Software Requirements Specification

for

<Taxonomic Hierarchy Comparator>

Version 0.1.0

Prepared by <Colin>

<2013/7/10>

Table of Contents

Revision History

Name	Date	Reason For Changes	Version

1. Introduction

1.1 Purpose

This document is for THC 0.1.0 (Taxonomic Hierarchy Comparator 0.1.0) which is to be a tool based on web. It is a new independent system intended for one of the targets proposed in 2013 EOL Rubenstein Fellow program that "using the multiple classifications harvested by EOL for analysis to obtain the degree of coverage and congruence among hierarchies and nomenclatures." This SRS summarizes the detail requirements of THC according to our previous survey and some experiences in other relative projects like Catalogue of Life China.

1.2 Document Conventions

No special conventions.

1.3 Intended Audience and Reading Suggestions

Developers and testers are the intended audiences. Developers should read all the sections of this document in detail; testers should focus on the sections about system features and user interface.

1.4 Project Scope

Taxonomic Hierarchy Comparator is intended for analyzing various biological taxonomic hierarchies from different sources. The targets of the project are as follows:

a. Propose a reliable method for comparing taxonomic hierarchies.
b. Implement a mature tool basing on comparison method for classifications comparison and visualization.
c. Find out the differences among classifications provided by EOL or from other sources, and propose quantitative indexes for measuring the degree of overlap and congruence among them.
d. Try to explore a possible method for taxonomists to mine differences of taxonomic views and find out potential taxonomic or nomenclatural acts.

1.5 References

1. Colin, 2012, Project Description in Application Documents for EOL 2013 Rubenstein Fellows.
2. Congtian Lin, Huijie Qiao, Jiangning Wang, Liqiang Ji*, 2012. Taxonomic Tree Tool for Managing and Comparing Taxonomic Trees (Abstract). 2012 TDWG Conference in China.

2. Overall Description

2.1 Product Perspective

THC is a new independent product initiated firstly and supported by EOL. The architecture of THC is shown in diagram 1.

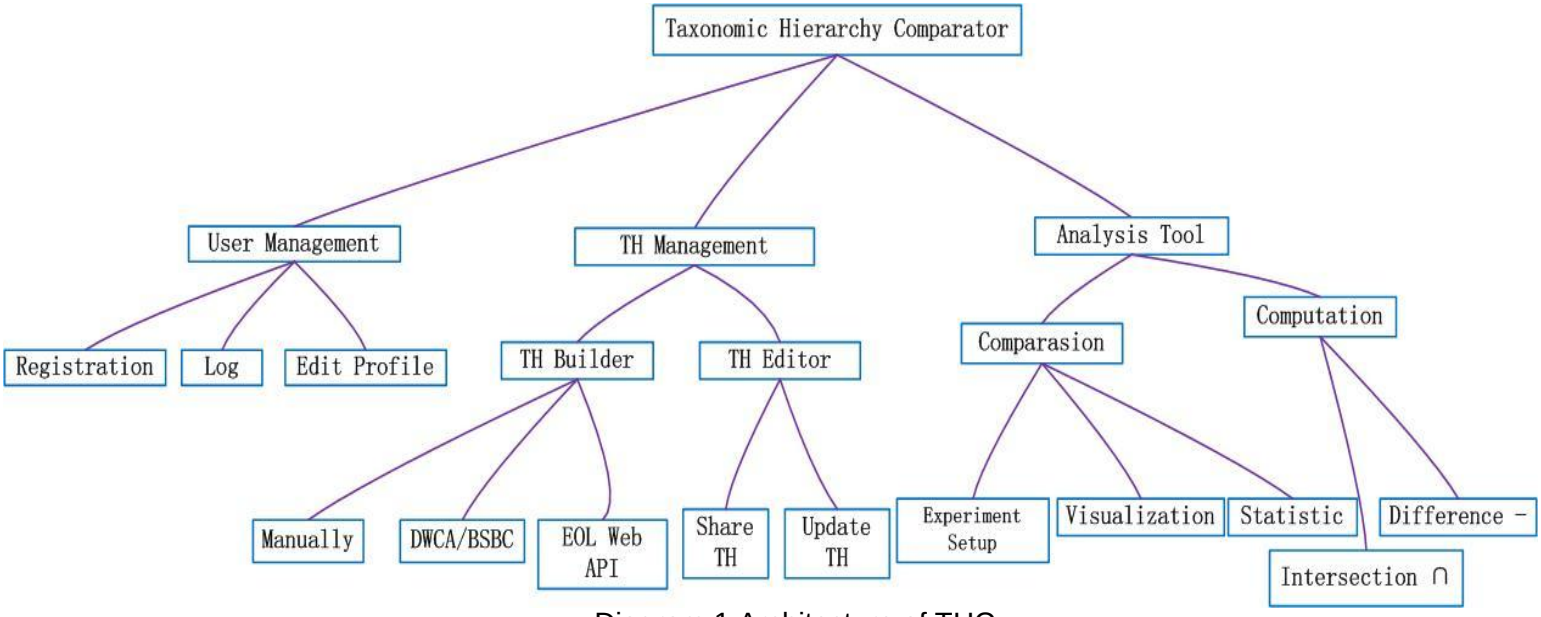

Diagram 1 Architecture of THC

2.2 Product Features

This section describes the major features of THC and provides some diagrams of major groups of related requirements.

(1) User management: diagram 2
(2) Taxonomic Hierarchy (or tree) management: diagram 3
(3) Experiment management: diagram 4

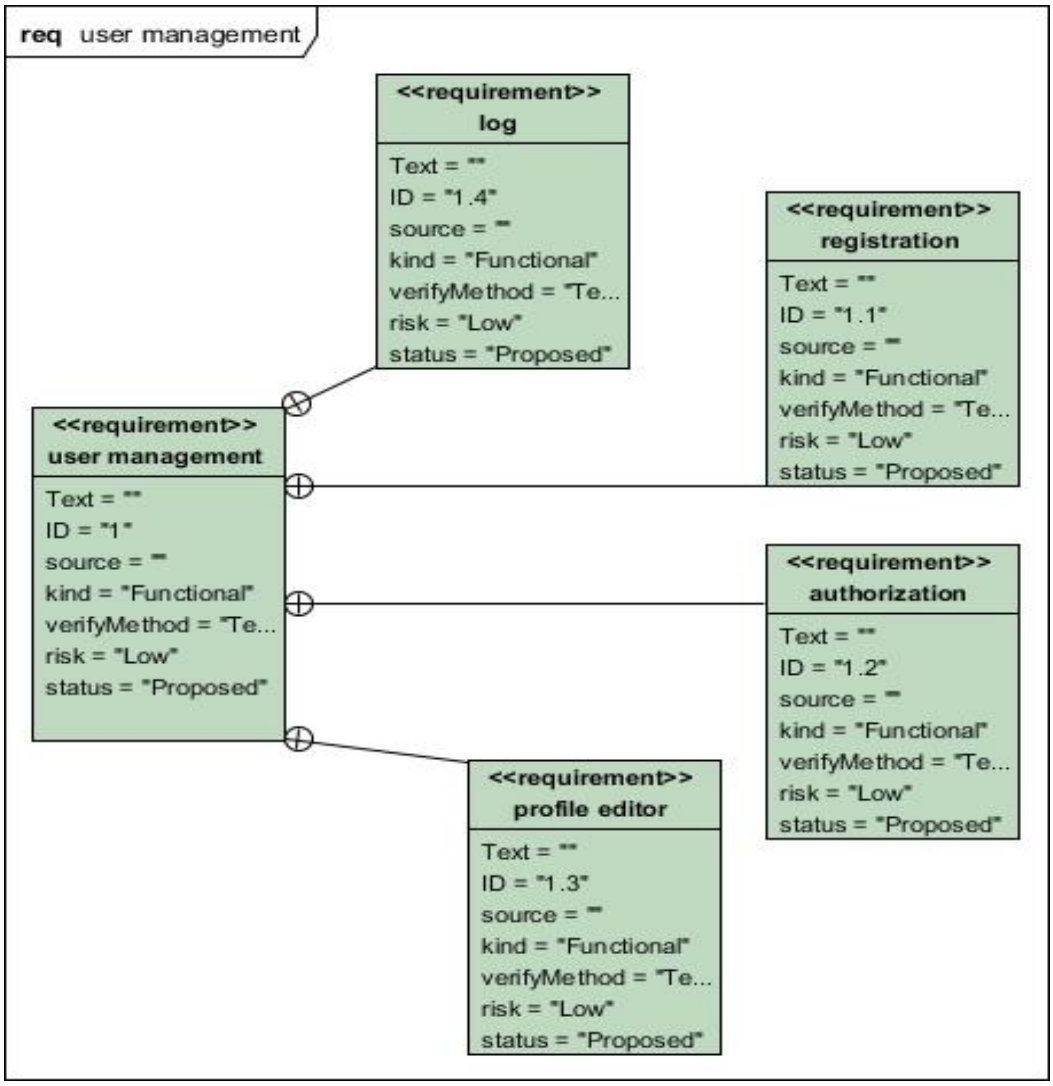

Diagram 2 Requirements of User Management

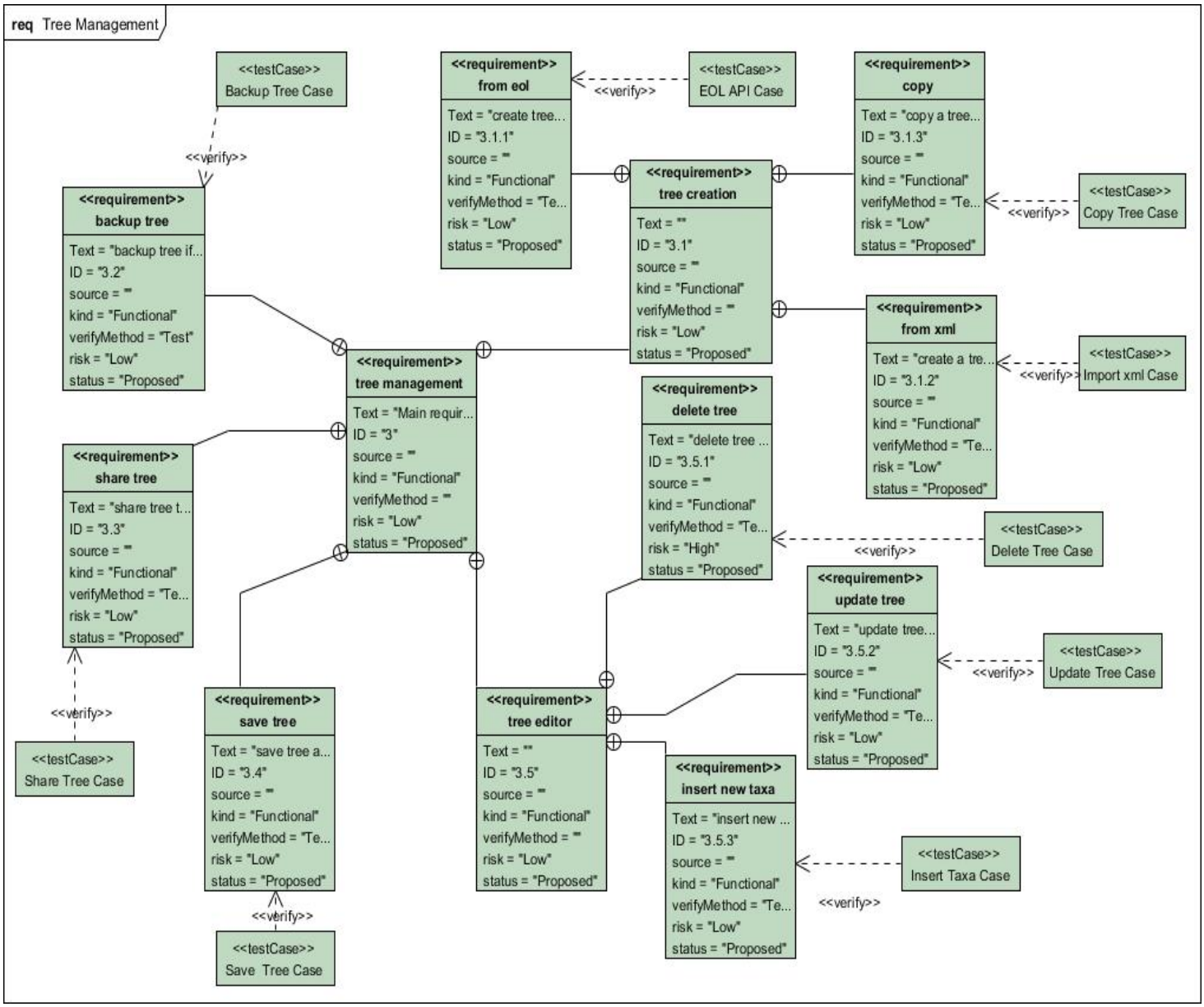

Diagram 3 Requirements of Tree Management

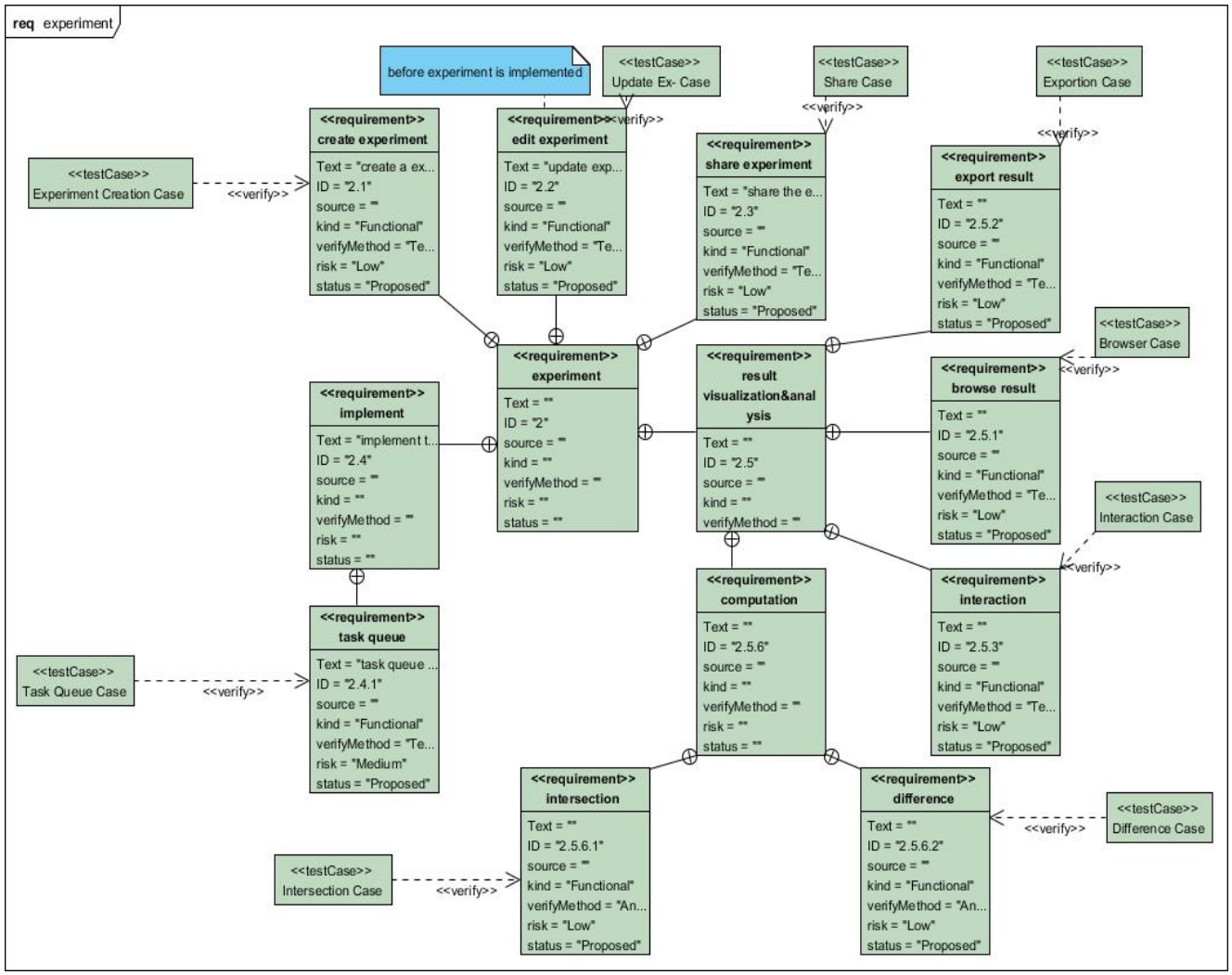

Diagram 4 Requirements of Experiment

2.3 User Classes and Characteristics

(1) Taxonomists my prefer the functions of comparison, visualization and tree editor that help them to find out the difference of views on biological classifications and revise their own tree.
(2) Biodiversity Informationists mostly concern the congruence and incongruence among the taxonomic hierarchies to find out what information is absent and what species are covered. They may prefer the functions of comparison, computation, result statistic, and result export.

All functions can be used by both potential user groups.

2.4 Operating Environment

Hardware: Lenovo Server, Intel(R) Xeon (R) CPU, 2.40GH, 12.0GB RAM
Operation system: Windows sever 2008 R2
Running environment: jdk6 or above, Tomcat

2.5 Design and Implementation Constraints

(1) Comparison and computation are time-consuming process, so the high performance should be considered when design and code.
(2) Some hierarchies are of large nodes that bring into the problem big RAM requirement.
(3) The database software should be mysql5.0 or above version.
(4) About the language, at least Chinese and English are required.

2.6 User Documentation

User manual in PDF and tutorials in PDF or video

2.7 Assumptions and Dependencies

(1) DWCA tools from GBIF for reading and validating DWCA files
(2) EOL Web API for retrieve taxonomic hierarchies.

3. Use Cases

This section is to describe the detail use cases and their relationship. The content has been arranged into another document named "Use Cases Specification for THC".

4. External Interface Requirements

4.1 User Interfaces

Refer to "user interface specification for THC".

4.2 Hardware Interfaces

No special interfaces are required.

4.3 Software Interfaces

(1) Mysql database, version 5.0 or above: THC read data from mysql and save data to mysql through JDBC API. Refer to the design document of the database for the data items.
(2) EOL web site, version 2.0: THC retrieves taxonomic hierarchies from the EOL web site through the web API.
(3) Common use of other software interfaces like operating systems, tomcat server and so on.

4.4 Communications Interfaces

THC is a web based tool and mainly use HTTP as communication standard.

5. Other Nonfunctional Requirements

5.1 Performance Requirements

Comparison is a time consuming process, so it is hard to implement many comparison tasks at the same time. It is necessary to utilize task queue and multiple tasks function to realize it.

5.2 Safety Requirements

(1) Use Redundant Arrays of Inexpensive Disks (RAID) as disk environment.
(2) Back up data every week regularly.

5.3 Security Requirements

(1) All the data retrieved from EOL web site will be published to public.
(2) The data sets created by user is own only to the user, and the owner have the right to share them or not.
(3) To prevent malicious task submission, manager has the right to stop any malicious task queue.