

GA-A26400

**CUSTOMIZABLE SCIENTIFIC WEB-PORTAL
FOR DIII-D NUCLEAR FUSION EXPERIMENT**

by
G. ABLA, N. KIM, and D.P. SCHISSEL

APRIL 2009



DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

GA-A26400

CUSTOMIZABLE SCIENTIFIC WEB-PORTAL FOR DIII-D NUCLEAR FUSION EXPERIMENT

by
G. ABLA, N. KIM, and D.P. SCHISSEL

This is a preprint of a paper to be presented at the 17th International Conference on Computing in High Energy and Nuclear Physics, March 21–27, 2009, in Prague, Czech Republic, and to be published in the *Proceedings*.

Work supported in part by
the U.S. Department of Energy
under DE-FC02-01ER25455

**GENERAL ATOMICS PROJECT 30106
APRIL 2009**



Customizable scientific web-portal for DIII-D nuclear fusion experiment

G Abla, E N Kim and D P Schissel

General Atomics, PO Box 85608, San Diego, California 92186-5608, USA

abla@fusion.gat.com

Abstract. Increasing utilization of the Internet and convenient web technologies has made the web-portal a major application interface for remote participation and control of scientific instruments. While web-portals have provided a centralized gateway for multiple computational services, the amount of visual output often is overwhelming due to the high volume of data generated by complex scientific instruments and experiments. Since each scientist may have different priorities and areas of interest in the experiment, filtering and organizing information based on the individual user's need can increase the usability and efficiency of a web-portal.

DIII-D is the largest magnetic nuclear fusion device in the US. A web-portal has been designed to support the experimental activities of DIII-D researchers worldwide. It offers a customizable interface with personalized page layouts and list of services for users to select. Each individual user can create a unique working environment to fit his own needs and interests. Customizable services are: real-time experiment status monitoring, diagnostic data access, interactive data analysis and visualization. The web-portal also supports interactive collaborations by providing collaborative logbook, and online instant messaging services.

The DIII-D web-portal development utilizes multi-tier software architecture, and Web 2.0 technologies and tools, such as AJAX and Django, to develop a highly-interactive and customizable user interface.

1. Introduction and motivation

The DIII-D Tokamak [1] is the largest experimental nuclear fusion research facility in the US. The mission of the DIII-D Research Program is to establish the scientific basis for the optimization of the tokamak approach to fusion energy production. The majority of experiments are carried out from the DIII-D main control room, which is a large hall equipped with tokamak monitoring, control, data acquisition and analysis devices, as well as up to 60 computer terminals. In a normal operation day, an experimental team, which consists of physicists, operators and engineers, carries out experiments. While most of the team members are located in the control room, it is common that some scientists lead or actively participate in experiments remotely. There are approximately 500 researchers worldwide who actively participate in DIII-D research.

The everyday fusion experiment is cyclical in nature. The DIII-D tokamak operates in a pulsed mode producing plasmas of 5 to 10 second duration every 10 to 15 min, with 25 to 35 pulses per operating day. Scientists analyze what happened during a pulse by looking at acquired and analyzed data and making decision on what hardware/software plasma control adjustments should be made as

required by the experimental science for the next pulse. The decision-making is a collective process that multiple scientists look at different data and different aspects of the experiment.

There are two ways for utilizing experimental run time more efficiently. The first way is to make machine status information, raw and analyzed data available for scientists as soon as possible. The second way is to provide a reliable and effective communication and collaboration mechanism among distributed team members. There are tools available which individually satisfy both points above and while some of them try to address data visualization and experiment status monitoring needs, others improve communication among members of the scientific team. However, each of those tools is often a standalone application and several applications need to be running to analyze, monitor and communicate at the same time. These general tools lack of customizable user interface. The user experience for scientists sometimes can be very inefficient since there is a need to switch multiple application windows from time to time in order to look at multiple data and collaborate with others.

Recent advances in internet technologies have made it possible to create a highly interactive web-based application. A group of interrelated web development techniques, named AJAX [2], paved the way to mimic feature-rich desktop applications on the web interface. A variety of tools and frameworks of Web 2.0 [3] also made the design and development of web-based interaction and collaboration software a relatively small effort. Therefore, a web-based portal for data analysis, experiment status monitoring and information sharing during the fast paced DIII-D research experiment seemed appealing.

While portals have provided a centralized gateway for multiple computational services, the amount of visual output often is overwhelming due to the high volume of data generated by complex scientific instruments and experiments. Since each scientist may have different priorities and areas of interest in the experiment, presenting information without filtering and organizing may hinder the effectiveness of the portal. This problem can be solved by making the portal customizable so that individual users can choose the application and layout of the portal based on their specific needs and interest. The DIII-D scientific web-portal is designed in a way that it can be customized in different ways. The layout of the portal, the number and position of the applications as well as the information filter in each application can be customized by users.

This paper describes the design and implementation of the scientific web-portal for DIII-D nuclear fusion experiment. In the section 2 the three-tier layered software architecture of the web-portal is discussed. Section 3 presents the applications in the web-portal and technologies behind them. Section 4 describes the current deployment status and related discussion, and finally followed by section 5 with future work.

2. Software architecture of the DIII-D web-portal

The DIII-D web-portal design utilizes multi-tier software architecture. It is mainly composed of three tiers: Data tier, logic tier, and presentation tier (figure 1).

2.1. Data tier

The data tier provides multiple data sources generated by the DIII-D tokamak and research activities of scientists. There are two types of data providers. The first type of data providers are data servers. The second type is tokamak control and monitoring systems that can act as data sources.

Data servers provide a variety of experimental data and meta data. Some of the data acquired directly from the DIII-D tokamak and others are generated by data analysis codes after each pulse. Meta data is data about data, which describes location, type and availability of data from tokamak operations. Some meta data also comes from the electronic log book application. The meta data is stored in "Relational DB" server and the experiment related data is stored in the "Raw and Analyzed Data" Server.

The tokamak control and monitoring systems are not specialized data storage. However, they provide data about real time machine status, data analysis status and control room activity. The real-time control servers provide the hardware control requests and results as well as tokamak status information. The media server streams audio and video data from control room operations. Analysis

monitoring system provides real time status information about computational resources and between pulse data analysis codes.

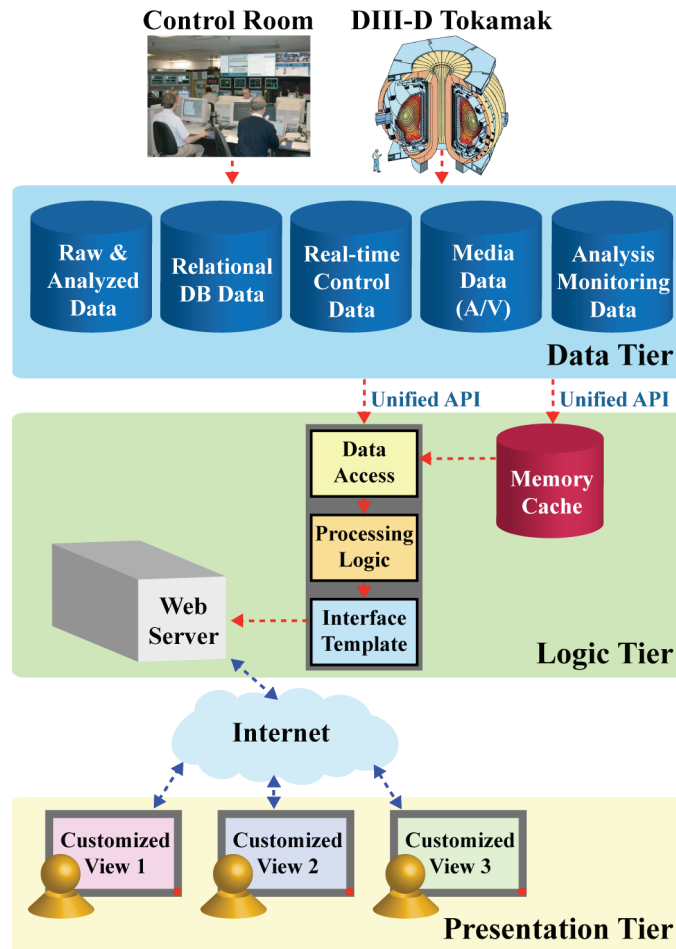


Figure 1. Multi-tier software architecture of DIII-D web-portal.

All the above data sources together provide data needs of the web-portal system. According to data characteristics, there are three different modes for serving data from the data tier. One mode is request/response, in which data is only served when requested. Another mode for serving data is subscription mode, in which the data tier provides the logic tier with a continuous data stream if subscribed. The last mode is spontaneous mode that triggers certain actions in the Logic Tier due to the data changes in the data providers. The three modes of communication use three corresponding standard APIs for data exchange.

2.2. Logic Tier

The logic tier is responsible for receiving data from the data tier, processing it and generating the web interface for the presentation tier. It consists of multiple “pluggable” application components and provides all the functionalities of the web-portal. Each function of the web-portal is represented with unique a component (figure 2).

The python-based Django framework [4] is used for the logic tier. Django is an open source web application framework, which loosely follows the model-view-controller design pattern [5]. The Django framework has three major parts: Data access, interface template, and processing logic. In the Django world these three parts are called model, template and view.

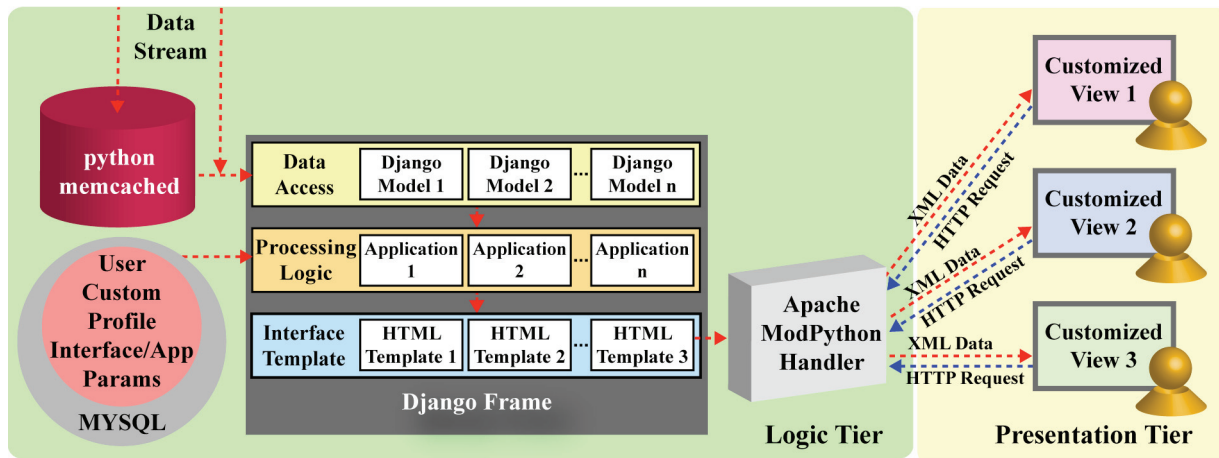


Figure 2. Detailed logic tier and presentation tier.

The logic tier also maintains a memory cache which is used for two purposes in the logic tier. The first one is caching data in order to minimize the number of data requests to the data tier for non-real-time data. The second application of the memory cache is using it for fast data sharing between multiple data sources and logic tier components. The software, memcached [6], is a high-performance, distributed memory object caching system and used for the memory cache of the logic tier.

2.3. Presentation tier

The presentation tier is the part of the web-portal that interacts with its users. It is responsible for generating requests to the logic tier and receiving corresponding presentation information. The requests from the presentation tier to the logic tier are simply described with HTTP GET and HTTP POST methods. The types of requests are: 1) Data viewing and visualization requests, 2) control requests, and 3) web-portal customization requests.

The presentation tier directly communicates with the users to offer a personalized working environment. The AJAX technology provides a unique and efficient way to observe all changes made on the client-side and record them from the server-side. These customization options include turning applications on/off, changing locations or updating parameters on a specific application. A list of all available applications and services are provided as links to independently turn them on or off. Once enabled, an application can easily be moved around the page by drag-and-drop (figure 3). When the location is redefined, it gets calculated and updated on the server-side. Some applications also offer their own customization which is discussed further in section 3. When a change is made on the client-side, JavaScript makes a HTTP POST request to update to the “User Custom Profile” database. This database includes multiple tables which share a unique key defined by Apache’s authentication credentials. The web-portal is protected by Apache’s basic authentication with secure HTTP. All users are required to provide a username and a password in order to gain access. Once the credentials are received, all entries identified with the username is selected and processed as the logic tier calculates the output for the user. Because all changes are recorded and saved on the server-side, the web-portal will keep the user’s customization in place even after the client session ends.

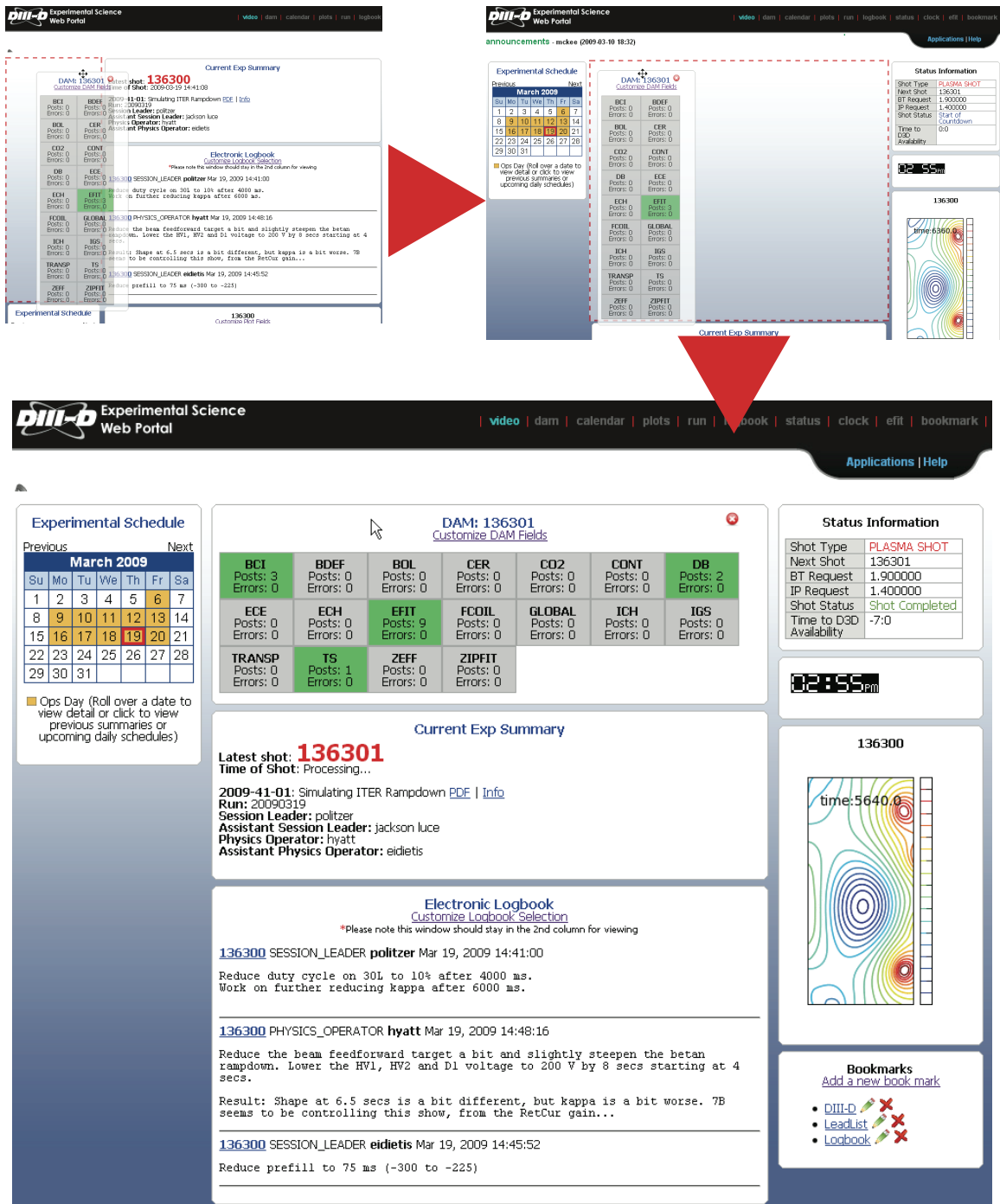


Figure 3. Application can easily be moved with drag-and-drop.

3. Applications

The web-portal is a library of applications and services that cover a wide variety of aspects of the DIII-D research program. Currently, there are 11 applications on the web-portal:

1. *Current Experiment Information* gives a summary of the experiment such as experimental team and details of the proposed experiment.
2. *Data Analysis Monitoring* outputs results of analysis codes in real-time.
3. *EFIT Animation* displays an animation of the most recently calculated plasma shape data.

4. *Electronic Logbook* provides a way of organizing notes for the experimental team
5. *Experimental Schedule* provides a detailed overview of all scheduled experiments.
6. *Signal Plots* outputs the latest plots of various measured and analyzed data.
7. *Tokamak Status Information* displays pulse information such as the configurations for the pulse as well as the status and the type.
8. *Video* provides a real-time view of the DIII-D control room.
9. *Bookmarks* provides a way for users to save websites.
10. *Clock* displays the current time.
11. *Instant Global Announcement* makes announcements available to all web-portal users.

The main goal of the web-portal is to present the large amount of visual data in a user-friendly and straightforward manner, hiding the complexity from the client-side. Part of hiding this complexity is that applications retrieve data using source specific methods and then process the data as required. In the rest of this section, we will describe how this is done in several applications of DIII-D web-portal.

The experimental schedule collects information from multiple databases in order to produce a calendar view of all experiments. The processing logic in the logic tier utilizes Python's `pymssql` class to connect, execute queries and fetch the information from the MSSQL databases. Once the information is received, it parses and processes the data into an HTML table. It configures each cell with a background color for operating days, formats details as a tooltip, and attaches a link to summaries of previous experiments or detailed schedule information on upcoming ones. This allows users to easily view all the needed schedule data in a small window just over 2 inches by 3 inches.

The Data Analysis Monitor (DAM) is a monitoring system used at DIII-D to detect discrepancies in the results from data analysis codes [7]. Existing systems such as DAM can easily be configured into the web-portal as long as the needed data is passed to the processing logic of the logic tier in a recognizable scheme. DAM acquires and analyzes the raw data produced by various data acquisition computers and saves the result in a data storage system. As the new updates are made to the storage system, DAM passes the values to memcached in the logic tier. The data access reads the new values and passes them to the processing logic that creates the color coordinated HTML-based output for users to easily view the data.

Because DAM provides information on around 40 different analysis codes, the output can be overwhelming and maybe not be suitable for all users. To satisfy the many different needs and interests of its users, the web-portal provides a customization (figure 4) for each user. The customization window provides a list of all available data analysis codes for users to select from. This information is stored in the "User Custom Profile" which is kept even after the client session is over.

After each pulse is completed, scientists always inspect a variety of raw and analyzed data to further interpret the experiment results. Currently, the most used visualization methods are 2-D time-based signal plots and a plasma shape contour plot, which can give a quick overview of the experiment. The DIII-D web-portal automatically creates the needed graphical images immediately after each pulse. For this purpose, the logic tier component is subscribed to DAM's data service to check if the needed data is available for the new pulse. If it receives a "data available" notification from DAM, it sends a request to "Raw and Analyzed Data Server" to fetch the signal data. The data will be read and 2-D data plots as well as plasma shape contour movie will be generated for the new pulse. Once the plot images and movie are generated on the web server, the web-portal picks them up for output. It reads a given list of available data, picks up only the ones specified by the user, and outputs them to the customized interface.

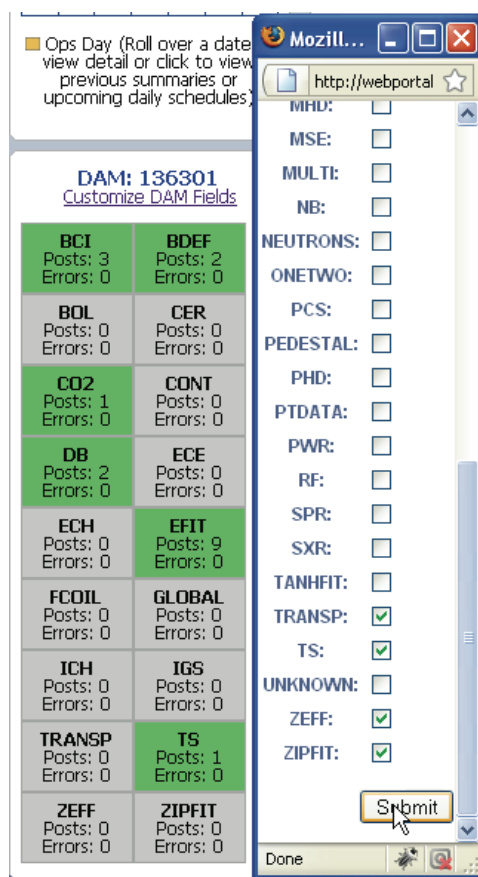


Figure 4. Data analysis monitoring application customization.

4. The initial deployment status and discussion

The DIII-D web-portal has been deployed during the year 2009 experimental campaign and used for two months. More than 60 scientists have viewed the web-portal with approximately 25 using it on a routine basis to monitor the ongoing experiment activities. The number of web-portal users not in the DIII-D control room exceeded the number of those in the control room. An informal survey found that most of the users agree that the customization capability provides a convenient user experience. Additionally, the scientists expressed that the web-portal is valuable for enhancing their science yet as is typical they desire even more applications and features.

During the development and improvement process we realized that the web-portal application has several benefits over regular desktop applications. The first benefit is the short deployment time, since there is no need to install any client software on the users' computers. The second benefit is the easy solution for modifying and improving the application, since the users always have access to the most-up-to-date version without any client installation effort. Therefore, it is significantly less intrusive to make minor enhancements, as no action is required by the user to obtain the new version.

The one problem with the customizable web-portal is that it is hard to know a priori the average computational load on the server and network bandwidth requirements. Constant monitoring is required during the initial deployment phase to insure that the server hardware and network are sufficient to meet the demand.

5. Future work

During the initial deployment period, users have already requested new features and provided ideas for extending the web-portal. As the next step, more applications such as instant messaging and visualization sharing capability will be added. A finer grained customization of portal layout also will be implemented as requested.

Another plan for the DIII-D web-portal is to improve its efficiency and scalability. Currently, the DIII-D web-portal uses a web browser's "client pull" mechanism, not the web server's "server push", to make data updates on most of the applications. This can be very inefficient as the number of users logging into the web-portal increases. Therefore, the "server push" mechanism will be implemented and the unnecessary network and server CPU load will be minimized. As the number of applications on the web-portal increase distribution of the Logic Tier computational load will be considered.

Template management will also be improved. Currently, the web-portal mainly targets web browsers as clients. However, the template system can be more versatile and it is possible to dynamically create presentation templates for clients other than web browsers. One such client is a handheld device such as smart phones. Recently handheld devices have become very versatile and it is certainly possible to develop client applications and interfaces which can be utilized on these platforms. Another client that will be considered is large public displays. The large US fusion experimental control rooms have display walls that use homegrown applications to display visual data from experiments [8]. While it is possible to display web pages on the display walls, the optimum color and size for visual information as well as the interaction method on the display walls are different from normal desktop screens or smaller devices. The web-portal on the display wall will be investigated. Taken together, the web-portal can serve multiple clients with minimal development.

Acknowledgment

This work was supported by the US Department of Energy SciDAC program under Cooperative Agreement DE-FC02-01ER25455.

References

- [1] DIII-D, <http://fusion.gat.com/global/DIII-D>
- [2] What is AJAX?, <http://www.riaspot.com/articles/entry/What-is-Ajax->
- [3] What is Web 2.0, <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
- [4] Django project web site, <http://www.djangoproject.com>, visited on January 22nd, 2009
- [5] Trygve M H Reenskaug, MVC XEROX PARC 1978-79, <http://heim.ifi.uio.no/~trygver/themes/mvc/mvc-index.html>
- [6] Memcached, <http://www.danga.com/memcached/>, visited on January 22, 2009
- [7] Flanagan S, et al., 2004 System to monitor data analyses and results of physics data validation between pulses at DIII-D *Rev. Sci. Instrum.* **75** 2043; DOI:10.1063/1.1753668, May 2004.
- [8] Abla G, et al. 2005 Shared Display Wall Based Collaboration Environment in the Control Room of the DIII-D National Fusion Facility, published in the *The Fifth Annual Workshop on Advanced Collaborative Environments*.