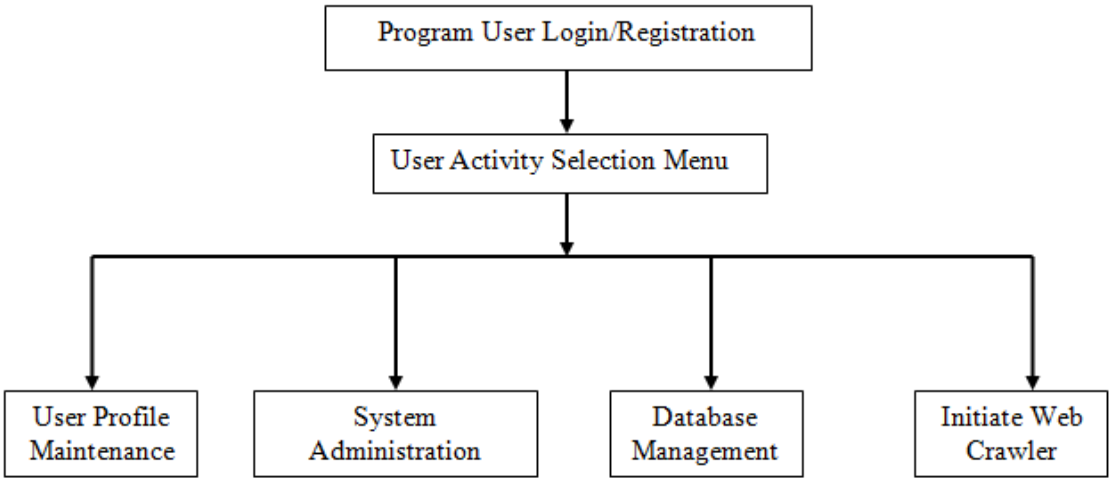
- User Registration
- User Login
- User Activity Selection
- User Profile Maintenance
- System Administration
- Database Management
- Initiate Web Crawler

The program operates on an existing Linux OS and utilizes existing hardware and software interfaces to communicate with Commercial Off-The-Shelf (COTS) equipments and programs.

2.1.3 Software Architecture

The SWC software architecture is Object Oriented and is comprised of seven modules, each of the modules is implemented as a separate application, some of which contain subordinate functions. The components encapsulate data and the operations that must be applied to manipulate the data. Communication and coordination between the components is accomplished via message passing. The individual components have their own separate Graphic User Interfaces (GUI)s that provide the interface between the user and the SWC program to implement activities and display information as shown in the following hierarchical architectural diagram.

FIGURE 2.1 SWC SOFTWARE COMPONENT ARCHITECTURE



2.1 Task Procedures

The Task Procedures section provides the user with the necessary information to carry out the desired operations. The task functions are arranged to follow the flow of the program architecture with each individual function being described in terms of its Scope, Materials required, Preparations, Cautions and Warnings, Methods and any related Information. First time activation of the program on the desktop generates display of the Registration Module, which serves as the starting point for first users. Therefore, this will be the first module discussed.

2.2.1 User Registration

2.2.1.1 Scope

The User Registration Module controls the registration for the SWC software. The purpose of the User Registration is to create a user profile for the SWC program. It ensures that the user creates a User Name, a password to provide access, and to provide an E-mail address for updates and information from the System Administrator.

2.2.1.1 Materials

To complete the login task the user must have a User Name, password, and Email address. The system administrator will accept the User Name of the user and an account is prepared. When the initial password is provided by the user, detailed instructions on the requirements for passwords such as length, character types, update rates etc are given.

2.2.1.3 Preparations

The SWC program must be in operation to execute the registration.

2.2.1.4 Cautions and Warnings

If the User Name provided by the user is already in use, the module will notify the user that the User Name is already in user and to choose another.

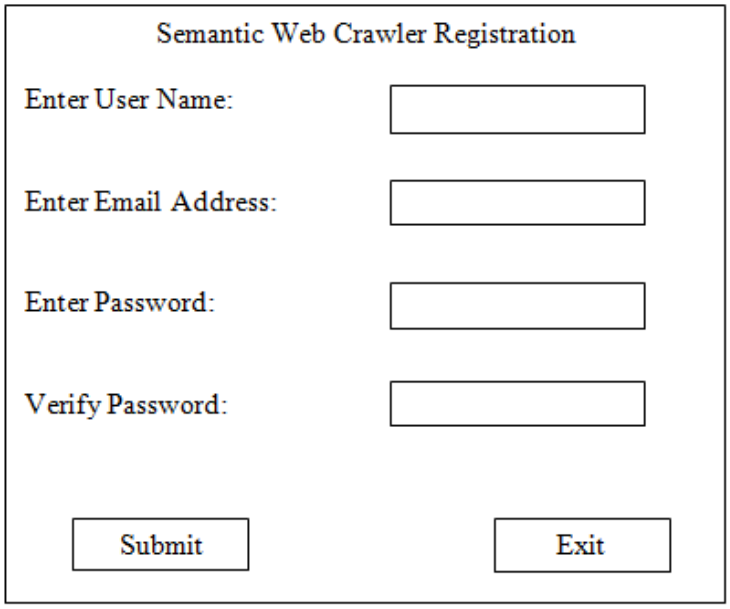
2.2.1.5 Method

Initiating the SWC program for the first time on the desktop accesses the User Registration Module. Once the Registration Module accepts the new user account, the new user must login to gain access to the individual SWC activity modules.

2.2.1.5.1 User Actions

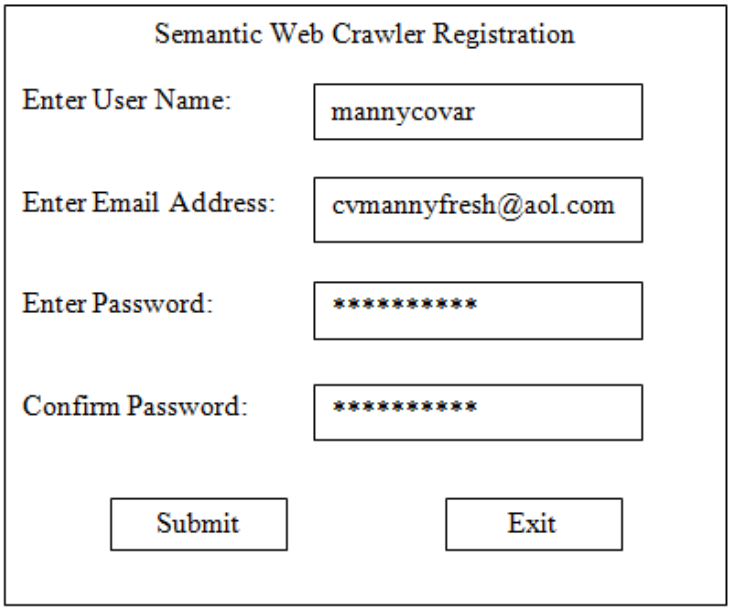
- The user initiates the operation of the SWC program either from the programs menu or a desktop icon.
- The application responds by returning the registration page that requests the user to register. Blank data entry windows for the User Name, password, and Email address are presented with the cursor in the first space of the User Name window. The window fields have appropriate labels next to them to indicate the intended content.

FIGURE 2.2 BLANK REGISTRATION GUI



- In the User Name field, enter your unique User Name using the keyboard. Enter the Email address you wish to choose for updates and notifications. Enter the password you to use and verify it.
- When the three entries have been made submit the data for error checking by clicking on the GUI Submit Key or the Enter Key on the keyboard. The function is cancelled with the Exit key and the program is returned to the background mode.
- Exit the function at any time by clicking on the Exit key.

FIGURE 2.3 REGISTRATION GUI WITH USER INFORMATION



2.2.1.5.2 Functions Invoked

- Upon submission, the security module validates the User Name against the list of authorized users to see if the User Name is already in use.
- If the data is correct, the User Registration Module creates an account for the user.
- When all requirements have been met the module returns the User Login page so that the user can login.

2.2.1.5.3 Errors and Resolution

If the User Name is already in use, the server will request the user to enter another User Name.

- What if I don't have a User Name or my User Name is not accepted by the system?
    - If you do not have a unique User Name or your name is not accepted by the system, contact the System Administrator at the designated number.
- What if after choosing another User Name, my User Name is still not accepted by the system?
    - If your User Name is still not accepted by the system, contact the System Administrator at the designated number or by Email.

2.2.1.5.4 Expected Results

After a successful registration, the Login GUI is presented to the user to login.

2.2.1.6 Related Information

For first time users, the User Registration is the first task in the SWC program operation.

2.2.2 User Login

2.2.2.1 Scope

The Login Module controls the access to the SWC software. It controls the access to the individual SWC activity modules. Based on the user login, the module will determine the class of the user by determining whether they are a system administrator or a user. As a result of determining the class of the user the appropriate activity selection GUI with available activity options is presented by the GUI manager.

2.2.2.2 Materials

To complete the login task the user must have a login User Name and password. The user will provide the login User Name and password.

2.2.2.3 Preparations

The SWC program must be in operation to execute login.

2.2.2.4 Cautions and Warnings

Excessive attempts to login with an incorrect password can result in the user account being locked out and intervention by a System Administrator will be required.

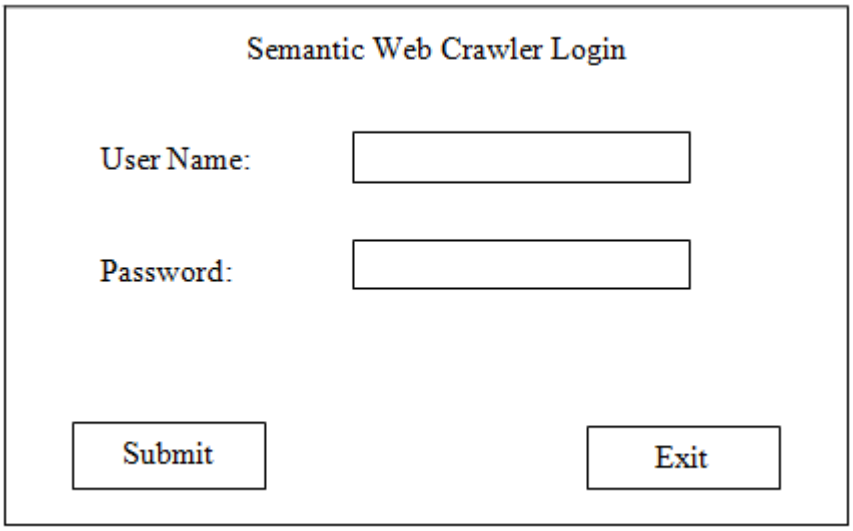
2.2.2.5 Method

Initiating the RCH program on the desktop accesses the User Login Module. As a result User Activity Selection GUI is presented by the GUI manager.

2.2.2.5.1 User Actions

- The user initiates operation of the SWC program either from the programs menu or a desktop icon.
- The application responds by returning the login page that requests entry of the User Name and password. Blank data entry windows for the User Name and password are presented with the cursor in the first space of the User Name window. The window fields have appropriate labels next to them to indicate the intended content.

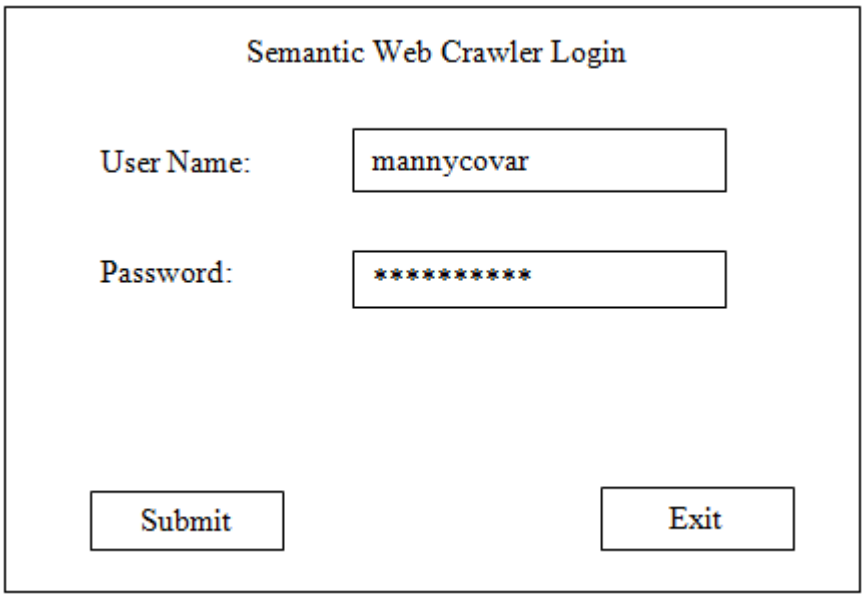
FIGURE 2.4 BLANK LOGIN GUI



- In the User Name field, enter your unique User Name using the keyboard. (You assigned this name to yourself when you registered into the system as an authorized user). Enter the password you have selected and recorded in the database. Following initial login the password may be changed by the user and depending on the security policies, may require periodic change.

- When both entries have been made submit the data for error checking by clicking on the GUI Submit Key or the Enter Key on the keyboard. The function is cancelled with the Exit key and the program is returned to the background mode.
- Exit the function at any time by clicking on the Exit key.

FIGURE 2.5 LOGIN GUI WITH USER INFORMATION



2.2.2.5.2 Functions Invoked

- Upon submission, the security module validates the User Name against the list of authorized users and if the User Name is valid, it checks the password against the security data for that user. If both are valid it checks the password update parameters to determine when the password should be changed.
- If the data is correct, the password is evaluated to determine whether it requires to be updated. If not the procedure continues. If the password requires to be updated the user is notified and the procedure initiates the change password function. When it has been completed the basic function continues.
- When all requirements have been met the module returns the User Activity Selection page to the user that has successfully logged in with a greeting to the individual with their proper name and a welcome page inviting them to proceed with selection of an application from the User Activity Selection GUI.

2.2.2.5.3 Errors and Resolution

If the data is not correct, the server will request the user to re-enter their data. The security function has a limit on the number of attempts permitted. As a result, if successful login is not accomplished within the permitted opportunities, the user is locked out and must contact the System Administrator for access.

- What if I don't have a User Name or my User Name is not accepted by the system?
    - If you do not have a unique User Name or your name is not accepted by the system, contact the System Administrator at the designated number or by Email.
- What if my password is not accepted by the system?
    - If your password is not accepted by the system, contact the System Administrator at the designated number or by Email.
- What if I'm locked out for exceeding the maximum number of password entries?
    - If you are locked out by the system you must contact the System Administrator at the designated number or by Email to have your password issue resolved.

2.2.2.5.4 Expected Results

A successful Login results in the User Activity Selection page being returned to the user. The user is presented with a greeting using their User Name and a welcome page inviting them to proceed with a selection of an application from the User Activity Selection GUI.

2.2.2.6 Related Information

For returning users, the User Login is the first task in the SWC program operation and is directly related to all subsequent tasks.

2.2.3 User Activity Selection

2.2.3.1 Scope

The purpose of the User Activity Selection GUI is to provide the capability to select the activity the user wishes to perform. The User Activity Selection GUI is accessed by successfully logging into the SWC program. The GUI manager accepts the User Name and password from the Login function and returns the User Activity Selection GUI.

2.2.3.2 Materials

No materials are required for this function.

2.2.3.3 Preparations

The use of this function requires a successful login to the system and operation within the GUI management program.

2.2.3.4 Cautions and Warnings

No Cautions and Warnings are associated with this function.

2.2.3.5 Method

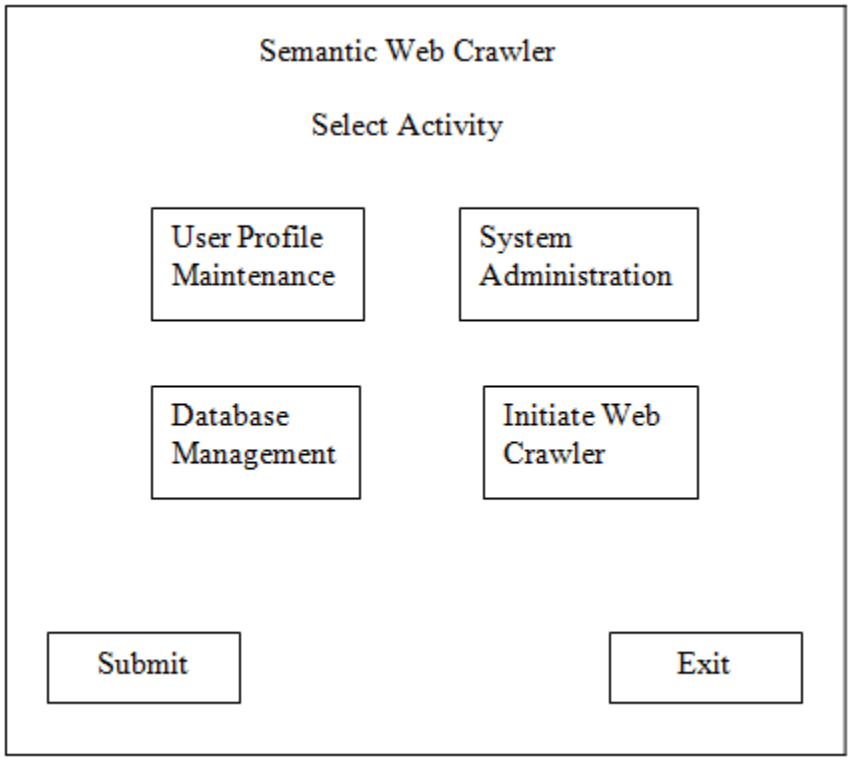
When the user first enters the SWC program or when they have completed a specified activity and wish to exit the program or change to a different activity, they navigate back to the User Activity Selection Menu which displays the following activities:

- User Profile Maintenance
- System Administration
- Database Management
- Initiate Web Crawler

2.2.3.5.1 User Actions

In response to a successful Login or upon exiting activities, the user is presented with the User Activity Selection Menu GUI.

FIGURE 2.6 USER ACTIVITY SELECTION MENU GUI



- The user selects the desired activity on the GUI either with the pointing device, such as a mouse, or by tabbing up and down with the Arrow keys and initiates the activity either with the GUI Submit button or depressing the Enter key.
- Selection of Exit button on the User Activity Selection GUI causes the application to close and returns the user to the previous status of the OS.

2.2.3.5.2 Functions Invoked

The application responds by returning the entry page of the selected module.

2.2.3.5.3 Errors and Resolution

There is no user data entry function associated with this activity.

- What if the functions I select does not take me to the proper application?
    - If the selected function does not take you to the proper application you must contact the System Administrator at the designated number or by Email.

2.2.3.5.4 Expected Results

Selection of a user activity results in display of the first GUI page of the chosen activity.

2.2.3.6 Related Information

The User Activity Selection Menu application is initiated by a successful login and provides the link to the activity that is chosen.

2.2.4 User Profile Maintenance

2.2.4.1 Scope

The purpose of the User Profile Maintenance Module is to allow the user to update or change any information of their personal user profile to keep it up to date. These items that the user can update are their Email address, password, and website preferences. The user modifies their user profile through the User Profile Maintenance GUI. The revisions to the profile are recorded in the database after error checking has been successful.

2.2.4.2 Materials

No materials are required for this function.

2.2.4.3 Preparations

Access to this function is through the User Activity Selection Menu.

2.2.4.4 Cautions and Warnings

There are no Cautions and Warnings that are associated with this function. None of the data entered by the user is considered critical to the system and therefore no error checking, other than field limits, is conducted.

2.2.4.5 Method

The User Profile Maintenance GUI is accessed by navigating through the SWC User Activity Selection Menu. When it is selected it displays the data currently in the database that the user has to update and allows the update of the data. The contents of the fields are modified by selecting the field with the pointer or by tabbing up or down and using the standard keyboard functions to edit the contents within the field. All, or some, of the following data options may be updated by the user:

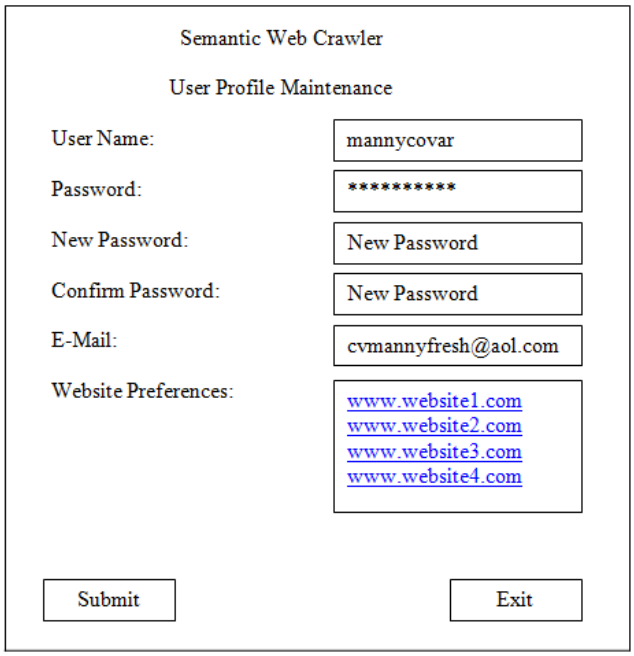
- E-mail Address
- Password
- Website Preferences

When the data has been changed on the GUI, the user submits the changes with either the Submit button or the Enter key. When the data has been successfully updated in the database, the module will return a message indicating success. If any errors are found during error checking, the field in error will be highlighted and the user will be requested to modify the highlighted data. When the user has completed updating their user profile, the user returns to the User Activity Selection Menu by selecting the Exit button.

2.2.4.5.1 User Actions

- Selection of the User Profile Maintenance on the User Activity Selection Menu results in the display of the User Profile Maintenance GUI containing the User Name and the fields the user is authorized to update with the data currently held in the database.

FIGURE 2.7 USER PROFILE MAINTENANCE GUI



- The user updates the data by tabbing up or down using the Arrow keys or using the pointer to identify the field to be updated and edits the data using the keyboard.
- When the user password is changed the Change Password function is called to process the new password with a confirmation routine.
- When editing is complete, the user submits the new data for error checking either by clicking to the Submit button or by pressing the Enter key.
- The system will respond to the submission request with either an error message and the erroneous field(s) highlighted or a verification request. The verification options are:
    - Yes
    - No, Continue Editing
    - Option selection is made either with the pointer or by tabbing up and down.
- When all the data has been corrected and changes verified the user resubmits the changes and the system will respond with either an error message or verification that the data has been updated.
- When editing has been completed the Exit button is selected to return to the User Activity Selection Menu.

2.2.4.5.2 Functions Invoked

- When the user password is updated the Change Password function is executed.
- The application updates the user profile information in the database.

2.2.4.5.3 Errors and Resolution

None of the data in this function is considered critical to the system operation and therefore is only checked for field limits. If the data entered is not consistent with the requirements of the database the field is highlighted and an error message is displayed. Data is not transferred to the database until the error checking is successful.

- What if I receive an error message and one or more of the fields are highlighted?
    - The data you have entered has an incorrect parameter.
        - Verify that the characters are of the correct type by consulting the HELP function.
        - Verify that the number of characters does not exceed the limits of the field by consulting the HELP function.
    - If the above actions do not resolve the situation contact the System Administrator at the designated number or by Email.
- What if I receive an error message while changing my password?
    - The password confirmation you entered did not match the new password you entered.
        - Reenter the new password and confirmation.
    - The password you entered did not meet the password parameter requirements.
        - Verify the new password meets the parameter requirements by consulting the HELP function.
    - If the above actions do not resolve the situation contact the System Administrator at the designated number or by Email.

2.2.4.5.4 Expected Results

The data fields will be updated with the new data.

2.2.4.6 Related Information

The User Profile Maintenance function is accessed from the User Activity Selection Menu. The change password activity is a sub-function.

2.2.5 System Administration

The System Administration Module provides the user the capability to manage the user profiles by creating new user profiles and editing or deleting existing user profiles. The These files are used by the Login function to control access to the program and by the GUI Manager to determine which user profile(s) are to be provided to the logged in user.

2.2.5.1 System Administration Menu

2.2.5.1.1 Scope

The System Administration GUI is accessed through the User Selection Activity Menu. The GUI provides a menu for the user to select Create New User Profile, Modify or Delete User Profile or Exit.

2.2.5.1.2 Materials

No materials are required for this function.

2.2.5.1.3 Preparations

Access to this function is through the User Activity Selection Menu.

2.2.5.1.4 Cautions and Warnings

There are no Cautions and Warnings that are associated with this function. No data entry by the user is required.

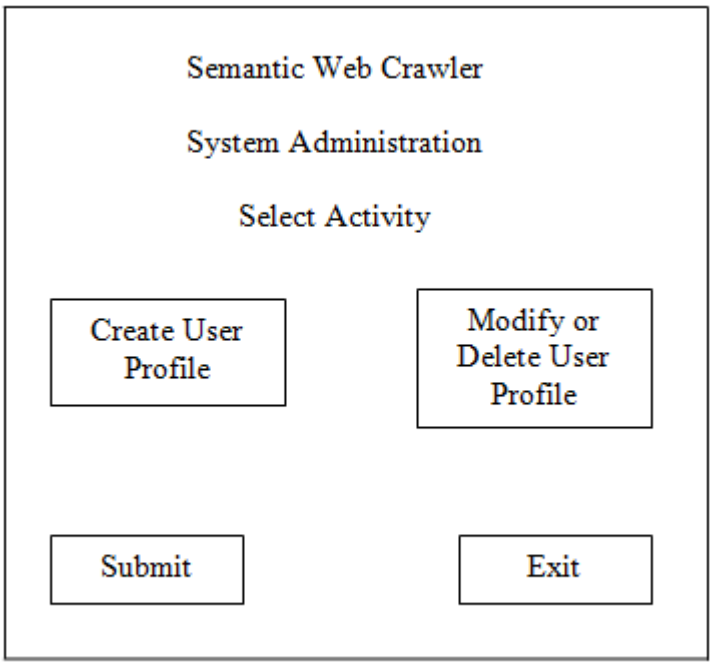
2.2.5.1.5 Method

The user selects the System Administration activity to be performed with the pointer and pressing the Submit button or the Enter key. Selecting the Exit button will return the user to the User Activity Selection Menu.

2.2.5.1.5.1 User Actions

- The user selects the System Administration Module from the User Activity Selection Menu.
- The module responds by displaying the System Administration GUI that provides the option of either Create New User Profile or Modify or Delete User Profile.

FIGURE 2.8 SYSTEM ADMINISTRATION GUI



- The user can either select the Create New User Profile, the Modify or Delete User Profile, or Exit.
- Selection of Exit will return the user to User Activity Selection Menu.

2.2.5.1.5.2 Functions Invoked

The System Administration Menu will transfer the user to the selected module or exits back to the User Activity Selection Menu.

2.2.5.1.5.3 Errors and Resolution

There are no errors that are associated with this function.

2.2.5.1.5.4 Expected Results

The user is transferred to the selected module.

2.2.5.1.6 Related Information

The System Administration Menu is accessed from the User Activity Selection Menu and provides access to the functions: Create New User Profile and Modify or Delete User Profile.

2.2.5.2 Create New User Profile

2.2.5.2.1 Scope

The Create User Profile Module allows the user to create a new user profile within the SWC system. The user profiles contain identifying factors, including contact information, of the user that can be used to identify them. The user profile is stored in the database and accessed as part of the user authorization process to validate the user's access to the system.

2.2.5.2.2 Materials

There are no materials that are required for this function.

2.2.5.2.3 Preparations

Access to this function is through the System Administration Menu.

2.2.5.2.4 Cautions and Warnings

The creation of a new user profile defines the access of the user and should be carefully considered with respect to what is really required. The critical entries in this application are the User Name that must be unique and will be subject to error checking. None of the other data entered is considered critical to the system and therefore there is no error checking, other than field limits, is conducted.

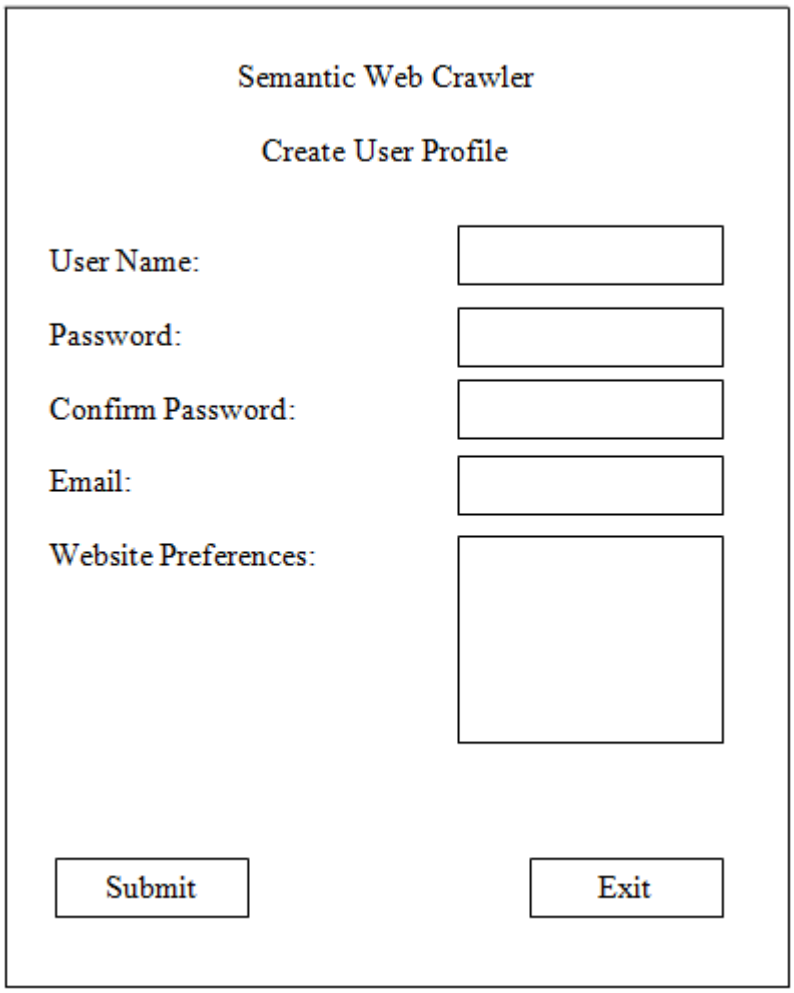
2.2.5.2.5 Method

The user selects the Create User Profile on the System Administration Menu calling the application and displaying a blank user profile GUI. The user fills in the fields with data entry. When the data entry is complete the user submits the data to the database with the Submit button. The module will error check the entered data and either updates the database or returns an error message with the erroneous data highlighted. The user returns to the System Administration Menu with the Exit button.

2.2.5.2.5.1 User Actions

- The User selects the Create User Profile on the System Administrator GUI and selects the Submit or presses the Enter key on the keyboard.
- The Create User Profile returns a blank user profile GUI for the insertion of information with data entry using the keyboard.

FIGURE 2.9 CREATE NEW USER PROFILE GUI



- The user uses the keyboard to enter the mandatory fields of a unique User Name, password, Email address, and the user's website preferences.
- When data entry is complete the user submits the form for error checking by either selecting the Submit button or by pressing the Enter key on the keyboard.
- The module error checks the form to ensure that the required fields of the User Name and password have met the established parameters.
- The module will then check the new User Name against the database to confirm that it is unique.
- When all of the data has been confirmed the new user profile is recorded in the database and a confirmation is returned to the user.
- If an error message is received the user can either correct the highlighted data or exit the module with the Exit button without modifying the database.

2.2.5.2.5.2 Functions Invoked

The functions used during the creation of a new user profile will determine whether the new User Name is unique and meets the parameters and that the password meets the parameters.

2.2.5.2.5.3 Errors and Resolution

For this function, all the data field entries will be error checked. The critical fields of the form are checked as follows:

- The User Name field is checked for uniqueness and complies with the field parameters.
- The password field is checked for compliance with the field parameters.
- The personal data fields are checked for compliance with field parameters.
- What if I receive an error message and one or more of the fields are highlighted?
    - The parameter of the data you entered is incorrect.
    - Verify that the entries are of the correct type by consulting the HELP function.

2.2.5.2.5.4 Expected Results

The new user profile is entered into the database and a confirmation message is returned to the user.

2.2.5.2.5.5 Related Information

The Create User Profile module is accessed from the System Administrator Menu.

2.2.5.3 Modify or Delete User Profile

2.2.5.3.1 Scope

The Modify or Delete User Profile Module allows the user to modify or delete an existing user profile within the SWC system. The user profile contains identifying factors of the user that can be used to identify them and has their contact information. The user profile is stored in the database and accessed as part of the user authorization process to validate the user's access to the system.

2.2.5.3.2 Materials

Entry of the User Name of the profile to be accessed is required to modify or delete the user profile.

2.2.5.3.3 Preparations

The System Administration Menu is used to access this function.

2.2.5.3.4 Cautions and Warnings

The critical entry in this application is the User Name, which is required for access. None of the other data is considered critical to the system and therefore no error checking, other than field limits, is conducted. The deletion of a user profile is permanent.

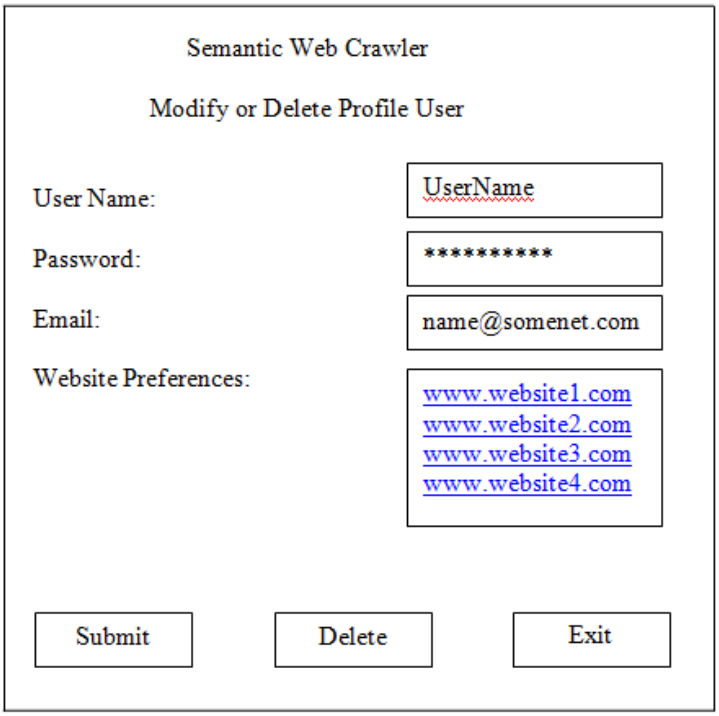
2.2.5.3.5 Method

The user selects the Modify or Delete User Profile on the System Administration Menu. The application requests entry of the User Name that is to be modified or deleted. The submission of the User Name will result in the Modify or Delete User GUI being returned with the data currently in the database. The user can either modify the data and submit it to the database or deletes it from the database. The module error checks the modified data and confirms that the modify action is to be accomplished prior to making any changes to the database. The user returns to the System Administration Menu with the Exit button.

2.2.5.3.5.1 User Actions

- The User selects the Modify or Delete User Profile on the System Administrator Menu and selects the Submit button or presses the Enter key on the keyboard.
- The Modify or Delete User Profile module will request the entry of the User Name to be modified or deleted.
- The user enters the User Name with the keyboard and submits the entry by either selecting the Submit button on the GUI or by pressing the Enter key on the keyboard.
- The Modify or Delete User GUI is returned with the data currently in the database for the User Name submitted.

FIGURE 2.10 MODIFY OR DELETE USER PROFILE GUI



- The user can delete the user profile by selecting the Delete button on the GUI.
- The module will request the confirmation of the action prior to deleting the user profile from the database.
- When the user profile has been deleted a confirmation message is displayed.
- The user can modify the user profile by selecting the fields to be modified with the pointing device and entering the data with the keyboard.
- When user profile has been modified the user submits the data for error checking by either selecting the Submit button on the GUI or by pressing the Enter key on the keyboard.
- The module error checks all the fields on the form and either returns an error messages or forward the entries to the database for update.
- If an error message is received the user can either correct the highlighted data or exit the GUI with the Exit button without modifying the database.

2.2.5.3.5.2 Functions Invoked

The functions used during the user profile modification will determine that the keyboard entries conform to the field limits.

2.2.5.3.5.3 Errors and Resolution

For this function all the data field requiring keyboard entries are error checked. The critical fields of the form that are checked as follows:

- The User Name field is checked for uniqueness and complies with the field parameters.

- The password field is checked for compliance with the field parameters.
- The personal data fields are checked for compliance with the field parameters.
- What if I receive an error message and one or more of the fields are highlighted?
    - The parameter of the data you entered is incorrect.
    - Verify that the entries are the correct type by consulting the HELP function.

2.2.5.3.5.4 Expected Results

The selected user profile is either modified or deleted and a confirmation message is displayed.

2.2.5.3.6 Related Information

The Modify or Delete User Profile is accessed from the System Administrator Menu.

2.2.6 Database Management

2.2.6.1 Scope

The SWC program modules and the GUI applications request data from the database and update data in the database either automatically or upon request. In order to manage these operations the SWC module requests deployment of the Commercial Off-The-Shelf (COTS) database application on the host computer to enable the user the capability to manipulate the database. When the database management functions are complete the COTS database application returns control to the SWC program.

2.2.6.2 Materials

For this function no materials are required.

2.2.6.3 Preparations

The System Administration Menu is used to access this function.

2.2.6.4 Cautions and Warnings

All the database activities are completed outside of the SWC program. There are no Cautions and Warnings that are associated with this function. No data entry within the SWC program is required by the user.

2.2.6.5 Method

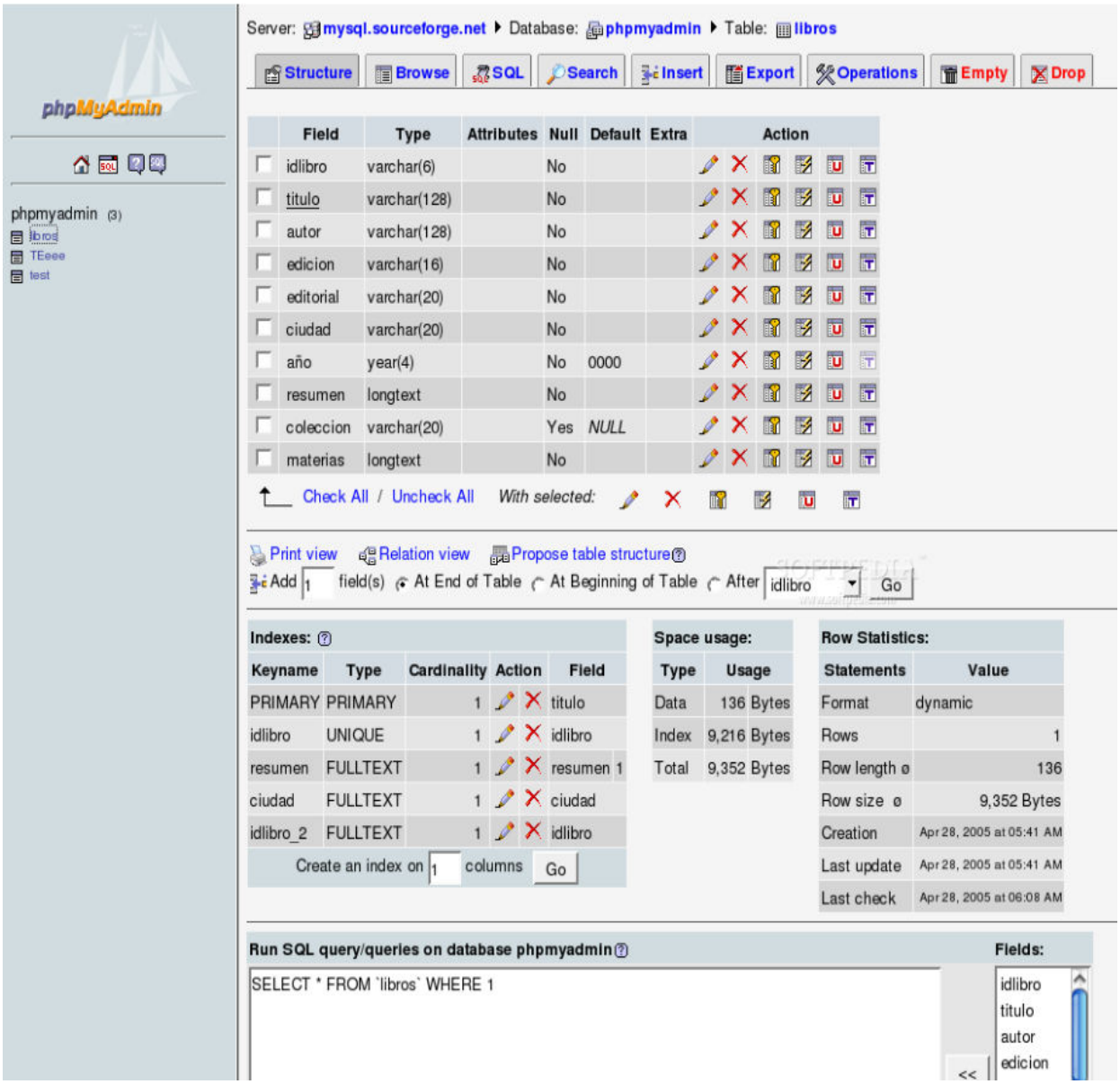
The purpose of the Database Management Module is to allow the user to manipulate the database. It is selected by pointing to the Database Management button on the User Activity Selection Menu and by pressing either the Submit key or the Enter key on the keyboard. The COTS database functions with the SWC program and when called will be presented to the user for use. All the capabilities provided by the COTS database product will be provided to the user.

During or upon completion of activity, the user can store the current task in the PC memory. When the work is complete the user closes the COTS database application. The system will prompt the user as to whether they wish to save the current work or close without saving and execute the selected action. The selection of the Exit button will return the user to the User Activity Selection Menu.

2.2.6.5.1 User Actions

- The user selects the Database Management Module from the User Activity Selection Menu.
- The module responds by returning the COTS database product.

FIGURE 2.11 SCREEN SHOT OF COTS DATABASE GUI



- The user utilizes the functions provided by the COTS database product to complete their desired tasks.
- During the session the user can save the current work periodically on any of the provided storage mediums.
- When the user completes their current task or wishes to exit the application, exiting the application is accomplished with the functions provided by COTS Database Application. When exit is initiated the program will ask the user if they wish to save their current work or exit without saving.
- The program executes the option selected by the user, terminates the current session, and returns the user to User Activity Selection Menu.

2.2.6.5.2 Functions Invoked

The COTS database application is invoked by the Initiate Web Crawler Module.

2.2.6.5.3 Errors and Resolution

There are no errors that are associated with this function.

2.2.6.5.4 Expected Results

The user is transferred to the COTS Database Application.

2.2.6.6 Related Information

The Database Management Module is accessed from the User Activity Selection Menu and provides access to the COTS database application.

2.2.7 Initiate Web Crawler

2.2.7.1 Scope

The Initiate Web Crawler Module provides the user the capability to crawl the Internet for information specified by the user. By setting the crawl parameters the user can crawl the Internet in search of the defined data. Once the Initiate Web Crawler Module is done with its crawl, the module calls the COTS Database Application to present the user with any matches found during the crawl.

2.2.7.2 Materials

Entry of the search parameters is required to initiate the web crawl.

2.2.7.3 Preparations

The System Administration Menu is used to access this function

2.2.7.4 Cautions and Warnings

Even though data entry is required by the user, there are no Cautions and Warnings that are associated with this function.

2.2.7.5 Method

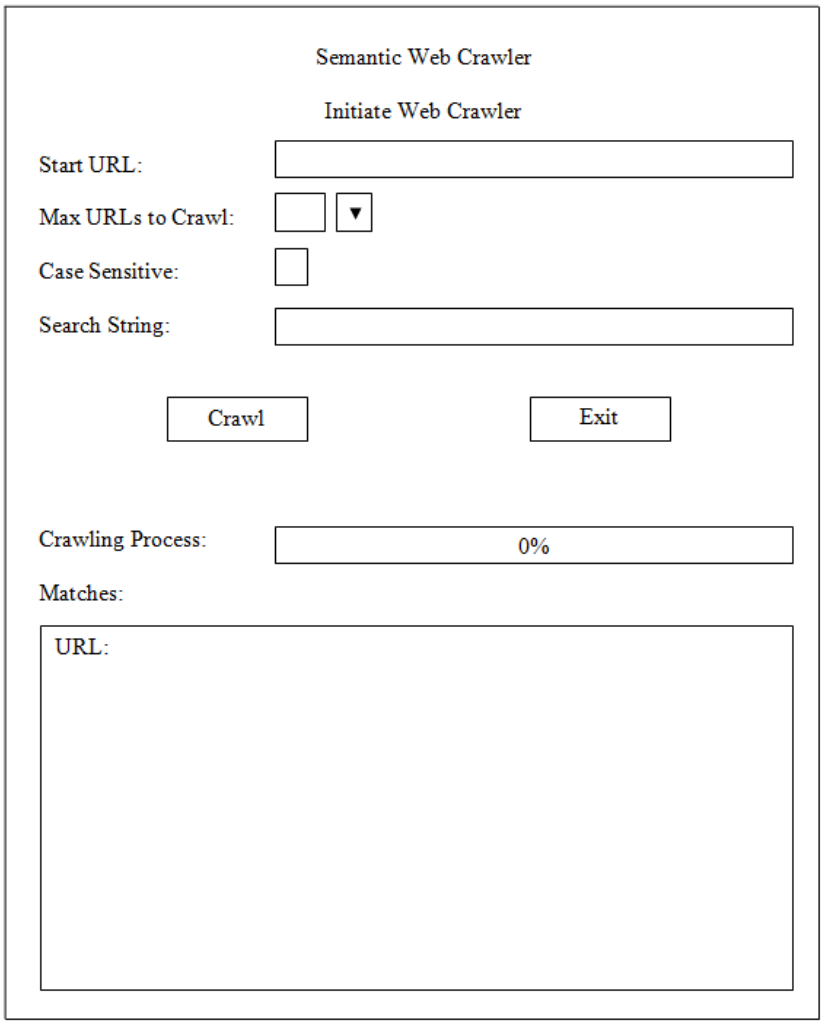
The Initiate Web Crawler GUI is accessed by navigating the SWC User Activity Selection Menu GUI. Through data entry the user can set the crawl parameters for the Internet crawl. The user initiates the crawl with the Submit button. When the Internet crawl is done, the Initiate Web Crawler Module calls the COTS Database Application to present the user with any matches

found. Using the functions from the COTS Database he user can store the current task in the PC memory during or upon completion of the activity. The user closes the COTS Database Application when the work is completed. The system shall prompt the user as to whether they want to save the current work or close the application without saving and execute the action selected. Exiting the COTS Database application returns the user to the Initiate Web Crawler GUI. The user can initiate another crawl or use the Exit button to return to the User Activity Selection Menu.

2.2.7.5.1 User Actions

- The user selects the Initiate Web Crawler Module from the User Activity Selection Menu.
- The module responds by returning the Initiate Web Crawler GUI.

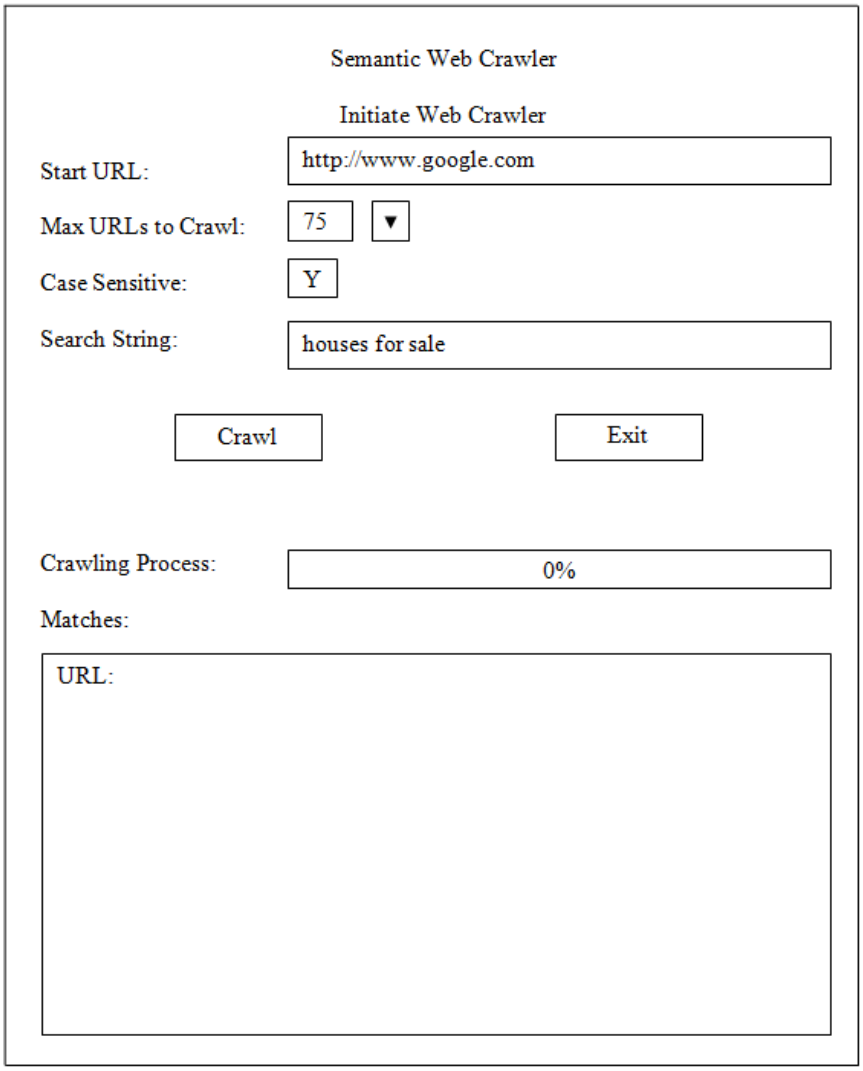
FIGURE 2.12 INITIATE WEB CRAWLER GUI



- Through data entry, the user sets the crawl parameters by entering the starting Uniform Resource Locator (URL), the number of URLs to crawl, and the search string.

- If the user wants to make the crawl case sensitive, then they enter Y, if not they leave it as is.

FIGURE 2.13 INITIATE WEB CRAWLER GUI WITH SEARCH PARAMETERS



- When the user is done defining the search parameters, they initiate the web crawl with Submit button or the Enter key on the keyboard.
- After the Internet crawl, the module responds by returning the COTS Database Application to present the user with any matches found.
- The user utilizes the functions provided by the COTS Database to complete their desired tasks.
- During the session the user can save the current work periodically on any of the provided storage mediums.
- When the user completes their current task or wishes to exit the application, exiting the application is accomplished with the functions provided by COTS Database Application. When exit is initiated the program will ask the user if they wish to save their current work or exit without saving.

- The program executes the option selected by the user, terminates the current session, and returns the user to User Activity Selection Menu.

2.2.7.5.2 Functions Invoked

The functions used for the search parameters will determine that the keyboard entries conform to the field limits. The COTS database application is invoked by the Initiate Web Crawler Module.

2.2.7.5.3 Errors and Resolution

For this function all the data field requiring keyboard entries are error checked. The critical field of the form that is checked is as follows:

- The Start URL field is checked for compliance with the field parameters to connect to the Internet.
- What if I receive an error message and one or more of the fields are highlighted?
    - The parameter of the data you entered is incorrect.
    - Verify that the entries are the correct type by consulting the HELP function.
- What if I receive an error message stating that the crawl could not be initiated?
    - There may be a problem the Internet connection, call your Internet Service Provider (ISP) and they will provide you information as to how to restore your Internet connection.

2.2.7.5.4 Expected Results

The user is presented with the matches found using the COTS Database Application.

2.2.7.6 Related Information

The Initiate Web Crawler Module is accessed from the User Activity Selection Menu and provides access to the COTS database application.

3.0 Error Conditions

The main error condition from the user's perspective could come in the case of improper arguments being given through the User Interface and directly into the SWC Core program. Since arguments are fed through command line input then an invalid or improperly built argument could confuse the core application and produce a substandard or not the desired crawl you originally wanted. This is actually built on purpose as it allows the users to make mistakes and get bad data giving them a chance to make the necessary changes to garner the desired crawl results.

3.1 Related Information

As the SWC Core is compiled and internal in its functions there are no real errors that can be fed to the user, just bad information from good information being crawled. The main related point

for the user is found in the User Interface, thus the UI is the only related function that applies to the user directly.

3.2 Error Messages, Known Problems and Error Recovery

There are no error messages listed for the User at this time, the main point of concern for the user is being able to list and or allocate what components from a website they wish to crawl. There is no need to understand regular expressions from the point of the end user as the User Interface handles the building out of what is necessary to activate the Core to crawl the site.

Appendix A

Bibliography

Bass, Clements, Kazman, Software Architecture in Practice, 2nd Ed., April 2003

Bastos, M., Covarrubias, M., Patel, D., Semantic Web Crawler Operational Concept Document March 2012

Bastos, M., Covarrubias, M., Patel, D., Semantic Web Crawler Software Project Management Project Document March 2012

Bastos, M., Covarrubias, M., Patel, D., Semantic Web Crawler Software Requirements Specification
Document April 2012

Pressman, Roger S., Project Engineering: A Practitioner's Approach, Seventh Edition, McGraw Hill, 2010.

Appendix B

Glossary & Definitions

Cache	A location used to store data in order to have faster subsequent retrievals
Cloud Based Environment	Refers to applications and services offered over the Internet. Data Centers that offer services are referred to as the 'cloud'
Cloud S3	A simple storage service
De-couple	Two or more systems that are able to transact without being connected (coupled). The systems do not interact with each other and a decoupled system allows changes to be made to any one system without having an effect on any other system
Debug	The process of locating and fixing or bypassing bugs (errors) in a computer program code or the engineering of a hardware device.
DynamoDB	A fully managed NoSQL database service that provides fast and predictable performance with seamless scalability
Encapsulate	The process of combining elements to create a new entity
Functional Model	A structured representation of the functions within the modeled system. Functions include activities, actions, processes, and operations
Hash	A function that maps large sets of data into smaller sets of data
Hash Value	A number generated from a string of text
Host Computer	A computer that is connected to a TCP/IP network, including the Internet, and has a unique IP address
NoSQL	A way of storing and retrieving data quickly, like a relational database except it isn't based on the mathematical relationship between tables as a traditional relational database does
Object Model	Description of an object-oriented architecture, which includes the details of the object structure, interfaces between objects, and other object-oriented features and functions
Polymorphism	A programming language's ability to process objects differently depending on their data type or class

Proxy	An online computer server that acts as an intermediary between an Internet user and their destination site
Sockets	Fundamental technology for programming software to communicate on TCP/IP networks. Sockets provide a bidirectional communication endpoint for sending and receiving data with another socket
SQL	A simple programming language used for accessing and managing data in relational databases such as SQL Server
URL	The global address of documents and other resources on the World Wide Web

Appendix C

Abbreviations & Acronyms

COTS	Commercial Off-The-Shelf
CSC	Computer Software Components
CSCI	Computer System Component Interfaces
CSU	Computer Software Units
CSV	Comma Separated Values
DNS	Domain Name Service
DSL	Digital Subscriber Line
ECO	Engineering Change Order
GB	Gigabytes
GHz	Gigahertz
GUI	Graphical User Interface
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol with Secure Sockets Layer
I/O	Input/Output
IP	Internet Protocol
ISP	Internet Service Providers
IT	Information Technology
KLOC	Thousand Lines of Code
MTBF	Mean time between failure
MTTR	Mean time to repair
OOA	Object Oriented Architecture
OS	Operating System
RAM	Random Access Memory
Regex	Regular expression
SDD	System Design Document
SQL	Structured Query Language
SRS	Software Requirements Specification
SSL	Secure Sockets Layer
SWC	Semantic Web Crawler
TCP	Transmission Control Protocol
UI	User Interface
URL	Uniform Resource Locator