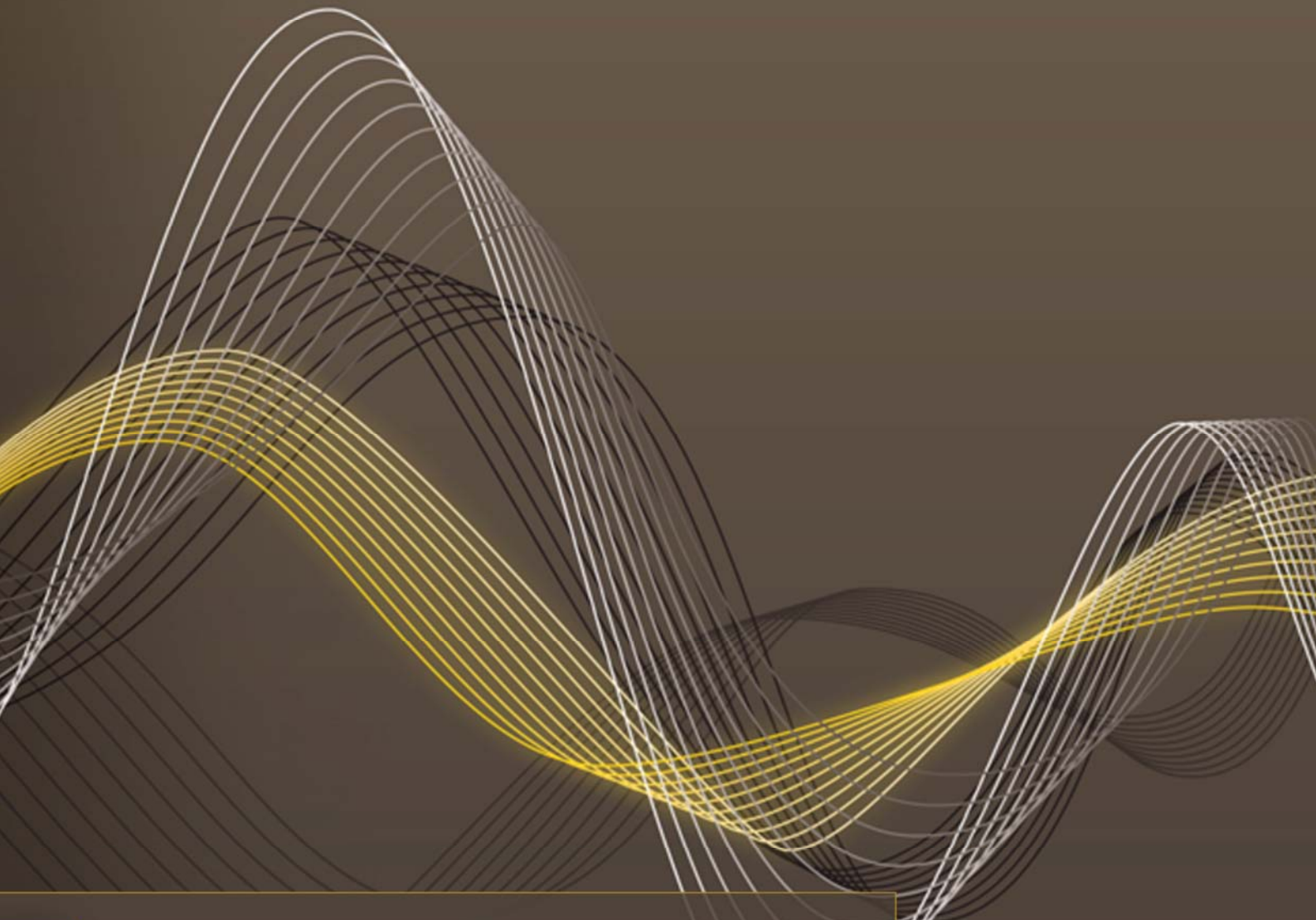
WEBSITE ANALYSIS

ETHIOPIAN CENTRAL STATISTICAL AGENCY (CSA)



Consultant: Araya Lakew
National Web Site Analyst and Developer
GCP/ETH/071/EC
Support to Food Security Information Systems in Ethiopia
arayalakew@yahoo.com
September 4, 2009

Table of contents

ABBREVATIONS

CMS: Content Management System

CSA: Central Statistical Agency of Ethiopia

RDBMS: Relational Database Management System

1 - An overview of CSA and its website

1.1 - Acknowledgements

I would like to thank to the FAO Chief Technical Advisor, Mr. Raphy Favre for his assistance in facilitating all my activities. I would also thank Mr. Yakob Mudesir for facilitating interviews and helping in distributing questionnaires. I also like to thank Eleni Kebede and Alemayehu G/tsadik for the effective and valuable response they offered in the analysis process. I would like to thank Mr. Thomas Gabrielle, who has been a lot of help on upgrading this document. CSA staff member Biruk was also helpful in giving out information. I look forward that this collaboration will continue throughout the implementation of the project.

1.2 - Background

The use of statistical information in Ethiopia can be traced back as far as the sixteen century. However, demand for statistical information for the purpose of economic management became an issue beginning 1957. It was in 1960 that statistics was put in place as regular government activity. This marked the first step in institutionalizing statistical work in the country. The basis for the establishment was the resolution of the Addis Ababa conference of the African Statisticians from UNECA member countries in 1960.

The statistical practice initiated at the time was organizationally set up within the then Ministry of Commerce, Industry and Tourism. But in 1963, the regular statistical activities became the mandate of a newly structured and autonomous organization called Central Statistical Office (CSO). At the beginning, CSO was responsible to the Ministry of Planning and Development and thereafter to the Planning Commission up until 1964. Then, CSO was reestablished in 1972 by proclamation number 303/1972 and was responsible for the then Planning Commission. Finally, CSO was restructured and became responsible to the Council of Ministers by the name of Central Statistical Agency, CSA, on March 9th, 1989. Furthermore, CSA became responsible to the Ministry of Economic Development and Cooperation on October 1996 and since September 2001 to the Ministry of Finance and Economic Development.

The Agency has carried out several socio-economic and demographic surveys that include agriculture, price, household income, consumption and expenditure, welfare monitoring, large and medium scale manufacturing and electricity industries, small scale manufacturing industries, cottage industries, construction, mining and quarrying, transport and communications, informal sector, distributive trade and services, manpower, demography, family and fertility, health and nutrition, child labor, etc.

The Agency runs the National Integrated Survey Program on annual basis and this operation involves quite substantial number of professional staff (Statisticians, Economists, Demographers, Mathematicians, Computer programmers, etc). There are also semi professionals that include statistical technicians, data editors and coders, data entry operators, field supervisors, enumerators and other supporting staff. The Agency also occasionally undertakes an ad-hoc survey that requires specialized personnel like only female enumerators, supervisors, field editors, etc. In such cases the office hires the field staff on temporary basis for the survey period and lays them off as soon as the field work of the survey in question is completed.

1.3 - About SFSIS

Support to Food Security Information Systems (SFSIS) in Ethiopia' was a project funded by the European Commission and implemented by the Food and Agriculture Organization of the United Nations (FAO). The project worked closely with the Government of Ethiopia (GoE) and in particular the Central Statistical Agency (CSA), Ministry of Agriculture and Rural Development (MoARD) and National Meteorological Agency (NMA).

1.4 - Objective of the website

The main objective of CSA's website is to provide information to target users and general public. A centralized RDBMS system is being implemented. Making this RDBMS data available on the web is also one of the website goals as most target users visit the website to query CSA's data.

1.5 - Methodology

The methodology that is used to collect the data for the project is:

- Reviewing the existing relevant documents
- Reviewing current website
- Reviewing similar existing systems
- Gather information from internal CSA staff members
- User survey using Questionnaire and interview

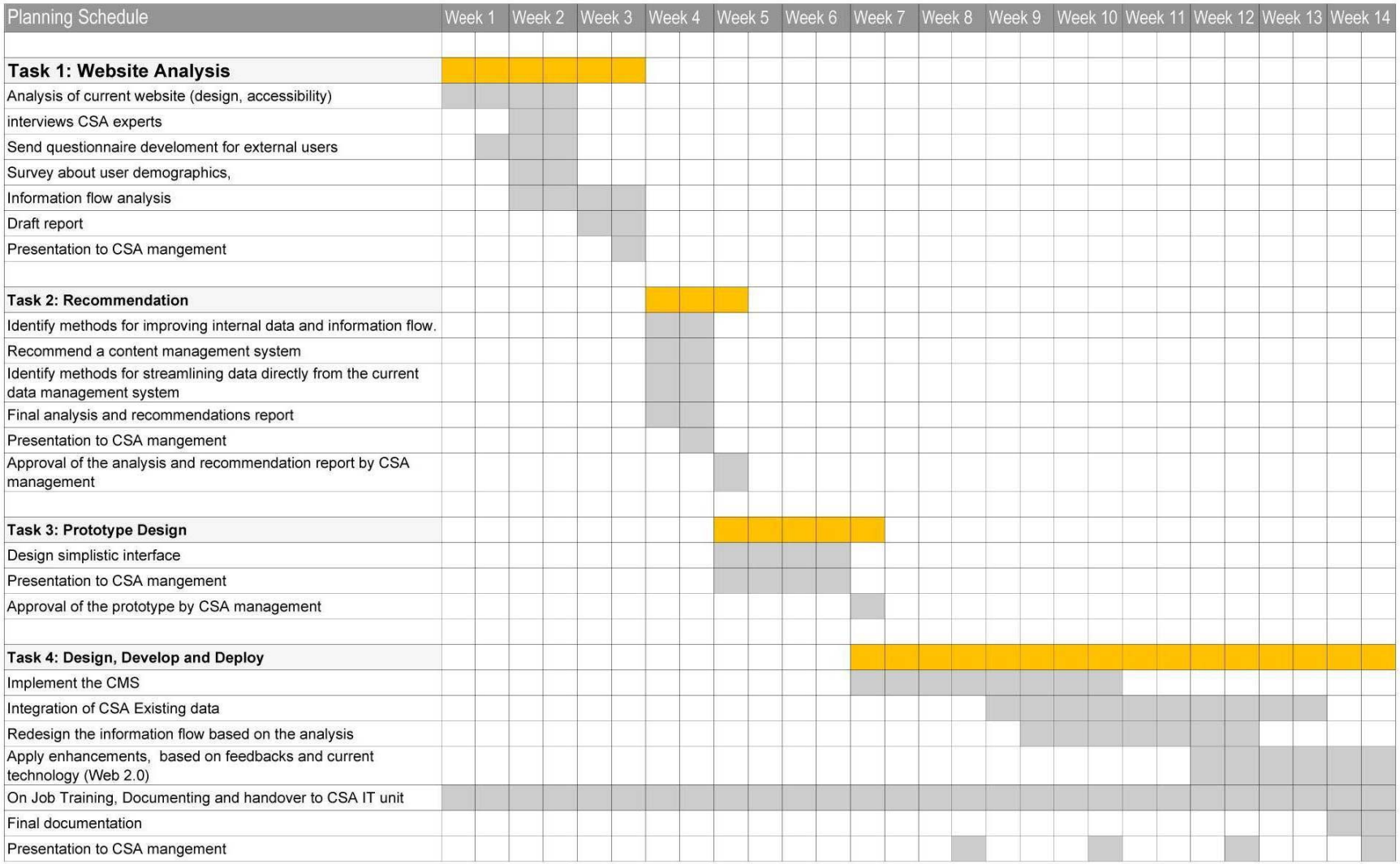
1.6 - Goal of the website

The main goal of the website is to make the website more dynamic and functional for the user. Integrating current data management systems into the website to make data accessible for the users and creating smooth and easy user experience is also expected.

1.7 - Scope of work

- ➔ Identify how internal and external users find the website useful and how it allows them to access CSA's data.
- ➔ Analyze the website regarding its ease of use (Navigation, presentation, layout and design),
- ➔ Analyze the information organization, document management, and speed of website to download.
- ➔ Analyze the current information flow of the website content. (Roles and responsibilities)
- ➔ Recommend best approach for the enhancement of the website
- ➔ Implement the recommendation when approved
- ➔ Integrate the website to the current database systems and the impending RDBMS system.

Work plan for CSA Website Analysis and Development



* Adjustments could be made depending on the RDBMS system being implemented

2 – Analysis of current CSA website

2.1 Overview of the website

The CSA website is one of the richest website of the Ethiopian Institutions in terms of its content. The website describes all the major activities and operational structure of CSA to anyone who is not familiar with CSA. Although it has lots of information it is difficult to find specific information due to its content organization and lack of functionalities. The current website is fully static and it is updated by modifying every static page whenever there is an update which makes it a tedious task. Data is currently accessible to internal and external users using various systems such as Dev-info, and in house developed database systems.

2.2 Users Analysis

The major audiences of CSA website are policy makers of the federal government and regional states, International organizations and planners at different level. Generally the following are also target users:

- ➔ UN and International organizations
- ➔ Researchers
- ➔ Students, particularly of higher learning institutes
- ➔ Journalists
- ➔ Sponsors of different Research on related subjects
- ➔ National and international survey sponsors
- ➔ Advocates / General Public

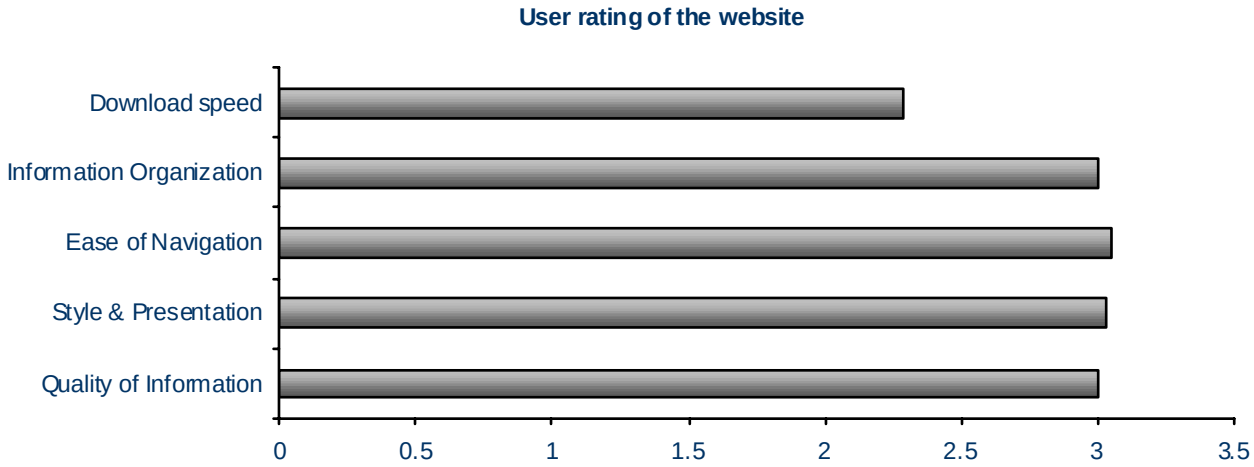
Analysis of CSA's website started from usability prospective to understand how user friendly the website is and to get the general overview of users experience using the website. The process started with user survey to find out the expectations and requirement of the website users. Online

user survey was carried out for about 3 weeks and more than 30 users participated to fill the online questionnaire. The questionnaire was prepared to get users view, usage experience and feedback of CSA website. An interview was also conducted for 11 organizational users. After interviewing 11 organizational users and collecting the online questionnaires, an overall user insight was gained into the needs of the users. The interviews provided an opportunity to receive detailed and feedback to what these target groups expect from the CSA website and how it would help or hinder their work practice.

The questionnaire consists of 9 questions which are a mix of check box and free text. A link for the questionnaire was posted on the website. There were various users participated on the online questionnaire. A journalist who visit the website to query data on a daily basis, a student doing his Ph.D, a visitor who was looking for news, a visitor who was looking for the vacancy page are some of the website users. Some of the users recommended for the news page to be updated often and making data available as they exist.

On the questionnaire users were asked to rate the website for quality of content, style and presentation, ease of navigation, information organization and download speed. Other questions on the questionnaire also include: How often you visit the website? What is the main purpose you visit the website for? Which pages do you visit the most? What do you like the most about CSA website?

Based on the survey result, users mainly visit the website to query data and various reports. The main constraint they faced with was the download speed although most of them were broadband internet users. Most users find the website as a useful source of information despite that it is difficult if not sometimes impossible to find specific information as there is no search feature. Information organization was also one constraint why users couldn't find documents easily.



2.2.1 - Some responses from the Online Questionnaire

More than 30 responses were received for the online questionnaire and some of the responses are listed below:

How often do you visit the website?
Occasionally: 15 responses
Weekly: 6 responses
Daily: 3 responses
Once in a while: 1 response

What type of Internet connection are you using at this moment?
Broadband: 18 responses
Dial up: 9 responses

Which page or pages do you visit the most?

- *Surveys/census section*
- *The pages that I mostly visit are: EthioInfo as UNICEF support the system, Surveys and Census and the price index . However I sometimes also go through other pages such as the publications and Organization part focusing on policy and strategy documents.*
- *i have been visiting your web for the past one year specially the vaccancy list but i always see one item which is out dated before one year.*
- *Price index/database, publications, etc*
- *population data,surevy resoult*
- *mostly data archive*
- *PRICE INDEX/DATA BASE/*
- *house hold consumption*
- *census & statistical survey reports*
- *Price Index*
- *avaliable survey*

- *AGRICULTURAL DATA*
- *An organized production and Import/Export data base.*
- *Publication*

What is the main purpose you visit the website for?

- *To access data*
- *The main purpose why I often surf and visit CSA website are-(a) UNICEF Supports the dissemination of surveys conducted by CSA via the socio - economic user friendly database called EthioInfo . Hence we check and follow up on the proper functioning of this specific page system. (b) for data inquiry and uploading of recent surveys and (c)to be up to date with new events , activities, political changes and etc that are taking place within CSA and beyond..*
- *1 to extract relevant data 2 to have an employment opportunity*
- *Looking for data*
- *getting data*
- *I usually visit this website to get official data at macro level*
- *RESEARCH*
- *to do Ph.D Dessertation*
- *to get relevant data*
- *To find material to write Wikipedia articles with. Many articles on the cities and administrative units of Ethiopia in Wikipedia could not have been written without the information the CSA provided on this website*
- *for our weekly Amharic Newspaper Economy to Transforming Your Data to our readers*
- *To search for secondary data on some aspects of agricultural sector*
- *obtaining organized data*
- *STUDIES*
- *To assess the market suitaion of the country using detail analysis.*
- *Get statistical data*

What do you like the most about this web site?

- *The fact that so many data are available.*
- *The site has got a wealth of information and I appreciate the fact that CSA is trying to bridge the gap of information in the country.*
- *almost all*
- *Well organzed, and the homepage is very fast to open.*
- *Figures*
- *I like that there are so many of pdf files on different survey result*
- *PUBLICATION SECTION*
- *having data at household level*
- *The reports from the 1994 census*
- *the data you collect and organaized on conssumer price*
- *nothing. It is very poor*

- *almost every thing exept the speed to download*
- *Agricultural survey report*
- *I used to statistical data such as price indexes and population census from this site for papers I am working on. No more available. All need money.*

What do you like the least about this web site?

- *The website is not regularly updated*
- *The home page (first page) is always the same (you always find the welcoming page), where as I think this page should capture new and timely interesting issues on statistics and events every now and then. - There are similar /duplication of features or functions on the page for instance ENADA list down and archives all surveys as well as the list of survey tab does the same thing. The homepage looks overloaded because of this reason. I assume that there are some tabs that could feed under one another instead of having them separately (i.e. site map tab could go under organization tab) - Coming to the layout of the page I don\'t like the background colour and the unfriendliness of the layout, the top part with the graph and number it is not appealing for the non statistician audience*
- *speed to download*
- *nothing*
- *The data base especially very slow to load and open, the queries do not respond at all*
- *Even with in the pdf files it is not easy to get numerical data which is the main reason for anyone to go to national statistical offices websites*
- *NO UPDATES*
- *The bandwidth of the servers. It can take an hour to download a PDF file*
- *the news section, becouse it didn\'t updated for so long*
- *Very very slow download process*
- *the speed*

What additional information would you like to see on CSA's website?

- *Link it with other useful websites and update it as often and regularly as possible.*
- *It would be interesting if you could include a search engine, networking forum where people from different area of profession and students could network and discuss on several statistical issues raised by the organization.*
- *Definitely the speed to access the website is a major challenge not even to think about downloading document . Hence CSA needs by any means to tackle this problem to increase the use of the website.*
- *It is sometimes good to go through other national statistics offices websites and adopt some of the web design and web presentation on your web system.*
- *I like to see data available in different format including excel format on different activities, like agricultural data, data on education, data on economic statistics, etc...easily available*
- *PLEASE UPDATE THE DATABASE OF PRICE MONTHLY MAKE IT ACCESSIBLE FOR RESEARCHERS EVEN LIVING ABROAD*

- *Volume 1 of the Afar report of the 1994 census is corrupt, and cannot be opened. Is it possible to create and upload a new PDF file that is readable?*
- *More material from the 2007 census, at least the population figures for the cities, towns and villages.*
- *pleas, update your news section also the produser price index which was updated three mounths a go.*
- *First, improve accessibility of the available ones. After that you can ask for additional data requirements.*
- *Please try to collect a full information on aparticular issue rahter than putting a partial information.Like consumption,production,trade,percapital consumption and per capital income etc of the major products so as to utilize the nationla resource efficeintly.Thank you*
- *Published & Avilable abstracts in the book sales shop(CSA main office)with: 1.Title 2.Contents 3.Fee for the soft copy(in CD)& hard copy(book)*

The result of the survey clearly shows that the download speed of the website was slow especially when users access data through the current database systems. Accessing DevInfo data is also difficult. The current website also lacks user interactivity such as contact forms. Some users who used the website on daily basis recommend that news should be updated often.

Users mostly find the digital libraries helpful to find data for their research or for their organizational works. Some users requested to find raw data for cross reference. Some users also find EthioInfo to be a useful resource but they are having problems accessing it.

Users could be categorized into different categories. These could be Organizational users, researchers, and others who are looking for data and mainly using the digital libraries. I found these group of users to be more technical having knowledge about IT. Even though they find their way in to the digital libraries somehow they are need more functionalities to make use of the website. The other group could be general users who are interested in getting reports, publications and other CSA information. These users are mainly basic internet users, and accessibility and ease of navigation are their major requirements.

Website Log analysis is one of the basic system which gives detailed information about the website users. The information found on the log files could be location of user, IP, what browser users used,

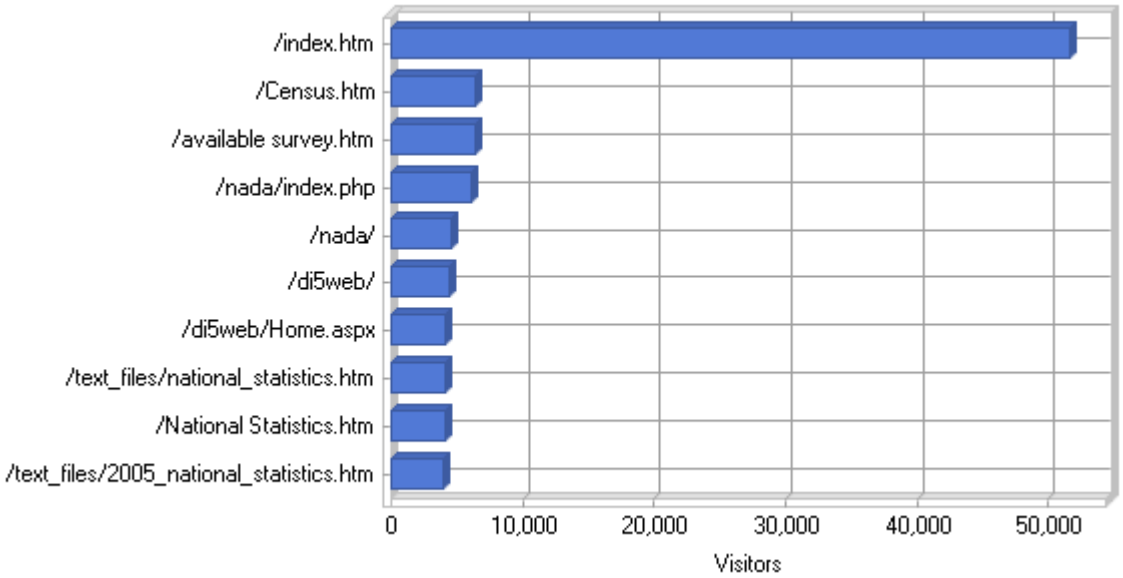
which pages they visited most and other details. Unfortunately website analytics system is not implemented on the current website.

2.3 Website Log Analysis

The Log files from the CSA's IIS web server were collected to get a more detailed picture of visitor information and activities on the website. The Logs give detailed information about visitors who were visiting the site for the last 20 months. Below are some of the findings of major areas of the website.

Hits	
Total Hits	7,589,529
Visitor Hits	6,765,959
Average Hits per Visitor	62.36
Cached Requests	245,119
Failed Requests	3,509,736
Page Views	
Total Page Views	2,142,152
Average Page Views per Visitor	19.74
Visitors	
Total Visitors	108,504
Total Unique IPs	48,492

Most Popular Pages

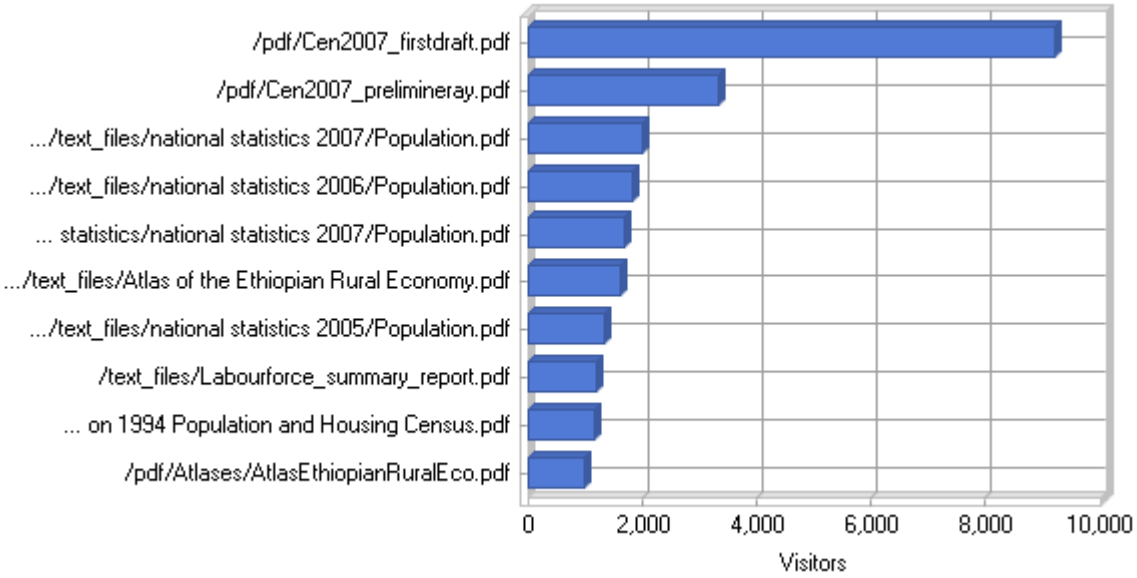


Most Popular Pages

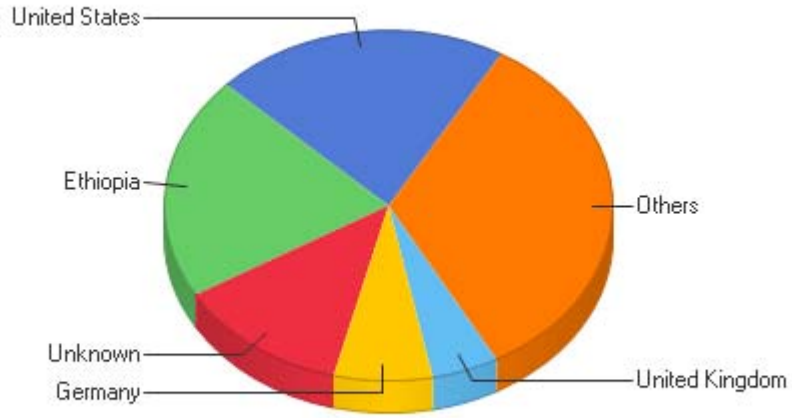
	Page	Hits	Incomplete Requests	Visitors
1	http://www.csa.gov.et/ index.htm	73,299	632	51,623
2	http://www.csa.gov.et/ Census.htm	8,117	42	6,440
3	http://www.csa.gov.et/ available survey.htm	8,718	55	6,410
4	http://www.csa.gov.et/ nada/index.php	22,315	0	6,086
5	http://www.csa.gov.et/ nada/	5,616	0	4,601
6	http://www.csa.gov.et/ di5web/	5,389	0	4,426
7	http://www.csa.gov.et/ di5web/Home.aspx	19,068	0	4,149
8	http://www.csa.gov.et/ text_files/national_statistics.htm	5,007	13	4,133
9	http://www.csa.gov.et/ National Statistics.htm	4,995	30	4,083
10	http://www.csa.gov.et/ text_files/2005_national_statistics.htm	4,663	19	3,939
11	http://www.csa.gov.et/ List of Current Survey.htm	4,293	30	3,557
12	http://www.csa.gov.et/ nada/templates/Theme 1/ styles.php	7,300	0	3,274
13	http://www.csa.gov.et/ text_files/publications.htm	3,308	8	3,067
14	http://www.csa.gov.et/ surveys/Population and Housing Census 1994/ survey0/	3,898	2	3,008
15	http://www.csa.gov.et/ text_files/EhioInfo.htm	3,295	8	2,995
16	http://www.csa.gov.et/ surveys/Population and Housing Census 1994/ survey0/navigation.html	3,837	15	2,948
17	http://www.csa.gov.et/ surveys/Population and Housing Census 1994/ survey0/banner.html	3,809	0	2,942
18	http://www.csa.gov.et/ surveys/Population and Housing Census 1994/ survey0/welcome.html	3,896	1	2,916
19	http://www.csa.gov.et/ surveys/Population and Housing Census 1994/	3,557	4	2,805
20	http://www.csa.gov.et/ di5web/devinfoapp.aspx	11,981	0	2,713
21	http://www.csa.gov.et/ text_files/available_surveys.htm	3,384	11	2,627

22	http://www.csa.gov.et/ Consumer_Price_Index.htm	3,119	11	2,531
23	http://www.csa.gov.et/ text_files/2007_national_statistics.htm	2,751	18	2,304
24	http://www.csa.gov.et/ PopulationandHousingCensus.htm	2,602	1	2,277
25	http://www.csa.gov.et/ nada/templates/Theme 2/ styles.php	4,797	0	2,270
26	http://www.csa.gov.et/ National Statistics 2008.htm	2,684	34	2,104
27	http://www.csa.gov.et/ surveys/Population and Housing Census 1994/ survey0/outputInformation/ reports.html	2,363	12	1,782
28	http://www.csa.gov.et/ nada/searchsurveys.php	8,267	0	1,734
29	http://www.csa.gov.et/ di5web/icallback.aspx	20,756	0	1,704
30	http://www.csa.gov.et/ text_files/2006_national_statistics.htm	1,953	11	1,666
31	http://www.csa.gov.et/ text_files/CPI.htm	1,862	5	1,646
32	http://www.csa.gov.et/ text_files/what's_new.htm	1,733	12	1,613
33	http://www.csa.gov.et/ text_files/List_of_surveys.htm	2,104	16	1,575
34	http://www.csa.gov.et/ Annual Agricultural Sample Survey and Enumeration.htm	1,935	26	1,500
35	http://www.csa.gov.et/ National Statistics 2007 .htm	1,904	42	1,465
36	http://www.csa.gov.et/ Vacancy and Notice.htm	1,678	30	1,401
37	http://www.csa.gov.et/ Population and Housing Census.htm	1,698	6	1,392
38	http://www.csa.gov.et/ text_files/about_CSA.htm	1,505	5	1,391
39	http://www.csa.gov.et/ Basic welfare Indicators.htm	1,648	30	1,383
40	http://www.csa.gov.et/ text_files/current_surveys.htm	1,402	5	1,312
41	http://www.csa.gov.et/ Departments.htm	1,605	26	1,283

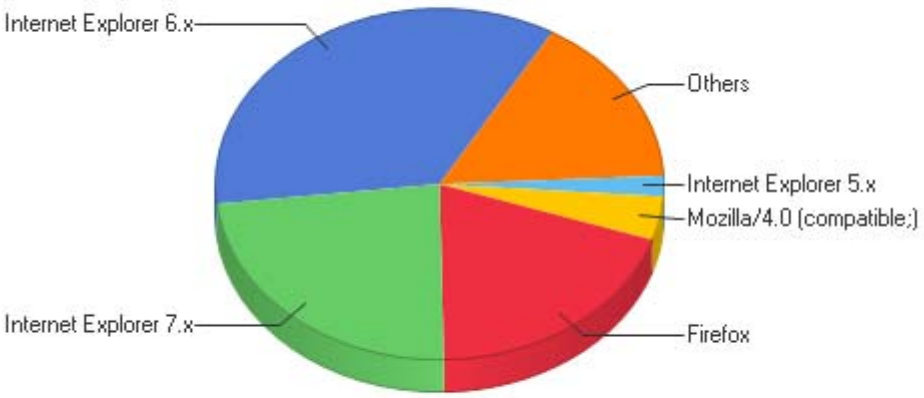
Most Downloaded Files



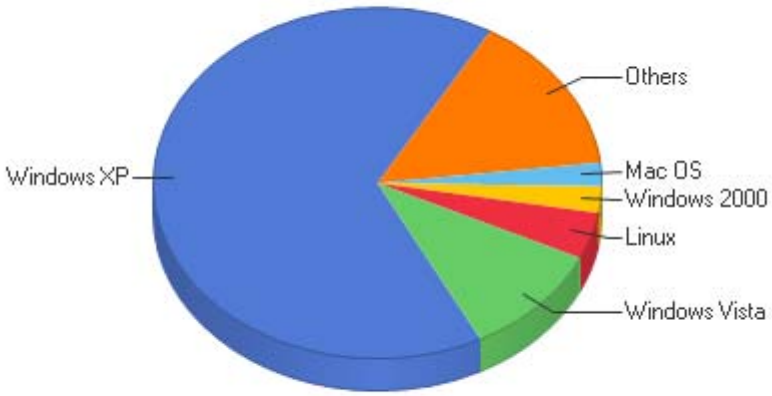
Most Active Visitors by Country



Most Used Browsers



Most Used Operating Systems



The website Log Analysis gives detailed information about visitors interests of the last 20 months. We can see the census page have a lot of visitors than others. The National Data Archive (NADA system) is more often visited by users. The NADA system is a system which enables users to search through the available surveys for detailed metadata using the IHSN system. Next to the NADA system, EthioInfo page is visited frequently. The analysis also shows us that 'the 2007 census result' is downloaded the most followed by 'the 2007 national statistics (population)' and 'The atlas of Ethiopian rural Economy. Based on the Log Analysis, 21.17% of visitors are from the United states

where as 20.68% of users are from Ethiopia. It is also found out that most users use Internet Explorer to browse CSA's website.

2.3 Current website assessment

The functionality of a site is imperative to not only attracting visitors but also retaining them. There are some basic rules for web sites.

- Website should be easy to read.
- Website should be easy to navigate.
- Information on website should be easy to find.
- Design should be clear and consistent.
- The site should be quick to load.
- The site should only be accessible to the authorized users.
- Website should have good search functionality

2.3.1 - Layout

The layout has been arranged in a three column layout presenting a good standard left column for hierarchical navigation, a middle column for featured content and a right column for quick links. Looking at its interface, it lacked a critical appeal to visitors as well as lacked potential functionality. Colors are used in a seemingly random manner which contributed to the lack of visual appeal and inconsistency to the website. Menu colors are not in cotrast and some lack standard mouse pointer effects. To achieve a good feel, it needs to be looked at how text, graphics, effects, color, layout, placement of links and speed are combined.

The home page must be informative and inviting. It should provide enough information for visitors to recognize what is being offered, enough to invite them to explore further, and simple navigation aids to make that task easy. CSA's website home page is informative and gives insight about insight pages.

2.3.2 -Fonts and color

Text should be easy to read by using default or standard fonts. The display properties of text (typeface, size and color of fonts) must be readable in both electronic and printed form. Not all fonts are supported or accessible by all users. Selecting standard font types (such as Arial, Verdana, etc.) where possible sans serif fonts (such as Arial) will increase readability. Font sizes should be easy to read the text on screen. The font used in CSA's website is standard and readable.

Link color should be consistent and should let the user know the text is a link. Some link colors in CSA's website are not user friendly as they are the same color as the normal text.

2.3.3 - Content and structure

Organizing content includes putting critical information near the top of the site, grouping related elements, and ensuring that all necessary information is available without slowing the user with unneeded information. Content should be formatted to facilitate scanning, and to enable quick understanding.

- Organize information at each level of the Web site so that it shows a clear and logical structure to typical users.
- Structure each content page to facilitate scanning: use clear, well-located headings; short phrases and sentences; and small readable paragraphs.
- Ensure that all needed information is available and displayed on the page where and when it is needed.
- Group all related information and functions in order to decrease time spent searching or scanning.
- Allow users to efficiently find what they want, design so that the most common tasks can be successfully completed in the fewest number of clicks.
- Limit page information only to that which is needed by users while on that page.

All the pages on the website have comprehensive content present, which is an advantage from both Search Engine & Usability prospective.

Page should not be over-filled. Visitors feel overwhelmed and be left wondering where on earth to start. It is fine to have lots of information, but it should be laid out logically and with enough white space in between to make it easy to compartmentalize.

2.3.4 - Navigation

a good navigation system should answer three questions:

1. Where am I?
2. Where have I been?
3. Where can I go?

CSA's current website is having proper internal linking structure & standard top navigation which makes it easy for user to locate information on the website. It is organized in a way a standard menu should be. The current navigation structure grouped related menus together which gives the user overview of the menus at a glance. Having some of the redundant menus grouped together and adding a breadcrumb which will allow visitor to know where they are currently, will add more functionality to the navigation system. Visitors should not get lost in the website, currently a user might not know where in the website they are. I recommend using the right menu for basic and focused areas such as indicators and census so the visitor should find it right away.

A robust site search feature helps visitors quickly locate the information they want. Breadcrumbs will let users which pages are they in currently. It shows the parent path from home of the current page. Currently breadcrumbs are not implemented.

Large or complex sites should always have a text-based site map in addition to text links. The site map at CSA's website is standard text based site map but it doesn't have links. Every page should contain a text link to the site map. Lost visitors will use it to find their way, while search engines spiders will have reliable access to all of your pages.

2.3.5 - Speed

People who use the internet are generally doing so because of speed and convenience. If we take those two elements away, we leave them with no reason to be there. For a website to load faster the major factor is the contents and elements in the page.

As the user survey show, short download times are of highest importance to them. CSA's website is rated low for download speed by users. Low bandwidth may contribute to the low speed of the

website as well as size of some PDF documents. Bandwidth intensive components should be avoided, such as large images, multimedia and large documents. Minimizing the size of larger documents or splitting them if necessary could solve the problem. Checking the uplink of the current connection should also be considered.

2.3.6 – Functionality

After reading the first page, visitors now want to browse further if the site seems to be of their interest. Perhaps they have specific information that they are looking for. In order for them to find what they need other than the navigational system they need a good search functionality. Although email is always possible, Contact forms directly from the website are also of good help if users want to contact CSA. Based on the user survey, users mostly visit the website to access data. And the database system functionality also plays a major role when considering the functionality of the website.

2.3.7 – Data Library

CSA has made digital libraries available through its website and most users tend to visit the website to get data from these digital libraries. Currently there are three database systems in the website. There is also RDBMS system being developed. The current website integrates these digital libraries but it lacks consistency in its look and feel. For example the National Data Archive section has a completely different look than the main layout which will confuse users. Integrating the IHNS system to have consistent look with the main layout is recommended. The data libraries current available are listed below.

2.3.7.1 DevInfo (Ethio-Info)

EthioInfo is in use in CSA at http://www.csa.gov.et/di5web/ . They are using it as a common platform for indicators related to Human Development, to facilitate data sharing and indicator harmonization at global, regional and country level by making statistics available to a wide audience. It allows presentation of data through Tables, Graphs and Maps. End users can get screens whereby they can enter some parameters for searching and information presentations. For example for the industry database one can use http://www.csa.gov.et/di5web/devinfoapp.aspx The site is using ACCESS Databases at the backend.ASP.NET programming language for the site at the front-end.

EthioInfo V 2.1 is the up-to-date version and contains the latest Ethiopia Demographic and Health Survey 2005 and also includes the following surveys;

- Welfare Monitoring Survey (WMS), 1996, 1998, 2000 & 2004
- Total Population for 2004 & 2005 (Population size, Area, & Density)
- Area and Production of temporary crops for 2004 & 2005
- Ethiopia Demographic and Health Survey 2000
- Household Income Consumption and Expenditure Survey (HICE), 1996,2000

It has been found out that EthioInfo is a customized adaptation of DevInfo, a world wide used user friendly software that helps to organize, present data in a result based environment with unique features linking to strategic monitoring and evaluation of policies such as MDG, National Poverty Reduction Strategies.

DevInfo is a powerful database system that is used to compile and disseminate data on human development. The software package has evolved from a decade of innovations in database systems that support informed decision making and promote the use of data to advocate for human development.

DevInfo was developed by UNICEF in cooperation with the UN System to assist the UN and Member States in tracking progress toward the Millennium Development Goals (MDGs). In 2002, DevInfo was proposed as a standard software package for the whole UN System. Its specific purpose is to store existing data, identify gaps in the MDG indicators, provide a single entry point for data on the MDG indicators, and disseminate information simply and attractively.

DevInfo is claiming that it is an integrated desktop and web-enabled tool that supports both standard and user-defined indicators. The standard set of MDG indicators is at the core of the DevInfo package. In addition, at the regional and country levels, database administrators have the option to add local indicators to their databases. The software supports an unlimited number of levels of geographical coverage: from global level to regional, sub-regional, national and sub-national down to sub-district and village levels (including schools, health centers, water points).

Data from DevInfo can be exported to XLS, HTML, PDF, CSV and XML files and imported from spreadsheets in a standardized format. DevInfo also has a data exchange module for importing data from industry-standard statistics software packages such as SPSS, SAS, Stata, Redatam, and CSPro.

DevInfo is distributed royalty-free to all Member States and UN agencies for deployment on both desktops and the web. The user interface of the system and the contents of the databases it supports include country-specific branding and packaging options which have been designed to ensure broad ownership by national authorities. UNICEF has absolutely no restrictions on the database and its use.

The most common DevInfo users include UN country teams, national statistical offices, planning ministries, and district planners. Frequent users also comprise members of the media (for reporting and tracking human development data), educational institutions (for analyzing data and helping students gain data access), as well as DevInfo administrators (in particular for customizing the system or adding data through advanced database administration modules).

2.3.7.2 IHSN Microdata Management Toolkit

The IHSN Microdata Management Toolkit developed by the World Bank Data Group for the International Household Survey Network aims to promote the adoption of international standards and best practices for microdata documentation, dissemination and preservation.

CSA is using all three modules of the Toolkit. They are using the Metadata Editor to document data in accordance with international metadata standards (DDI and Dublin Core). The Data Documentation Initiative (DDI) is an international project to create a standard for information describing social science data. The DDI specification, written in XML, provides a format for content, exchange, and preservation of information. Version 3 of the DDI standard is expected to be released in 2008.

DDI 3.0 is used to Support Preservation, Management, Access and Dissemination Systems for Social Science Data.

The DDI-tree contains five main branches, or sections:

1. The Document description, which consists of bibliographic information describing the metadata document and the sources that have been used to create it
2. The study description, which contains information about the data collection.
3. The Data files description, which describes each single file of a data collection (formats, dimensions, processing information, missing data information etc.)
4. The variable description, which describe each single variable in a datafile (format, variable and value labels, definitions, question texts, imputations etc.).
5. Other Study-Related materials, which can include references to reports and publications, other machine-readable documentation

CSA is taking the advantage of DDI to classify, describe, and organize datasets of its surveys.
The Dublin Core metadata standard is a widely recognized meta-language to describe information resources. It contains fifteen elements such as coverage, creator, date, description, format, etc.
The Explorer, free reader for files generated by the Metadata Editor is also in use. The module allows users to view the metadata and to export the data into various common formats (Stata, SPSS, etc).
CSA through its ICT department is widely and frequently using its CD-ROM Builder to generate user-friendly outputs (CD-ROM, website) for dissemination and archiving.

CSA is usually using this toolkit to publish CD-ROMs and website on each of the datasets.
Some of the users are claiming that they are using Nesstar and some others are claiming that they are using IHSN Microdata Management Toolkit. So, is there any difference between the two systems?

Nesstar is a commercial software system for data publishing and online analysis. The software consists of tools which enables data providers to disseminate their data on the Web. Nesstar handles survey data and multidimensional tables as well as text resources. Users can search, browse and analyse the data online.

Nesstar Publisher's feature is being integrated with IHSN. This component consists of data and metadata conversion and editing tools, enabling the user to prepare these materials for publication to a Nesstar Server. The Nesstar Server is built as an extension to a normal web server. As well as

providing all the usual facilities for publishing web content, this server provides the ability to publish statistical information that can be searched, browsed, analysed and downloaded by users. This is done either by using a standard web browser or using Nesstar WebView.

The Nesstar represents a system of software architecture that makes it easy to create, locate, access and operate remotely on metadata and corresponding data. At the same time it does this while maintaining a high level of compatibility with the WWW.

Nesstar has the following analytical tools which are not available in IHSN:

- Cross tabulations
- Correlations
- Regressions
- Compute and recode
- Graphical representations of data in customizable forms
- Application of variable weights

In fact, IHSN toolkit has taken a lot of features from Nesstar. In fact, IHSN is using a lot of components from Nesstar as they are. In reporting and dissemination of reports and datasets archiving, CSA is using IHSN. The staff members of Dissemination team are using IHSN to produce a single output particularly for CD-ROM version, then they are publishing same version to the website without optimizing it for web.

The ICT department is burning CD-ROMs for each of the surveys based on the number of target audiences of each of the surveys. Whenever more or extra copies are requested, they are producing requested CD-ROMs copies and sent them to public relations for distributions.

2.3.7.3 – CountryStat

CountrySTAT is a statistical framework and applied information system for analysis and policy-making designed in order to organize, integrate and disseminate statistical data and metadata on food and agriculture coming from different sources. CountrySTAT gathers and harmonizes scattered institutional statistical information so that information tables become compatible with each other at the country level and with data at the international level. The main objectives are to facilitate

decision-maker's access to information and to bind data sources that are currently spread throughout the different institutions.

The CountrySTAT approach is based on the application of data and metadata standards of FAOSTAT and SDMX (Statistical Data and Metadata Exchange promoted by IMF, WB, UNSD, EUROSTAT, FAO, OECD, BIS and ECB) and GAUL (Global Administrative Unit Layers). The web-based system has been developed since May 2004 using PX-Web at FAO Headquarters. Many countries have shown interest and are adopting it into their national statistical system. Furthermore, CountrySTAT is accompanied by a capacity-building strategy at country level to make the system sustainable in the long-term. CountrySTAT is networking with FAOSTAT and other sister information systems like GIEWS workstation.

2.3.7.4 - NADA

NADA (Ethiopia National Data Archive) is a service provided by the Central Statistical Agency of Ethiopia (CSA) to the data user community. The NADA Data Catalog allows browsing and searching of CSA's catalog of surveys. Survey overviews, Survey specific web sites, PDF reports providing detailed metadata of the surveys, data Access policies, and Data request form are included in the NADA system.

2.3.7.5 - GeoNetwork

GeoNetwork opensource is a standardized and decentralized spatial information management environment, designed to enable access to geo-referenced databases, cartographic products and related metadata from a variety of sources, enhancing the spatial information exchange and sharing between organizations and their audience, using the capacities of the internet. This approach of geographic information management aims at facilitating a wide community of spatial information users to have easy and timely access to available spatial data and to existing thematic maps that might support informed decision making.

Maps, including those derived from satellite imagery, are effective communicational tools and play an important role in the work of various types of users:

Decision Makers: e.g. Sustainable development planners and humanitarian and emergency managers in need of quick, reliable and up to date user-friendly cartographic products as a basis for action and better plan and monitor their activities.

GIS Experts in need of exchanging consistent and updated geographical data. Spatial Analysts in need of multidisciplinary data to perform preliminary geographical analysis and reliable forecasts to better set up appropriate interventions in vulnerable areas.

The main goal of the GeoNetwork opensource software is to improve the accessibility of a wide variety of data, together with the associated information, at different scale and from multidisciplinary sources, organized and documented in a standard and consistent way.

The challenge is to enhance the data exchange and sharing between the organizations to avoid duplication, increase the cooperation and coordination of efforts in collecting data and make them available to benefit everybody, saving resources and at the same time preserving data and information ownership.

FAO and WFP, and more recently UNEP, have combined their research and mapping expertise to develop GeoNetwork opensource as a common strategy to effectively share their spatial databases including digital maps, satellite images and related statistics. The three agencies make extensive use of computer-based data visualization tools, known as Geographic Information System (GIS) and Remote Sensing (RS) software, mostly to create maps that combine various layers of information. GeoNetwork opensource provides them with the capacity to access a wide selection of maps and other spatial information stored in different databases around the world through a single entry point.

GeoNetwork opensource has been developed to connect spatial information communities and their data using a modern architecture, which is at the same time powerful and low cost, based on the principles of Free and Open Source Software (FOSS) and International and Open Standards for services and protocols (a.o. from ISO/TC211 and OGC).

2.3.7.6 – Price Indexes

CSA has price index databases and these are available through the website. The pages allow retrieving consumer and producer price data with easy steps.

2.3.7.7 - Adobe PDF

CSA has adopted PDF to streamline document management and reduce reliance on paper. They are using it as the standard format for the electronic document management and dissemination of output of all surveys.

The PDF formats are generated by ICT development department. Reports and questionnaires are available in PDF format for surveys over Internet. This format is used to keep file format by preserving the fonts, images, and layout of source documents created on a wide range of applications and platforms. PDF is the standard for the secure, reliable distribution and exchange of electronic documents and forms. Adobe PDF files are compact and complete, and can be shared, viewed, and printed by anyone with free Adobe Reader software.

The PDF documents can be opened either in Acrobat or in a web browser. In Windows, users can configure their web browser to open PDF documents. All PDF documents which were made available to users over the website do not have a feedback collection form which is an electronic-based document that can collect data from a user and then send that data via email or the web to CSA.

This implies that electronic document generating staffs are not making use of some important features of the dissemination system. For instance, all PDF documents do not have feedback forms in the body of the document.

2.3.8 - Consistency

Establishing a Visual identity and applying it throughout the website is important. The branding of a website can be established by incorporating common design elements such as colors, logos, styles, etc., into every page. This presents a professional and consistent visual identity. Internal data systems should be consistent with the main layout of the website which it lacks currently. The NADA pages looks completely of other website than part of CSA's website.

2.3.9 – Information flow

Updates and changes are made to the current website in a random manner. The information flow currently used is, the website administrator will receive the information to be updated somehow and it will be updated. There is no timely update like for news and publications. There should be roles and responsibilities for content authors and content managers of the websites.

2.3.10 - Roles and responsibilities

Web site staff must have a clear understanding of how the agency publishes online information. Those staffs who are responsible for releasing and approving web content should know their roles and responsibilities in this capacity. Extending beyond this, other staff in the agency should know and be aware who to contact, either individuals or the relevant department for requesting content to be placed on, deleted or altered on the agency's web site(s). Content authors (contributors) should be able to prepare content to be published and notify the content manager (website administrator) to be published.

2.3.11 - Search Engine Optimization

It is an essential issue to optimize your website to make it search engine friendly to get more traffic. Search engine optimization techniques include current technologies, including Google site submission, XML sitemaps, and Meta tag information's. The current website does not have 'keyword' and 'description' Meta Tags.

3 - Recommendation

Through careful analysis and usability testing of the design of CSA's website, I recommend what I believe would be effective modifications to the website. These recommendations would increase the site's usability and appeal. Website redesigning is important to improve the website's overall appeal and functionality to its visitors and to help the website optimization process.

3.1 - Overall look and presentation

As a data source website I recommend the website to look simpler with less animation and graphics. If the user is presented with a simple yet catchy layout, their experience of browsing through your website either to query data or looking for information will be much more exciting. Download speed should also be kept in mind while designing the website as most of the users complain about it. Optimization of Query time of the RDBMS system should also be considered.

3.2 – Static website verses dynamic website

Depending on the type of information put online as well as how you manage your online content, you may either opt for static HTML documents or for an integrated database solution. Currently CSA's website is organized with static html pages. Static HTML pages are usually easier to manage at the initial stage and seem to be suitable for small websites having limited number of pages and having static content which is updated not more than once or twice a year. With larger Web sites, the complexity of issues is aggravated, and therefore dynamic page servers (as well as content management systems) may be of great importance. Also, if you have a relatively consistent but regularly changing section on your site, it may be a better option to go for dynamic (database driven) pages. A database driven website with a Content management System is vital for CSA's website as the data is to be updated constantly.

3.3 – A Content Management System

A content management system (CMS) is a system used to manage the content of a Web site. Typically, a CMS consists of two elements: the content management application (CMA) and the content delivery application (CDA). The CMA element allows the content manager or author, who may not know Hypertext Markup Language (HTML), to manage the creation, modification, and removal of content from a Web site without needing the expertise of a Webmaster. The CDA

element uses and compiles that information to update the Web site. The features of a CMS system vary, but most include Web-based publishing, format management, revision control, and indexing, search, and retrieval.

Some advantages of a Content Management System are stated below:

- **Browser based**

  Users are able to update and maintain their content from any computer connected to the Internet.

- **Designed for nontechnical users**

  People with average knowledge of word processing can create and maintain content. HTML coding is optional.

- **Consistent site design**

  Authored content is stored in a database separate from the site design so pages from all users are presented on the public website with the same consistent look and feel.

- **Automatically generated navigation**

  Site menus are generated stored in a database and web administrators can re-arrange the way they want it to be or can remove or add a new navigation menu depending on their needs.

- **Database content storage**

  Central storage means that content can be reused/featured in many places on the website and formatted for any device, including, web browser, documents and printer.

- **Dynamic features**

  Modules such as site search, news management, photo gallery, members' only pages, newsletter management, polling systems and others can be added to the website to provide advanced functionality and services not possible with a static website.

- **Daily content updates**

  Content modifications and additions can be made without involving a web agency or programmer for every update. The owner of the website is in control of the website.

- **Content scheduling**

  Content is time-controlled with expiration dates set for automatic removal for such entries as events and job postings to keep the site current without requiring user intervention.

- **Review and activation**

  Content and pages can be saved and published later to allow for works in progress and editorial reviews and approvals.

Open-source CMS solutions offer lower upfront, implementation, customization, and maintenance costs than proprietary CMS systems because of the way the software is developed. By leveraging the expertise of developers worldwide, open-source vendors deliver a better product at lower cost. I recommend Joomla CMS to be implemented for CSA's website as it is one of the best and has the largest number of skilled personnel in the country and internationally.

3.4 Content Management System for CSA

It is essential to have a CMS to gain full control of a website the easy way. Currently CSA's website is not using any CMS. To implement a CMS will ease the process of managing the website. A web administrator will be trained on the way of development so it will not be difficult to have the transition. Besides that to update the current website (static HTML pages), it requires knowledge of HTML language, where as if a CMS is implemented knowledge of HTML is an advantage but not a requirement. Any person with basic word processing skill could update the website.

3.5 – Overview about some CMS's

3.5.1 - Drupal

Drupal is a free and open source Content Management System (CMS) written in PHP. It is used as a back-end system for many different types of websites, ranging from small personal blogs to large corporate and political sites. It is distributed under the GNU General Public License.

The standard release of Drupal, known as Drupal core, contains basic features common to most CMSs. These include the ability to register and maintain individual user accounts, administration menus, RSS-feeds, customizable layout, flexible account privileges, logging, a blogging system, an Internet forum, and options to create a classic brochureware website.

Drupal was also designed to allow new features and custom behavior to be added by third parties. For this reason, Drupal is sometimes described as a content management framework. Although

Drupal offers a sophisticated programming interface for developers, no programming skills are required for basic website installation and administration.

Drupal can run on any computing platform that supports both a web server capable of running PHP version 4.3.5+ (including Apache, IIS, Lighttpd, and nginx) and a database (such as MySQL) to store content and settings.

3.5.2 – Joomla

Joomla! is a content management system platform for publishing content on the World Wide Web and intranets as well as a Model–view–controller (MVC) Web Application Development framework. The system includes features such as page caching to improve performance, RSS feeds, printable versions of pages, news flashes, blogs, polls, website searching, and language internationalization. It is written in the PHP programming language and uses the MySQL database system to store information. Joomla is free and open source software. Joomla specializes in managing web content with a simple What-You-See-Is-What-You-Get interface so even a non-technical person can post articles and images

Joomla is an open source content management system referred by tens of thousands as the best CMS in the world. It is written in PHP for publishing and managing content on the web and intranets using a MySQL database. Joomla includes features such as page caching to improve performance, web indexing, RSS feeds, printable versions of pages, newsflashes, blogs, polls, calendars, website searching and more.

Once installed sections, categories, content items, polls and other extensions can be setup. When new content is created, a WYSIWYG (What You See Is What You Get) editor allows for simple edits without the knowledge of HTML. While it is not mandatory the user have knowledge of HTML, it is recommended to engage in a few HTML primers in the event that something needs to be edited on a more custom level.

3.5.3 – TYPO3

TYPO3 is a free and open source content management system written in PHP. It is released under the GNU General Public License. It can run on Apache or IIS on top of Linux, Microsoft Windows,

OS/2 or Mac OS X. Along with a set of ready-made interfaces, functions and modules TYPO3 has a large repository of extensions, allowing flexibility and extendibility. More than 3900 extensions are available for download under the GNU General Public License from a repository called the TYPO3 Extension Repository, or TER.

TYPO3 has a web frontend, which presents a TYPO3 based website to its users, along with a web based backend, used by authors and site administrators to manage content for the website. TYPO3 can run on Apache or IIS on top of Linux, Microsoft Windows or Mac OS X. It uses PHP 4 or 5 (starting with TYPO3 4.2.0 all releases will require at least PHP 5.2) and Any database system supported by the TYPO3 DBAL including MySQL,Oracle, PostgreSQL and others. The system can be run on any web server, with a modern CPU and 256 MB RAM. The frontend can be displayed in browsers such as Mozilla Firefox on any OS, with JavaScript

3.5.4 – PHP-Nuke

PHP-Nuke is a web-based automated news publishing and content management system based on PHP and MySQL. The system is fully controlled using a web-based user interface. PHP-Nuke was originally a fork of the Thatware news portal system.

PHP-Nuke was originally released under the GNU General Public License as free software however, versions after 7.5, are pay for use.

PHP-Nuke requires a web server which supports the PHP extension (such as the Apache HTTP Server), as well as an SQL database (such as MySQL, mSQL, PostgreSQL, ODBC, ADABAS, Sybase or InterBase).

PHP-Nuke is a content management system allowing webmasters to create community-based portals (websites), in which users and editors can post news items (user-submitted news items are selected by editors) or other types of articles. Modules can be added to the PHP-Nuke system allowing additional features such as an Internet forum, Calendar, News Feed, FAQ's , Private Messaging,Content, Downloads, Encyclopedia, FAQ, Feedback, Forums, Journal, News, Search and others. PHP-Nuke supports many languages and its look and feel can be customized using the Themes system, but drastic changes requires knowledge of PHP, HTML and CSS.

3.5.5 – SilverStripe

SilverStripe is a free and open source content management system (CMS) for creating and maintaining websites. The CMS provides a web-based administration panel that enables users to make modifications to parts of the website. This panel includes a WYSIWYG website editor.SilverStripe is written in PHP5, and takes advantage of PHP5's object-oriented design capabilities. It is based on a model-view-controller pattern and uses an object-relational mapper. The CMS generates markup using a custom template language. SilverStripe contains Sapphire - a custom PHP framework.

SilverStripe is extensible through modules, widgets, themes, customization and ModelAdmin. SilverStripe generally relies on a code customization model over configuration.

Modules extend the core functionality of SilverStripe. Some existing modules include: Blog, Advanced Workflow Management, eCommerce, Forum, LDAP/OpenID authentication. Modules are available from the SilverStripe modules repository.

SilverStripe Widgets are small pieces of functionality that can be dragged and dropped into SilverStripe modules (notably the blog module). Examples of widgets include: tag clouds, flickr photos, or word of the day.

The SilverStripe themes directory provides a number of community-contributed, freely available themes. These themes can be quickly added to most SilverStripe sites.

Modules, Widgets, and Themes are all available as free downloads under the BSD license, and the majority of them are community contributed.

3.5.6 – Frog CMS

Frog CMS simplifies content management by offering an elegant user interface, flexible templating per page, simple user management and permissions, as well as the tools necessary for file management.

Frog requires a MySQL database or SQLite 3 with PDO, a web server (Apache with mod_rewrite is highly recommended) and is distributed under the MIT software license.

Frog is unique because of its simple templating code. Because it uses PHP directly, there is no need to learn yet another scripting language. This approach has two main advantages:

3.5.7 – MODx

MODx is an open source CMS that helps you take control of your website. It makes updates easy, empowering end-users with as much control as you desire over website content and update frequency.

Techies call MODx a Content Management System (CMS for short). It's also a pretty schwanky Application Framework. A robust and flexible API and an event override system makes building engaging web projects and changing core functionality without hacking its code! a breeze.

Not only does MODx help you build sites fast, but it also hides its tremendous power unless it's needed. As far as end users know, MODx is just an easy-to-use CMS with tons of freely available resources and one heck of an end user community.

3.6 – Criteria for selecting a CMS

- **Ease of use** – How easy it is for non-technical person to use it
- **Extensibility** – the ability to extend to have different functionality
- **Maintainability** – The availability of experts in the country to maintain, expand or upgrade in the future.

3.7 - Recommended CMS

A content management system should be easy to use for a non-technical person. Looking at the CMS's I found that Drupal, Joomla and frog CMS to be simple to use. When looking for extensibility having a significant base of ready-to-go plug-ins and components such as forums, photo galleries, contact forms, directories and newsletters ready to go without a lot of custom web development and with a significant wide range of community, **Joomla is the best choice.** Joomla's wide adoption has meant that a lot of web developers have created a lot of free and open source plugins for use with the content management system and now the extensions that can be built have expanded dramatically because of the improved API built into Joomla 1.5. Joomla has always adopted the principle of extensibility, meaning that it is very easy to integrate additional functionality into the system, but the 1.5 version of Joomla offers a significantly expanded API so that developers can communicate with the Joomla Core system much easier and make better use of its functionality.

Joomla is popular for various reasons. Mainly, the environment is very flexible which meets the needs of different user accordingly. It has completely database driven site engine. Each and every element in Joomla is stored in a centralized database. The popularity and flexibility of Joomla modular architecture has made it possible to develop a wide range of components and associated modules. Joomla has built-in SEO (Search Engine optimization) functionality. It is also supported by a community of thousands of developers around the world. Joomla have a security team, and a bug squad comprised of core contributors that actively monitor the bug reports, and turn around security fixes

Furthermore Joomla is widely used in our country and has the large number of experts available nationally and internationally.

3.8 - Access control and Security

As a web based system CSA's website can be accessed by users from different locations at a time throughout the world. However, users without a privilege should not access some of the website area. Access control restricts users from accessing a part of the website. When the administrator wants to use the backend of the system, he/she should first access the backend of the website and s/he must login to the system with a user name and a password and submit it. The system then checks for the validity of the user name and password by looking up in the database. If the user is found to be valid, he/she will be allowed to login to the system. After this the system provides a menu for the customer. Once the menu is displayed, the user can select whatever he/she wants from the menu and continues to use the functionality that is provided by the system. And whenever the user wants to terminate the system he/she needs to click the Logout link.

Access Control Matrix of the Joomla Framework

Back -End	Manager	Administrator	Super Administrator
Site	x	x	x
Control Panel	x	x	x
User Manager		x	x
Media Manager	x	x	x
Global Configuration			x
Logout	x	x	x
Menu	x	x	x
Menu Manager		x	x
Menu Trash		x	x
Main Menu	x	x	x
Content	x	x	x
Article Manager	x	x	x
Article Trash		x	x
Section Manager	x	x	x
Category Manager	x	x	x
Front Page Manager	x	x	x
Component	x	x	x
Banners	x	x	x
Contacts	x	x	x
News Feeds	x	x	x
Polls	x	x	x
Search	x	x	x
Web Links	x	x	x
Extensions		x	x
Install/Uninstall		x	x
Module Manager		x	x
Plugin Manager		x	x
Template Manager			x
Language Manager			x
Tools		x	x
Read Messages			x
Write Message			x
Mass Mail			x
Global Check-in		x	x
Help	x	x	x
Joomla! Help	x	x	x
System Info	x	x	x

Joomla is very secure script, all core extensions use very safe ways of getting user input data. The core Joomla security team posts updates every time so updating and taking backup on a timely basis is also recommended. While Joomla itself is inherently safe, wrong configuration, vulnerable components, poorly configured server, and weak passwords can all contribute to the security issue. Backend of the website will only be limited to the website administrator. Physical Security of the web server is currently available so it will not be a security issue.

3.9 – The Proposed Solution

To start with the recommended CMS a new interface design will be consistent all through the website. After the design is ready and approved the next step will be implementing the CMS. Implementation of the CMS with the new approved design will give the basic functionality of the CMS. Navigational flow, creating menus, and categorizing contents will be done after the implementation. After categorizing content, the CMS will be integrated with the data libraries. Next adding different functionalities to the CMS will be done. The proposed solution will then be deployed.

Maintaining and managing the proposed solution will be done by CSA IT staff which will be trained at the same time of the development. A manual/guideline on how to manage and maintain the website through the CMS will also be prepared so any CSA staff will be able to update the website. Some section of the website content might also be updated and maintained by different CSA sections. Updating and maintaining the website does not require sophisticated technical knowledge of IT, but anyone with a basic computer skill could do it. In the future upgrading the version of the CMS as available and applying different patches is recommended and this will also be included on the manual/guideline.

3.10 - Enhancing functionality

Joomla has core feature which give the website great functionalities. Some of the features are Contact Management, Menu Manager, Media Manager, Banner Management, Content manager, polls. Although the Joomla CMS has the major features needed for the website, I recommend additional features to be integrated with the CMS to gain more functionality and usability. The following are some of my recommendations.

3.10.1 - Newsletter management

In its simple meaning, a newsletter is a form of advertising that is sent periodically. It is a regularly distributed publication that is mostly about one main topic, and is of interest to its subscribers. Newsletters are delivered via email. Newsletters are normally used by owners of the websites to communicate and advertise themselves to their readers or subscribers.

An online newsletter has many advantages. One of the main advantages is that sending a newsletter via email is cost effective e.g. you can reach a very broad audience instantly with very little cost. Newsletters are an effective communication tool. Creative newsletters can be appreciated by the subscribers and hence increase the communication between organization and users.

Users will be able to subscribe to the newsletter through a link on a home page which they will put their email address to be included on the subscribers list. They also could unsubscribe through the link from the home page or through a link sent to them with every newsletter.

3.10.2 - Document management

Currently CSA's website has lots of PDF documents. Documents should be arranged and categorized so users could find it easy to search documents available. This system will arrange documents in different category and users could filter documents easily.

3.10.3 - Vacancy management

This system will enable CSA to post available vacancy to the general public in an organized way.

3.10.4 – Commenting system

This system will enable users to comment on articles of the website. This could be administrated by the website administrator from the back end to publish or remove comments.

3.10.5 – Discussion Forum

This feature will manages discussion forums for the site. Users could discuss on different topics. An Internet forum, or message board, is an online discussion site which will add the website more interactivity to give it a web2.0 feature.

3.11 – Integration with the Digital Libraries

Digital Libraries are useful part of the CSA website as most of the users are visiting the website to get data from them. The recommended CMS is capable of displaying any external system or website in

its wrapper component. Integration of these digital libraries will be seamless without affecting the consistent look of the website. I have looked at the **Survey Data RDBMS** System being developed, the latest **countrySTAT** site, the latest **NADA** pages and **Ethio-INFO**. These systems will be displayed within the new CSA website. Adding any Newly designed system into the website will also be possible from within the CMS.

Currently there are different data sharing and presentation systems implemented and on the way to be implemented. These systems have their own purposes and presents different types of data in different manner. For example the NADA system is an interface to search the surveys which are presented with the IHSN system. The survey results are accessible in PDF format on the website but the NADA system gives users who need more than the survey report with a detailed metadata about the survey. EthioInfo/DevInfo allows presentation of data through Tables, Graphs and Maps. CountryStat, a system on the way to be implemented, also allows presentation of Data through tables, Graphs and Maps. Yet again another database system, to present the survey results through the Relational Database (RDBMS) is under way.

User analysis of the usage of some of these systems is conducted depending on the website log analysis. NADA/IHSN and DevInfo were on the website for some time now, and the analysis shows that the NADA system is used often. As some of the systems are not implemented yet, user opinion and usage analysis could not be compared to determine the accessibility of the systems. I recommend implementing those systems and get the user feedback on the systems. The IT staffs at CSA are currently analyzing DevInfo and CountryStat systems as they present data somehow similarly and using both systems might be a redundancy.

Managing these systems might be difficult for CSA and it might also be frustrating for website users to go through all the systems. The CMS will arrange these different data libraries under one menu and making the environment consistent, which gives the user more comfort. Selecting one system for each kind of presentation (i.e. surveys, tables, and GIS data) is recommended and this needs thorough analysis on each database systems. Each system has different functionalities, as also contains different types of data. The CMS will handle website contents and will serve as a common environment for these systems.

3.12 - Backup system

Backups are useful in case of data loss, if your security has been compromised by malicious attackers, or simply if you made an error and need to revert the site back to a previous state. There needs to be a periodical back up system. The backup system could be of two types. The database backup and file system backup.

3.13 - Hardware and software consideration

Necessary hardware and software for the realization of the system includes:

3.13.1 - Web server

A web server is currently available at CSA. The server currently runs IIS configured to run PHP.

3.13.2 - Database

3.13.2.1 – MySQL

MySQl is a relational database management system (RDBMS) which has more than 6 million installations. MySQL stands for "My Structured Query Language". The program runs as a server providing multi-user access to a number of databases.

The project's source code is available under terms of the GNU General Public License, as well as under a variety of proprietary agreements. MySQL is owned and sponsored by a single for-profit firm, the Swedish company MySQL AB, now a subsidiary of Sun Microsystems, which holds the copyright to most of the codebase.

MySQL is commonly used by free software projects which require a full-featured database management system, such as WordPress, phpBB and other software built on the LAMP software stack. It is also used in very high-scale World Wide Web products including Google and Facebook.

Performance of MySQL is very fast for common DB operations. We can enjoy large community support for MySQL. In fact nearly every problem we face has been seen by

someone else which makes it literally great. The latest version 6+ has triggers, replication, and high-availability on UNIX with heartbeat, stored procedures and views.

3.13.2.2 – MSSQL

Microsoft SQL Server is a relational model database server produced by Microsoft. Its primary query languages are T-SQL and ANSI SQL.

The code base for MS SQL Server (prior to version 7.0) originated in Sybase SQL Server, and was Microsoft's entry to the enterprise-level database market, competing against Oracle, IBM, and, later, Sybase itself. Microsoft, Sybase and Ashton-Tate originally teamed up to create and market the first version named SQL Server 1.0 for OS/2 (about 1989) which was essentially the same as Sybase SQL Server 3.0 on Unix, VMS, etc. Microsoft SQL Server 4.2 was shipped around 1992 (available bundled with Microsoft OS/2 version 1.3). Later Microsoft SQL Server 4.21 for Windows NT was released at the same time as Windows NT 3.1. Microsoft SQL Server v6.0 was the first version designed for NT, and did not include any direction from Sybase.

About the time Windows NT was released, Sybase and Microsoft parted ways and each pursued their own design and marketing schemes. Microsoft negotiated exclusive rights to all versions of SQL Server written for Microsoft operating systems. Later, Sybase changed the name of its product to Adaptive Server Enterprise to avoid confusion with Microsoft SQL Server. Until 1994, Microsoft's SQL Server carried three Sybase copyright notices as an indication of its origin.

3.13.2.3 - PostgreSQL

PostgreSQL is a powerful, open source relational database management system. It has been in development for more than 20 years and has a strong reputation for excellent architecture and world-class reliability, data integrity, and correctness.

PostgreSQL is an enterprise-class database that boasts sophisticated features, such as Multi-Version Concurrency Control (MVCC), point-in-time recovery, tablespaces, asynchronous replication, nested transactions (savepoints), online/hot backups, a sophisticated query planner/optimizer, and write-ahead logging for fault tolerance. It supports international

character sets, multi-byte character encodings, and Unicode, and it is locale-aware for sorting, case sensitivity, and formatting. PostgreSQL is highly scalable, both in the quantity of data it can manage and in the number of concurrent users it can accommodate.

PostGreSQL is a high-end Oracle like database, in fact at installation of the instance, it requests the type of installation you need (oracle type or ...), the web based monitoring and logging is excellent. It is really a high-end of the currently available free RDBMSs. As of recent time it is ready for windows. It is truly open source RDBMS.

3.13.2.4 - Some limitations of the databases

Database	Max DB size	Max Table size	Max row size	Max columns per row	Max Blob/Clob size	Max CHAR size
Microsoft SQL Server	524,258 TB (32,767 files * 16 TB max file size)	524,258 TB	8060 B (SqlServer 2000)	1024	2 GB	8000 B
MySql 5	Unlimited 2 GB	2 GB (Win32 FAT32) to 16 TB (Solaris)	64 KB	3398	4 GB (longtext, longblob)	64 KB (text)
PostgreSQL	Unlimited	32 TB	1.6 TB	250-1600 depending on type	1 GB (text, bytea) - stored inline	1 GB

3.13.2.6 - Proposed Database

The proposed CMS, Joomla, is developed specifically using MySQL database. Some queries in the code are specific for Mysql. Making Joomla work for other database at the moment consists of looking through all queries, checking which ones are Mysql specific and change them to others type or generic ANSI SQL. This is not recommended as it creates instability on Joomla. Joomla is tested for stability on MySQL. I recommend using **MySQL** for joomla.

3.14 - Performance Characteristics

- Speed constraints E.g. Connection speed – recommended to upgrade the connection speed to at least 1MB/s.
- Power interruption should not keep the website unavailable. Uninterruptible power supply should be configured to sync with the generator.

3 - Referred Documents

- ***Government of Federal Democratic Republic of Ethiopia - Web Site Standards and Guidelines***
  By Ethiopian Information and Communication Technology Development Agency (EICTDA)

- ***Web Content Accessibility Guidelines 1.0***
  http://www.w3.org/TR/WCAG10/

- ***Feasibility study for the development and implementation of RDBMS***
  By Birru Dorri

- ***Referred websites***

  - *www.wikipidia.com, www.opensourcecms.com, www.w3.org, www.bls.gov, www.statcan.gc.ca,* data.un.org and statistical office websites of Finland, Netherlands, South Africa, UK, Canada, USA and others