Lehrstuhl fuer Systeme und Informationsverwaltung

Prof.Dr.Ing. Klemens Boehm
Hoang Vu Nguyen
Marco Neumann

Praxis der Softwareentwicklung (PSE)
WS 2012/2013

Visualizing and Statistically Analyzing Access Behavior to Scientific Databases

Functional Specification

Hier koennte unser Logo oder Bild hinkommen. Die Sectionsueberschriften und Subsections bekommen die Tage noch n schoeneren Look.

November 11, 2012

Functional Specification **Alexander Noe**
Design **Jonathan Klawitter**
Implementation **Anas Saber**
QA / Testing **Niko Moraitakis**
Final **Lukas Ehnle**

Contents

1 Introduction

This is the functional specification for Group 14 in the PSE assignment 10 in winter semester 2012/13. In this project we have to analyze and visualize the serverlog of the Skyserver, which is a database for astronomical data. For this assignment we are going to write two separate applications. The first application is called Server Log Parser (SLP) and is going to convert the CSV-formatted serverlog from Skyserver into a data-warehouse. The second application works on this data-warehouse and allows it's users to create various charts, for example scatterplots and histograms. Due to the fact that we split the assignment into two separate applications, many parts of this specification are going to be splitted in two parts, one for each of those.

We have to state that no one of us has much experience in working with databases and javascript. Therefore we can't guarantee for the correctness of all information stated in this specification. It could happen that we will alter some of the specifications when designing or implementing the actual program.

2 Goals

2.1 Criteria

2.1.1 Server Log Parser (SLP)

- The SLP can read a CSV-formatted log file from sdss.skyserver.org and fill a data-warehouse with multi-dimensional user information. The dimensions of this data are location of the user, location of the database and time of the server-access.
- SLP will recognise invalid logs and won't add them to the data-warehouse. Every log with a mistake won't be accepted, because an error-message is not as bad as a corrupted data-warehouse.

2.1.2 Warehouse Analyzing Tool (WHAT)

- WHAT can use the data-warehouse which is filled with user data from sdss.skyverser.org to create various charts. Those charts are created by user command and are visible on a javascript-webpage.
- WHAT can create scatterplots and histograms from data in the warehouse.

2.2 Optional criteria

2.3 Exclusion criteria

2.3.1 Server Log Parser (SLP)

- The SLP can only use CSV-formatted logs from sdss.skyserver.org. It can neither read logs in another format nor logs from another source.
- There is no way to avoid using SLP when adding data to the warehouse. This can stop corrupting the warehouse to guarantee correct data in the warehouse.
- SLP doesn't correct mistakes in the log file.

2.3.2 Warehouse Analyzing Tool (WHAT)

- WHAT takes correct data in the warehouse for granted. It doesn't check the data because it already checked by SLP.

3 Usage

3.1 Applications

3.2 Target groups

3.3 Operating conditions

4 Operating environment

4.1 Software

4.2 Hardware

4.3 Orgware

4.4 Product interfaces

5 Functional requirements

Here we go with a nice enumeration list with a (cheated) 10 steps counter.

5.1 First Whatever

/F10/ blabla

/F20/ tadaa

5.2 Next Whatever

/F30/ 30!

/F40/ 40!

6 Data

6.1 First Whatever

/D10/ blabla

/D20/ tadaa

6.2 Next Whatever

/D30/ 30!

/D40/ 40!

7 Nonfunctional requirements

7.1 First Whatever

/NF10/ blabla

/NF20/ tadaa

7.2 Next Whatever

/NF30/ 30!

/NF40/ 40!

8 Global test cases

9 Models

9.1 User stories

9.2 Object models

9.3 Dynamic models

9.4 Web interfaces

10 Glossary