

An Automation based approach to improve CW Data Management

Submitted By

Priya John

14MCEI18



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2016

An Automation based approach to improve CW Data Management

Major Project

Submitted in partial fulfillment of the requirements

for the degree of

Master of Technology in Computer Science and Engineering

Submitted By

Priya John

(14MCEI18)

Guided By

Prof. Pimal Khanpara



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2016

Certificate

This is to certify that the major project entitled “**An Automation based approach to improve CW Data Management**” submitted by **Priya John (Roll No: 14MCEI18)**, towards the partial fulfillment of the requirements for the award of degree of Master of Technology in Computer Science and Engineering of Nirma University, Ahmedabad, is the record of work carried out by her under my supervision and guidance. In my opinion, the submitted work has reached a level required for being accepted for examination. The results embodied in this major project part-I, to the best of my knowledge, haven't been submitted to any other university or institution for award of any degree or diploma.

Prof. Pimal Khanpara
Guide & Assistant Professor,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

Dr. Sharada Valiveti
Associate Professor,
Coordinator M.Tech - CSE INS
Institute of Technology,
Nirma University, Ahmedabad

Dr. Sanjay Garg
Professor and Head,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

Dr. P. N. Tekwani
Director,
Institute of Technology,
Nirma University, Ahmedabad

Certificate

This to certify that **Priya John (14MCEI18)**, Institute of Technology, Nirma University, Ahmedabad is working in this organization since 17/06/2015 and carried out her thesis work titled “**An Automation based approach to improve CW Data Management**”. She was working as a graduate intern under our guidance. She is allowed to submit her dissertation report. The results embodied in this project, to the best of our knowledge, haven't been submitted to any other university or institution for the award of any degree or diploma. We wish her success in future endeavours.

Amit C Zala
Guide & Software Engineer,
MIG SPO BASE ,
Intel Technology India Pvt. Ltd.,

Giridhar Bajpe
Sr. Engg. Program & Operations Manager,
MIG SPO BASE
Intel Technology India Pvt. Ltd.

Statement of Originality

I, **Priya John**, Roll. No. **14MCEI18**, give undertaking that the Major Project entitled “**An Automation based approach to improve CW Data Management**” submitted by me, towards the partial fulfillment of the requirements for the degree of Master of Technology in **Computer Science & Engineering** of Institute of Technology, Nirma University, Ahmedabad, contains no material that has been awarded for any degree or diploma in any university or school in any territory to the best of my knowledge. It is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. It contains no material that is previously published or written, except where reference has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

Signature of Student

Date:

Place:

Endorsed by
Prof. Pimal Khanpara
(Signature of Guide)

Acknowledgements

I extend my gratitude to my manager **Giridhar Bajpe**, Intel for his valuable insights during the course of the project. Also, my mentor **Sangharsh Dubey** and colleagues **Amit Zala, Deebiga B., Manvendra Singh, Madhu Namboodiri, Ravi Sandesh, Shuiab Sadiq** and **Mohit Malhotra** at Intel for teaching me skills that are pivotal for a good software engineer.

I would like to thank **Prof. Pimal Khanpara**, Assistant Professor, Computer Science Department, Institute of Technology, Nirma University, Ahmedabad for her valuable guidance throughout this work.

It gives me an immense pleasure to thank **Dr. Sanjay Garg**, Hon'ble Head of Computer Science and Engineering Department, Institute of Technology, Nirma University, Ahmedabad for providing state-of-the-art infrastructure for research and development.

For his support throughout the project, I would like to thank **Dr P. N. Tekwani**, Hon'ble Director, Institute of Technology, Nirma University, Ahmedabad. Also, I would like to thank the Institution, all faculty members of Computer Engineering Department, Nirma University, Ahmedabad for their special attention towards the project work.

Thanks to my family for being my pillar of strength. Special thank you to all my colleagues and friends for sharing their suggestions that have improved the project.

- **Priya John**
14MCEI18

Abstract

We live in the information age. Data is increasing in all dimensions – velocity, variety and volume. Storing and analysing big data continues to be a topic for further research. Current topic deals with the efficient management of data of contingent workforce. Since multiple databases are maintained, it becomes very cumbersome to consolidate data for meaningful interpretation. The purpose of the project is to eliminate this problem of data integration and duplication by providing a one-stop solution with the help of data mining techniques and latest technologies.

Abbreviations

CW	Contingent Worker / Workforce.
ASP	Active Server Pages.
SQL	Structured Query Language.
WiFi	Wireless Fidelity.
RTT	Round Trip Time.
OLAP	Online Analytical Processing.
IoT	Internet of Things.
UI	User Interface.

Contents

Certificate	iii
Certificate	iv
Statement of Originality	v
Acknowledgements	vi
Abstract	vii
Abbreviations	viii
List of Tables	xi
List of Figures	xii
1 Introduction	1
1.1 Company Introduction	1
1.1.1 Team Introduction	1
1.2 Problem Statement	1
1.3 Proposed Solution	2
1.4 Objective	2
1.5 Scope Of Work	3
1.6 Business Need	3
1.7 Features	4
2 Literature Survey	5
3 Design and Implementation	7
3.1 Languages and Frameworks	7
3.1.1 HTML	7
3.1.2 AngularJs	7
3.1.3 Bootstrap	7
3.1.4 ASP.NET C#	7
3.1.5 Web Handler	8
3.1.6 AJAX	8
3.1.7 High Charts	8
3.1.8 SQL	9
3.1.9 MVC Architecture	9
3.2 System Architecture	9

3.3	Complex inner join queries	10
3.4	Database Design	10
3.5	Snapshots	11
4	Related Work	14
4.1	Optimization	14
4.1.1	Web Page Performance	14
4.1.2	Network utilization	15
4.1.3	Memory utilization	16
4.2	Knowledge Discovery Process	17
5	Further Work	18
5.1	EC Efficiency Initiatives	18
5.1.1	Objective	18
5.1.2	Problem Statement	18
5.1.3	Proposed solution	18
5.1.4	Implementation	19
5.2	BASE Website	22
5.2.1	Objective	22
5.2.2	Problem Statement	22
5.2.3	Proposed Solution	23
5.2.4	Implementation	23
5.2.5	Snapshots	24
6	Conclusion	26
6.1	Conclusion	26
	Bibliography	27

List of Tables

2.1	Table - Summary of Literature Survey	6
5.1	Comparison of results before and after query optimization	20

List of Figures

3.1	Ajax Call	8
3.2	MVC Architecture	9
3.3	System Architecture	10
3.4	Complex inner join queries	10
3.5	Table Interaction Diagram	11
3.6	Snapshot of main interface page	12
3.7	Snapshot of one of the modals used to enter CW data	13
4.1	Web performance graph - before	15
4.2	Web Performance graph - later	15
4.3	Network utilization graph - before	15
4.4	Network utilization graph - later	16
4.5	Memory utilization graph - before	16
4.6	Memory utilization graph - later	16
4.7	KDD process application in project development stage	17
5.1	Query Execution plan	21
5.2	Team's portfolio where the names of tools and description have been replaced with other text	24
5.3	Interface to display dashboards to keep track on the status of requests	25

Chapter 1

Introduction

1.1 Company Introduction

Intel (Integrated Electronics) is one of the biggest semiconductor companies in the world. It continues to be a pioneer in the field of high end processors. Intel is engaged in not only making semiconductor chips, but also in extending the limits of technology to enrich lives. In addition to 2-in-1 laptops, Intel is also involved in the field of wearables, drones, IoT, high speed computing, big data analytics, cognitive learning and security [1].

1.1.1 Team Introduction

IPG MIG Business Automation teams is engaged in creating enterprise web applications. Projects are undertaken with the aim of simplifying the otherwise complex and tedious business processes. The manual steps that are to be done repeatedly can be automated so as to enable the actions at the click of a single button. A wide range of technologies are used to implement solutions for the identified problems. The team works closely with the client to understand the requirements in depth and create applications that lead to customer delight.

1.2 Problem Statement

- Data regarding contingent workers resides in multiple data sources.
- Updating data related to a single CW at multiple locations becomes a tedious process.
- Duplicated data same record in multiple databases inefficient memory utilization -

problem.

- Reports are generated manually.
- No automation in the process of report generation in the current scenario.
- While searching, results are obtained from multiple database sources.
- Information is to be searched, interpreted and aggregated manually more effort and time.

1.3 Proposed Solution

- Create a single data source for CW data so that all the data is available at one location.
- Algorithms for faster and efficient data access to eliminate data duplication.
- Implement access control mechanisms to prevent unauthorized users from viewing or modifying data.
- Provide a reporting solution for comparison of current data sets and getting futuristic estimations.
- Prepare a single interface to enable viewing aggregated data regarding contingent workers.
- Provide a solution so that it is easy to search for CW based on search terms.
- Maintain logs for audit.

1.4 Objective

The main objective of the project is to develop and maintain a unified source of information for CW data. In future, when CW data is required, it can be fetched easily. The solution should be able to generate reports and intuitive graphs for evaluation of data sets. The results obtained should be highly interactive and easy to comprehend for the user. The access to the project should be given only to a few selected users and unauthorized users should be restricted from using the system. If an error occurs, the system should notify the user about the same.

1.5 Scope Of Work

The following points clearly define the scope of work pertaining to the project.

- Learning latest technologies
- Designing attractive interfaces
- Generating reports
- Generating automated indicators
- Writing scripts for automation
- Normalizing tables
- Removing data redundancy
- Implementing access control

The following points are beyond the scope of the project.

- Tracking CW contribution in project
- Evaluation of CW performance
- Scheduling programs for CW skill development
- Monitoring CW activity

1.6 Business Need

The following are the requirements for which the solution is proposed:

- Reduce data duplication
- Automate the process of updating relevant records by cascading changes in CW data
- Data consistency and integrity
- Reduce manual effort on maintenance of CW data

1.7 Features

- The CW data that is maintained in different sources like Excel, HAT, CDIS and External tools are consolidated at one place.
- It will help in futuristic budget estimation for groups whose projects involves the participation of contingent workers.
- It will enable easy comparison of budget allocation for a particular group and actual spending
- Reporting solution to represent data in the form of charts.
- Implementation of access control mechanisms for better security and prevention of unauthorized access.
- Implementation of techniques to optimize page load time to enhance user experience and reduce waiting time.
- Logging of user actions to avoid non-repudiation.
- Customized exception handling to user so as to not to overwhelm the user with excess technical information.

Chapter 2

Literature Survey

The following table is a synopsis of the research papers studied during the project. Each of these papers present a thought provoking perspective on visualization of data.

Title	Publication and Year	Author/s	Key Learnings
Information Visualization and Visual Data Mining	IEEE Transactions on Visualization and Computer Graphics(2002)	D. A. Keim	Data types to be visualized: 2-D and m-D, Visualization techniques : interactive filtering [2].
Data, Information and Knowledge in Visualization	IEEE Computer Graphics and Applications(2009)	M. Chen, D. Ebert, et. al.	Information about results : color, Histogram [3].
Research on Visualization techniques in data mining	Computational Intelligence and Software Engineering (2009)	H. Liu, H. Jin	Visualization technologies: Image maps, Pixel mapping value to pixel, and Projection scatter plots [4].

Adaptive Privacy- Preserving Visualiza- tion Using Parallel Coordinates	IEEE Trans- actions on Visualization and Computer Graphics(2009)	A. Dasgupta, R. Kosara	Privacy preserving in- formation visualization on the basis of screen space metrics control- ling attributes to control information displayed [5].
Data Mining Information Visualiza- tion Beyond Charts and Graphs	Proceeding of the Sixth Interna- tional Conference on Information Visualization	N. Robinson, M. Shapcott	Virtual Data Mining envi- ronment data equivalent to liquid filtered and purified in ID3 algorithm [6].

Table 2.1: Table - Summary of Literature Survey

Chapter 3

Design and Implementation

3.1 Languages and Frameworks

3.1.1 HTML

The most basic language used in web development is HTML or Hyper Text Markup Language. HTML5 along with CSS3 has been used to design the front end of the project.

3.1.2 AngularJs

For versatile display in web applications, AngularJS is used in handling of ever changing content. It is a comprehensive toolkit consisting of varied functionalities. It was designed to support the viewing of dynamic content.

3.1.3 Bootstrap

It is a combination of HTML, CSS and Javascript framework [7]. Used for front end, it scales the applications as per the device from mobile phones to tablets and laptop to personal computers. It has a range of visually appealing form controls to layouts and positioning. It has a wide range of themes that can be implemented in the project by invoking the name in the class attribute.

3.1.4 ASP.NET C#

C# is a programming language used to make web application that can be executed on .NET frame-work [8]. Being an object-oriented language, it supports features such as inheritance, polymorphism and encapsulation. The .NET framework used in the project

is 4.5.

3.1.5 Web Handler

An ASP.NET HTTP handler is the procedure (habitually alluded to as the "endpoint") that keeps running because of a solicitation made to an ASP.NET Web application [9]. The most widely recognized handler is an ASP.NET page handler that procedures .aspx records. At the point when clients ask for an .aspx record, the solicitation is prepared by the page through the page handler.

3.1.6 AJAX

Asynchronous Javascript and XML is used to load a section of the page again without the entire page being loaded once again [10]. This aids in creation of faster, responsive and more efficient web pages. AJAX calls have been used in the project to carry out update operations. The limitation of AJAX call is that in case of failure from server side, it will take too much time to update leading to failure in execution.

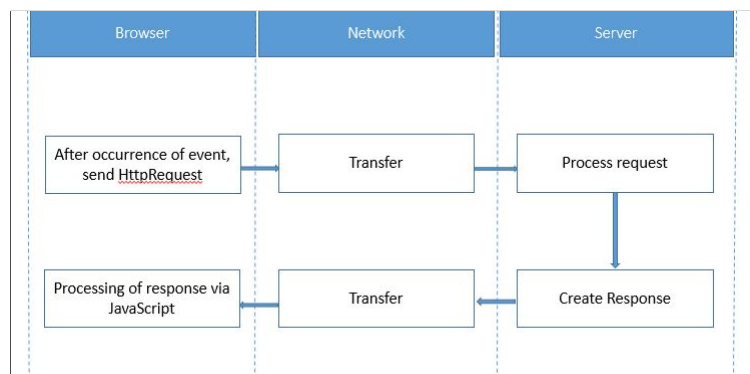


Figure 3.1: Ajax Call

3.1.7 High Charts

It is another popular framework used for charts to engage the user [11]. It offers a wide range of options of charts such as line, bar, pie and area being few of them. Label, captions, legends, drill down, tool tip are some of the features that are inherently offered by High-Charts. A wide range of themes, colours and styles are available that can be used while implementing HighCharts in the project.

3.1.8 SQL

Microsoft Structured Query Language (SQL) Server has been used for the purpose of implementation of database. Data manipulation and database definition operations have been performed using SELECT, INSERT, UPDATE and INNER JOIN. The tables have been normalized so as to avoid data duplication. The data from SQL tables is bound to the Angular UI grid in such a way that if an edit operation is carried out, the same is reflected from the database.

3.1.9 MVC Architecture

The MVC example offers assistance with creating applications that different the diverse parts of the application (information rationale, business rationale, and UI rationale), while giving a free coupling between these components [12]. The example determines where every sort of rationale ought to be situated in the application. The UI rationale has a place in the perspective. Information rationale has a place in the controller. Business rationale has a place in the model. This division offers assistance with managing many-sided quality when you manufacture an application, on the grounds that it empowers you to concentrate on one part of the execution at once. The free coupling between the three principle parts of a MVC application likewise advances parallel improvement.

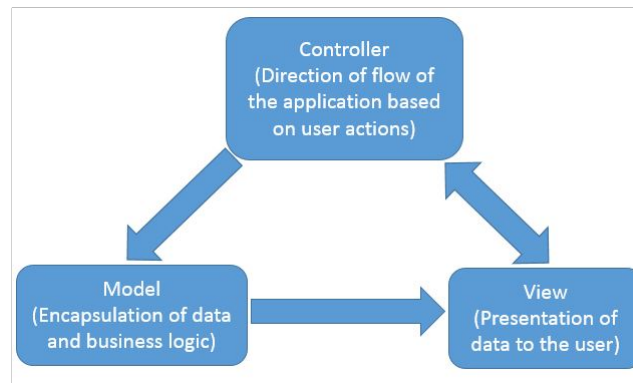


Figure 3.2: MVC Architecture

3.2 System Architecture

The following diagrams shows the interaction of the system components with each other. The CWT database is consolidated from different sources like EXCEL worksheets and other Intel databases. The data is fetched from the database and displayed using Angu-

larJs grid. The data can be viewed, updated and corresponding changes are reflected on the database.

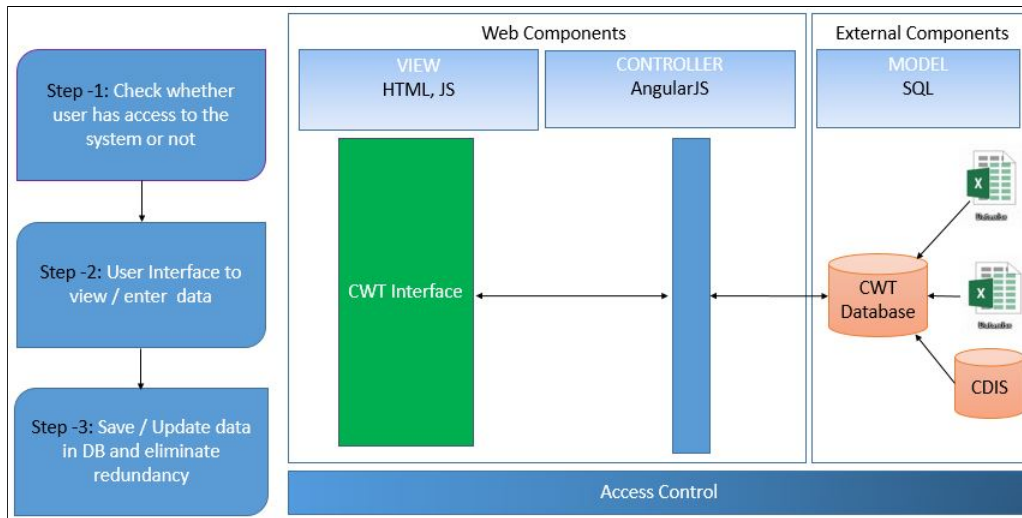


Figure 3.3: System Architecture

3.3 Complex inner join queries

The inner join queries used to display the data in the interfaces involves multiple tables. Further optimizing the queries by adding indexes is to be included as a part of further work.

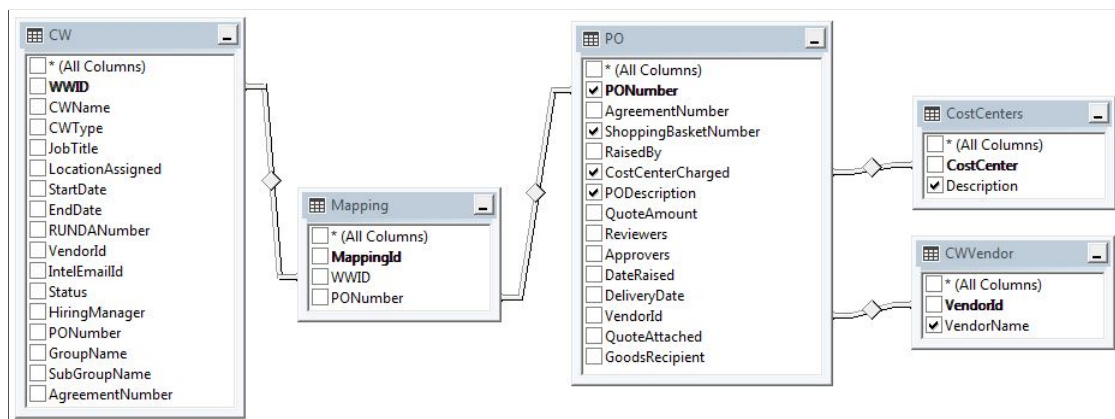


Figure 3.4: Complex inner join queries

3.4 Database Design

The following diagram depicts the relationship between the tables. The relationship can be 1:1, 1:n or m:n. The m:n relationship tables have been normalized so as to avoid data duplication and anomalies while insert or update operations are carried out.

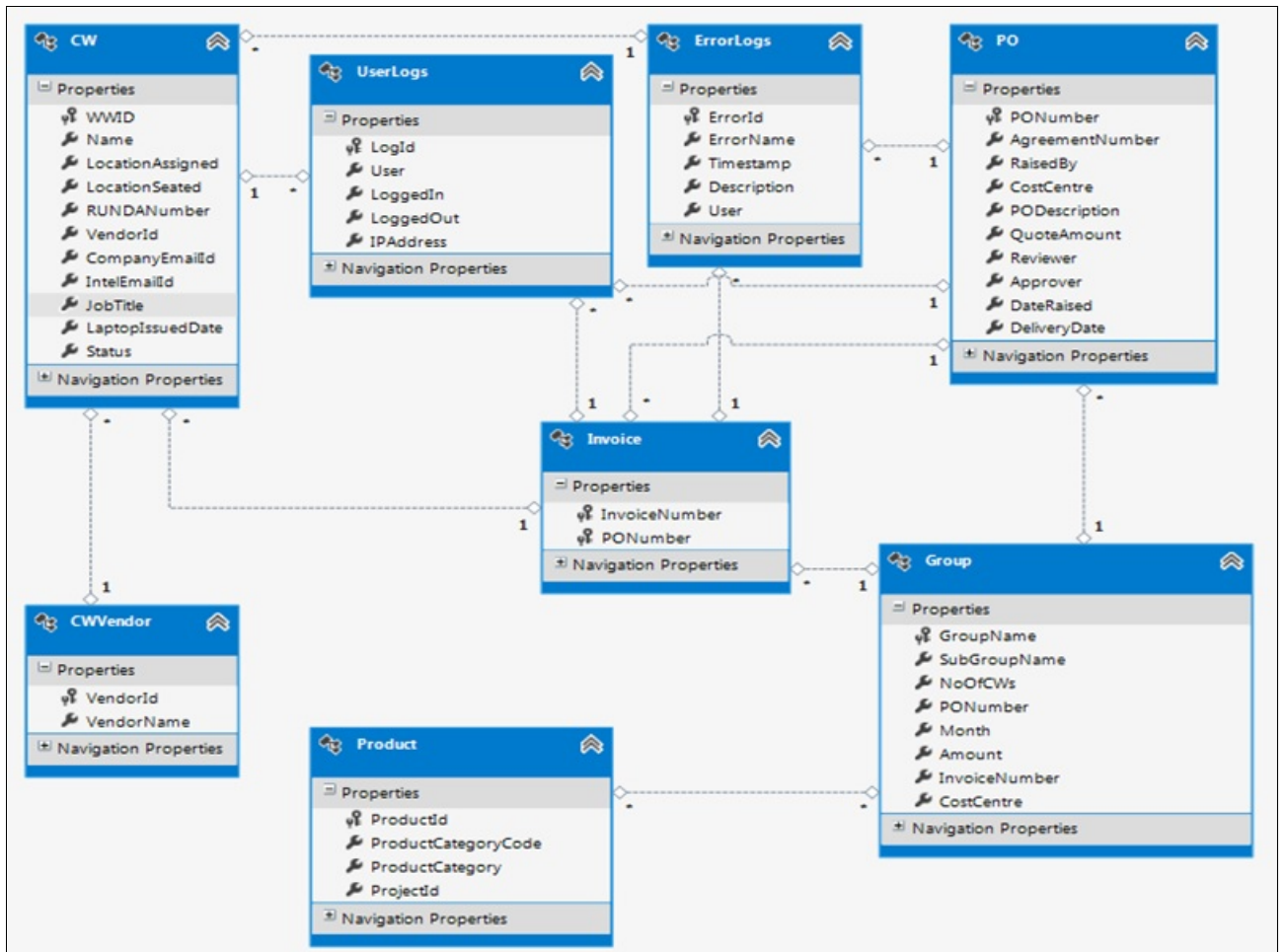


Figure 3.5: Table Interaction Diagram

3.5 Snapshots

The following are the snapshots taken from the CW Tracker tool. The first snapshot is the CW Main interface in which the entire set of data pertaining to a CW can be viewed at a single place. By clicking on View Details button, the PO and invoice details will be displayed in pop-up windows. Similarly, edit operations can be performed so as to update the information. The second snapshot depicts the interface to enter CW data in the modal.

Add new CW

WWID: **CW Name:** **CW Type:** ▼

PO Number: **Vendor:** **RUNDA Number:**

Status: ▼

Location Assigned:

Job title:

Hiring Manager:

Group: ▼ **Sub Group:** ▼

Start Date: **End Date:**

Figure 3.7: Snapshot of one of the modals used to enter CW data

Chapter 4

Related Work

One of the key points to be kept in mind is user experience. The following section deals with the improvement of performance of the web page in terms of loading time, network utilization and memory heap utilization. The second section contains how knowledge discovery process can be applied at each stage of the project development and implementation. Visualization techniques can be applied to provide a unique perspective to the user in terms of interaction and drill-down, rather than displaying two - dimensional charts.

4.1 Optimization

Optimization is one key aspect that needs to be involved in the project for higher efficiency and enhanced customer experience. No user would appreciate longer page loading time. These results have been obtained with the use of Developer tools in Google Chrome.

4.1.1 Web Page Performance

Web Page optimization or site enhancement is the field of information about expanding the rate in which website pages are downloaded and showed on the client's web program [13]. Quicker site download velocities have been appeared to expand guest maintenance and unwaveringness and client fulfillment, particularly for clients with moderate web associations and those on cell phones. Web execution likewise prompts less information traversing the web, which thus brings down a site's energy utilization and natural effect. Some of the factors which can influence the pace of page burden include program/server reserve, picture streamlining, and encryption (for instance SSL), which can influence page

rendering time. The execution of the website page can be increased through systems, for example, multi-layered store, light weight outline of presentation layer parts and non-concurrent correspondence with server side segments.

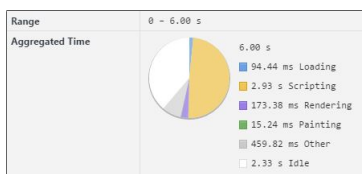


Figure 4.1: Web performance graph - before

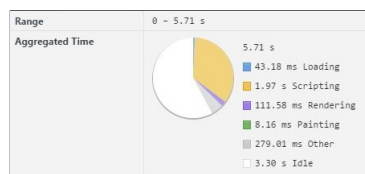


Figure 4.2: Web Performance graph - later

4.1.2 Network utilization

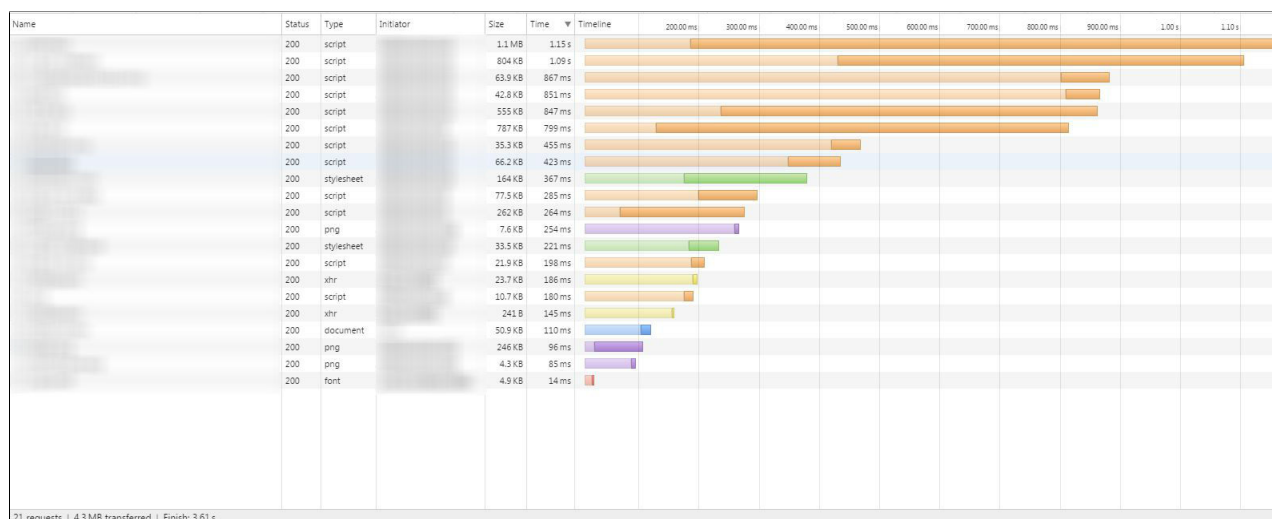


Figure 4.3: Network utilization graph - before

Network utilization is the proportion of current system activity to the most extreme movement that the port can deal with [14]. It shows the data transfer capacity use in the system. While high system use demonstrates the system is occupied, low system use shows the system is unmoving. At the point when system use surpasses the edge under typical condition, it will bring about low transmission speed, irregularity, solicitation postponement etc. Systems of distinctive sorts or in diverse topology have distinctive hypothetical look esteem under general conditions. On the other hand, this doesn't imply that the higher the network utilization, the better. For enhance user experience, it must be ensured that there is no bundle misfortune when system use achieves a sure esteem. Through checking utilization of network, we can comprehend whether the system is unmoving, typical or occupied. It additionally helps us to set legitimate benchmark

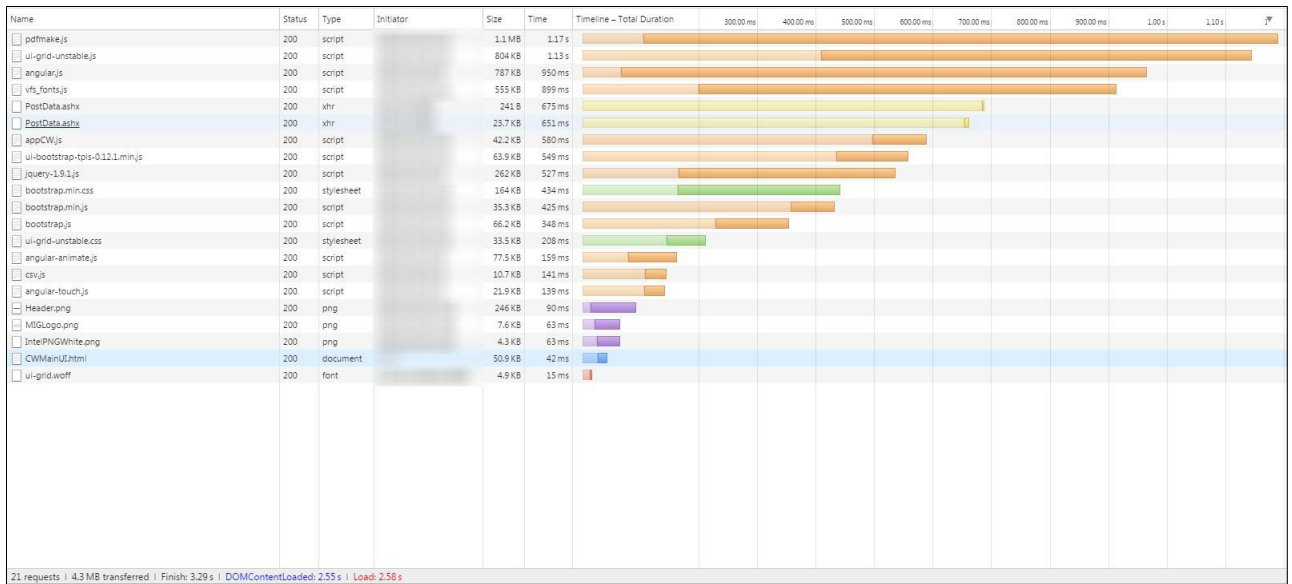


Figure 4.4: Network utilization graph - later

and investigate system disappointments. The following results have been obtained in 30 Mb/s Wifi network and RTT of 2ms.

4.1.3 Memory utilization

The following results have been obtained for memory utilization by one of the pages in the project.

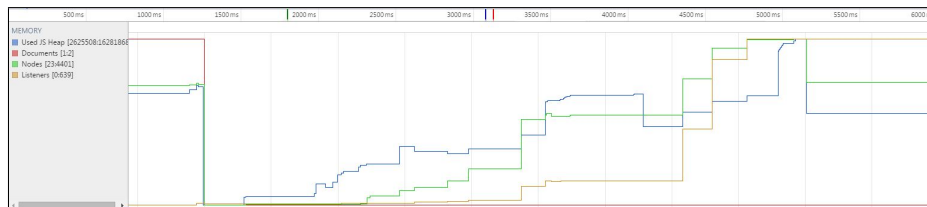


Figure 4.5: Memory utilization graph - before

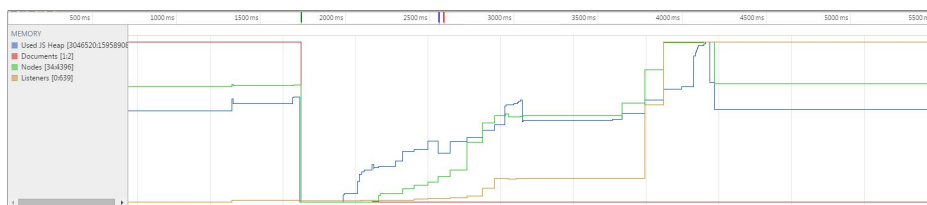


Figure 4.6: Memory utilization graph - later

4.2 Knowledge Discovery Process

The term Knowledge Discovery in Databases, or KDD for short, alludes to the wide procedure of discovering information in information, and stresses the "abnormal state" use of specific information mining techniques [15]. It is of enthusiasm to specialists in machine learning, databases, measurements, and information perception. The bringing together objective of the KDD procedure is to concentrate information from information in the connection of vast databases. It does this by utilizing information mining routines (calculations) to concentrate (distinguish) what is regarded learning, as indicated by the determinations of measures and edges, utilizing a database alongside any required preprocessing, subsampling, and changes of that database.

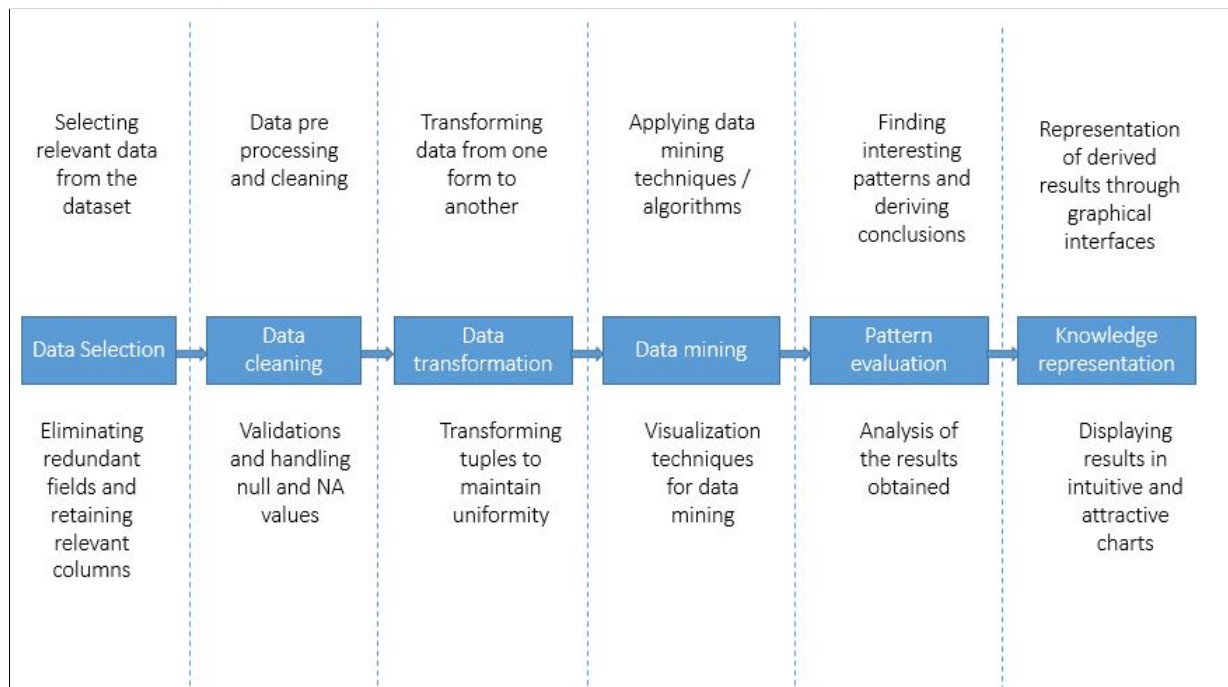


Figure 4.7: KDD process application in project development stage

Chapter 5

Further Work

5.1 EC Efficiency Initiatives

5.1.1 Objective

The objective of this project is to perform estimation of current resource utilization that will further help in forecasting future resource requirements

5.1.2 Problem Statement

- The required data resides at different sources.
- Computing CPU time, run time and disk usage for an IP considering defined use cases.
- Ability to analyse data considering a set of predefined use cases.

5.1.3 Proposed solution

- Building data model from external sources and performing computations to generate relevant data set
- Comprehensive Pivot Table enabled with interactive filtering and drill down features
- Analysing and visualizing data with Business Intelligence tools

5.1.4 Implementation

Query Optimization

Query optimization is an element of numerous social database administration frameworks. The query optimizer endeavors to decide the most proficient approach to execute a given query by considering the conceivable question plans. Once inquiries are submitted to database server, and parsed by the parser, they are sent to the query optimizer where improvement happens. Some database engines permit directing the inquiry streamlining agent with insights.

The streamlining agent endeavors to create the best execution arrangement for a SQL query. The best execution arrangement is characterized as the arrangement with the least cost among all arrangements. The cost calculation represents components of inquiry execution like input-output, CPU, and other resources. The best strategy for execution relies on upon heap conditions including how the inquiry is composed, the span of the information set, the design of the information, and which structures exist. The SQL query analyzer decides the best arrangement for a SQL articulation by looking at numerous entrance techniques and distinctive join strategies. Since the database has numerous inward insights and devices available to its, the analyzer is more often than not in a superior position than the client to decide the best technique for explanation execution. Hence, all SQL queries utilize the optimizer. The streamlining agent might choose that a full table output is generally productive. Perusing a record through table access by a column might be more proficient than a full table output [16].

Query optimization is the general procedure of picking the most productive method for executing a SQL proclamation. The streamlining agent is allowed to combine, rearrange, and prepare in any request since SQL is a non procedural dialect. The database expands every SQL clause in light of measurements gathered about the required information. While creating execution plans, the streamlining agent considers diverse access ways and join techniques. Components considered by the analyzer include:

- Framework assets, which incorporates I/O, CPU, and memory
- Number of columns returned
- Size of the underlying information set [16].

An execution plan portrays a prescribed strategy for execution for a SQL proclamation. The arrangements demonstrates the operations database uses to implement a SQL query. Every stride either recovers lines of information physically from the database or sets them up for the client issuing the query. An execution plan shows the expense or cost of the whole arrangement, and every different operation. The expense is an inner unit that the execution plan just shows to take into account arrangement correlations.

In this project, the query optimization techniques implemented are:

- Appropriate data types for columns
- Using select with required columns [17]
- Adding indexes for most frequently used columns [18]
- Using UNION instead of OR .

Table 5.1: Comparison of results before and after query optimization

Fields	Values (Before)	Values (After)
Rows returned	4	4
Type of SELECT	Simple	Primary, Union, Union result
No. of table scans	1	1
Index	NA	Block (Non unique, non-clustered)
Rows returned	4	4
CPU Time	16ms	16ms
Elapsed Time	46ms	6ms
Cost of Table Scan	65%	43%

Data Visualization

Data Visualization or data representation includes the creation and investigation of the visual representation of information, signifying data that has been preoccupied in some schematic structure, including traits or variables for the units of information. An essential objective of information perception is to convey data unmistakably and effectively to clients by means of the measurable illustrations, plots and data graphics charts. Viable representation offers clients some assistance with analysing and reason about information and confirmation. It makes complex information more open, reasonable and usable. Clients might have specific systematic undertakings, for example, making correlations or

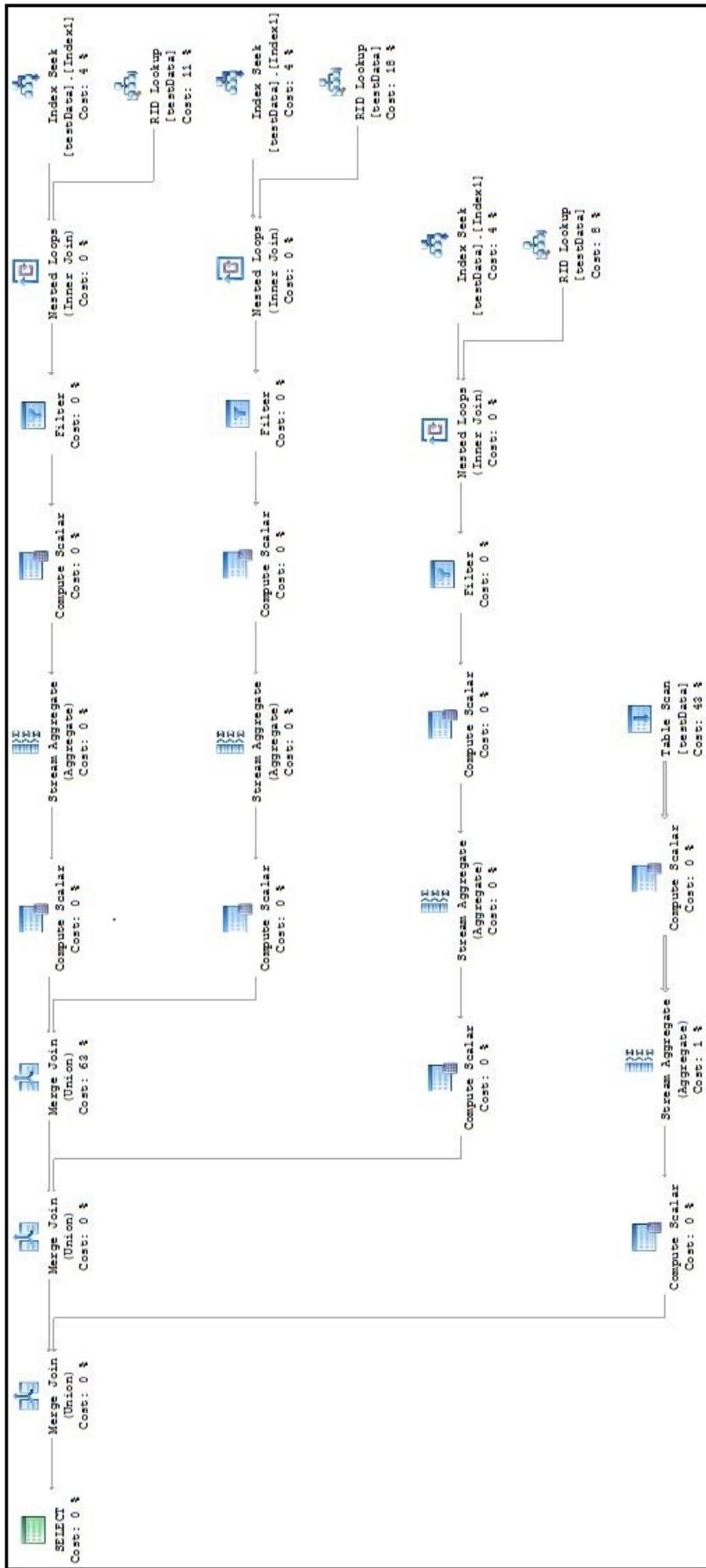


Figure 5.1: Query Execution plan

comprehension causality, and the configuration rule of the realistic (i.e., demonstrating examinations or indicating causality) takes after the assignment. The rate at which information is created has exponentially increased. Information made by web action and an extending number of sensors in nature are alluded to as "Big Data". Handling, dissecting and conveying this information exhibit an assortment of moral and diagnostic difficulties for information perception. The field of information science and professionals called information researchers and data scientists have developed solutions to address this test.

Information perception alludes to the procedures used to impart information or data by encoding it as visual items (e.g., focuses, lines or bars) contained in representation [19]. The objective is to impart data obviously and proficiently to clients. It is one of the progressions in information investigation or information science. Developers neglect to accomplish a harmony in the middle of structure and capacity, making exquisite information representations which neglect to fill their principle need "to convey data".

In this project, visualization techniques implemented are:

- Interactive filtering that allows users to select filters instead of predefined filters
- Data Aggregation which deals with summarizing data for further statistical analysis
- Drill down in which data is depicted as fine grained and coarse grained data

5.2 BASE Website

5.2.1 Objective

The main objective of this project is to design and implement new user interface for the team's portfolio by using recent UI frameworks. If a potential customer wants to request for a new tool, he/she should be redirected to new requester's dashboard. For existing customers, bug reporting and feedback mechanism is enabled for effective handling and monitoring of requests.

5.2.2 Problem Statement

- Requests for new tool, reports on system bugs, enhancement and feedback on existing tools were received on emails.

- Tracking, monitoring and managing requests became a cumbersome process.
- Updating the overall look and feel of the team’s portfolio.
- Elaborate email signature with hyper links for new requests, bug reporting and feedback.

5.2.3 Proposed Solution

- New design for user interface by implementing recent web technologies
- Web page with different hyper links to contact the team on the basis of the type of the request.
- Single interface showcasing the tools developed by the team during the tenure.

5.2.4 Implementation

Font Awesome

Font Awesome gives versatile vector symbols that can in a split second be altered size, shading, drop shadow are some of its properties of vector symbols that can be changed via CSS [20]. Font Awesome doesn’t require JavaScript. Similar to CSS, it allows effortlessly style symbol shading, size, shadow, and anything that is conceivable with CSS. Font Awesome symbols are vectors, which mean they’re exquisite on higher digital quality screens. Its adaptable vector representation enables each symbol to look magnificent at any size. In this project, Bootstrap’s glyphicons and Font Awesome icons have been used to represent tools and web pages.

Sweet Alert

Sweet alert is a wonderful substitution for standard alert message [21]. SweetAlert consequently focuses itself on the center of the page and looks extraordinary regardless of the device being used - desktop PC, smartphones or tablet. It can be customized exceptionally to prompt messages, automatically close after interval and including icons. In this project, sweet alerts have been used to notify the user in case a required field is left blank and successful submission of feedback.

Jquery UI

jQuery UI is a curated set of user interface collaborations, impacts, gadgets, and themes based on top of the jQuery JavaScript Library [22]. JQuery UI is useful for building web applications with highly intuitive and interactive interfaces. jQuery UI is a collection of additional functionalities for user input. In this project, jQuery UI plugin for enabling multi-select via check box was implemented. The user can select multiple recipients for sending feedback by simply checking the box next to their names.

5.2.5 Snapshots

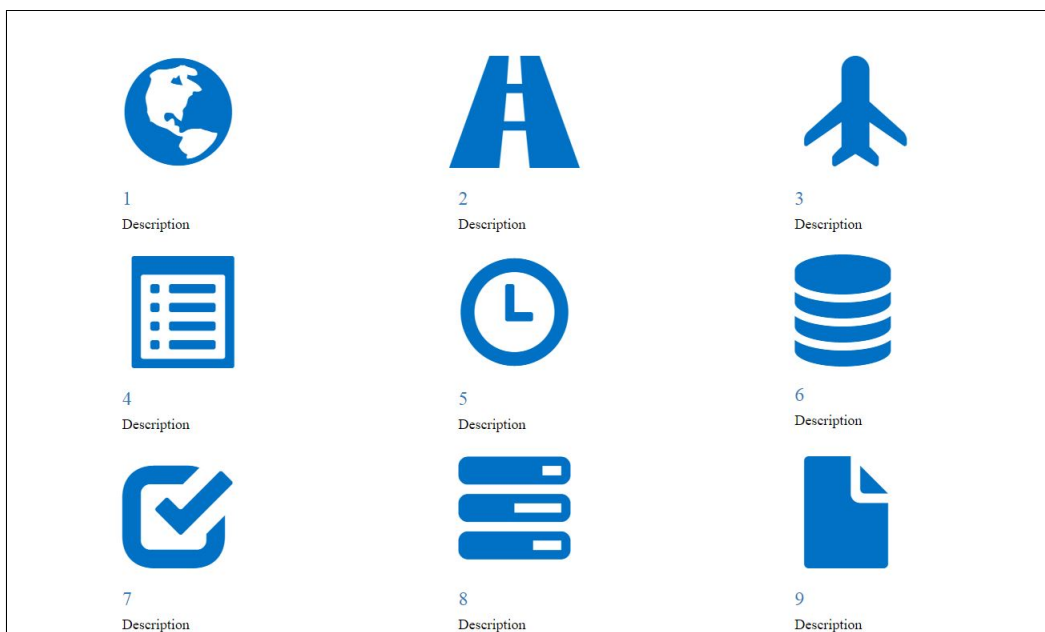


Figure 5.2: Team's portfolio where the names of tools and description have been replaced with other text

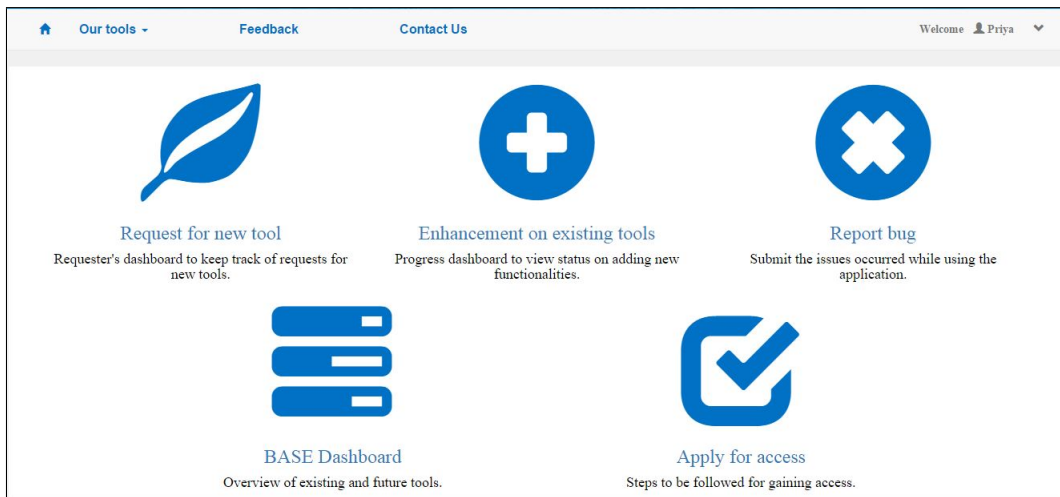


Figure 5.3: Interface to display dashboards to keep track on the status of requests

Chapter 6

Conclusion

6.1 Conclusion

The work undertaken is to eliminate the problem of data integration and duplication by consolidating a single database and normalizing it. The user interface has been created so as to not to overwhelm the user with too much information. The data grid also allows update operations and ensures no insertion anomalies. Customized emails are sent in case of occurrence of error in order to notify the user. Performance testing was carried out to optimize the page rendering time, memory utilization and network bandwidth utilization. Complex join queries are optimized by adding indices that help in optimizing the database query. Performance enhancing and optimization techniques have been used to improve the overall user experience.

Further work was carried out to continue performance enhancement through query optimization and visualization. By observing the query execution plan, the major operation were identified to be table and sort operations. Query optimization techniques such as adding indices and using UNION instead of OR statement for better database performance. With this, the cost of table scan operation was reduced from **65% to 43%**. The total execution time was reduced by **7 times**. Data Visualization was achieved by implementation of interactive filtering and drill down techniques. Value added to team's portfolio by using recent web technologies by updating overall look and feel of the UI.

Future work could be done in the area of query optimization for SQL queries that include multiple OR statements. Another area is to identify root causes of table fragmentation and resolving them to achieve better performance.

Bibliography

- [1] “Intel Company Overview.” <http://www.intel.in/content/www/in/en/company-overview/company-overview.html/>. [Online; accessed 14-December-2015].
- [2] D. Keim, “Information visualization and visual data mining,” *Visualization and Computer Graphics, IEEE Transactions on*, vol. 8, pp. 1–8, Jan 2002.
- [3] M. Chen, D. Ebert, H. Hagen, R. Laramée, R. van Liere, K.-L. Ma, W. Ribarsky, G. Scheuermann, and D. Silver, “Data, information, and knowledge in visualization,” *Computer Graphics and Applications, IEEE*, vol. 29, pp. 12–19, Jan 2009.
- [4] H. Jin and H. Liu, “Research on visualization techniques in data mining,” in *Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on*, pp. 1–3, Dec 2009.
- [5] A. Dasgupta and R. Kosara, “Adaptive privacy-preserving visualization using parallel coordinates,” *Visualization and Computer Graphics, IEEE Transactions on*, vol. 17, pp. 2241–2248, Dec 2011.
- [6] N. Robinson and M. Shapcott, “Data mining information visualisation - beyond charts and graphs,” in *Information Visualisation, 2002. Proceedings. Sixth International Conference on*, pp. 577–583, 2002.
- [7] “Bootstrap framework .” <http://getbootstrap.com/>. [Online; accessed 14-December-2015].
- [8] “Introduction to the C Language and the .NET Framework .” <https://msdn.microsoft.com/en-IN/library/z1zx9t92.aspx/>. [Online; accessed 14-December-2015].

- [9] “HTTP Handlers and HTTP Modules Overview.” <https://msdn.microsoft.com/en-us/library/bb398986.aspx/>. [Online; accessed 14-December-2015].
- [10] “AJAX - Send a request to a server .” http://www.w3schools.com/Ajax/ajax_xmlhttprequest_send.asp/. [Online; accessed 14-December-2015].
- [11] “Understanding HighCharts.” <http://www.highcharts.com/docs/chart-concepts/understanding-highcharts/>. [Online; accessed 14-December-2015].
- [12] “Basic MVC Architecture.” http://www.tutorialspoint.com/struts_2/basic_mvc_architecture.htm/. [Online; accessed 14-December-2015].
- [13] “Web Performance Optimization.” <http://www.stevesouders.com/blog/2010/05/07/wpo-web-performance-optimization//>. [Online; accessed 14-December-2015].
- [14] “How to analyze network utilization rate.” http://www.colasoft.com/capsa/network_bandwidth_analyzer.php/. [Online; accessed 14-December-2015].
- [15] “Overview of the KDD Process.” http://www2.cs.uregina.ca/~dbd/cs831/notes/kdd/1_kdd.html/. [Online; accessed 14-December-2015].
- [16] “Query Optimizer concepts.” https://docs.oracle.com/database/121/TGSQL/tgsql_optcncpt.htm#TGSQL194/. [Online; accessed 02-April-2016].
- [17] “SQL Database tuning.” <http://www.toptal.com/sql/sql-database-tuning-for-developers/>. [Online; accessed 02-April-2016].
- [18] “Optimizing data access.” <http://www.codeproject.com/Articles/34372/Top-steps-to-optimize-data-access-in-SQL-Server/>. [Online; accessed 02-April-2016].
- [19] “Analyzing and Visualizing Data with Excel .” <https://www.edx.org/course/analyzing-visualizing-data-excel-microsoft-dat206x-0/>. [Online; accessed 02-April-2016].
- [20] “Font Awesome Home Page.” <https://fontawesome.github.io/Font-Awesome/>. [Online; accessed 02-April-2016].

[21] “Sweet Alert Home Page.” <http://t4t5.github.io/sweetalert/>. [Online; accessed 02-April-2016].

[22] “jQuery UI.” <https://jqueryui.com/>. [Online; accessed 02-April-2016].