Delivering Real-time Satellite Data to a Broader Audience*

Lan Zhao, Shuang Wu, Rakesh Veeramacheneni, Carol X. Song, Larry Biehl
Rosen Center for Advanced Computing
Purdue University
West Lafayette, IN 47907
{lanzhao, rveerama, carolxsong, biehl}@purdue.edu

ABSTRACT
This paper presents our work of enhancing the existing cyberinfrastructure using Web 2.0 technologies for delivering real time satellite data to a broader audience. As a resource provider on the TeraGrid, Purdue University hosts several data collections for earth and environmental research, including a large number of satellite data products that are received and generated in real time at the Purdue Terrestrial Observatory (PTO). A science gateway portal, named PRESTIGE, has been developed as the access point to services related to the satellite data generated at PTO. It remains a major challenge to help potential users easily discover the data products that are available and access them based on their needs. In an effort to address this challenge, we developed different interfaces for different user groups. We use the Google gadget as a new way to deliver satellite data and information to a broad user community. Three satellite data viewer gadgets have been developed and deployed. They connect to a data generation and processing system at the backend. The gadgets embrace the recent Web 2.0 technologies in bringing the rich access and sharing capabilities available in the general cyberspace to the scientific user communities.

Categories and Subject Descriptors
H.3.3 [**Information Storage and Retrieval**]: Online Information Services – *Data sharing, Web-based services.* H.5.2 [**Information Interfaces and Presentation (e.g., HCI)**]: User Interfaces – *User-centered design.*

General Terms
Design

Keywords
Remote Sensing Data, Satellite Data, Real-time Data Delivery, Data Service, User Interface, Google Gadget.

1. INTRODUCTION
With the rapid advancement of data collection and storage technologies, vast volumes of data are being generated daily from sensors, observatories, instruments, models and simulations. Data providers continue to face the challenge in managing the growing data volumes efficiently and reliably and making them easily accessible by users. Existing research and commercial systems, along with modern high performance storage systems, helped to address this issue. As a research center in a university, we also face the challenge of reaching out to the largest possible audience who may be interested in the data collections available with limited cyberinfrastructure resources, to go way beyond the traditional method of word-of-mouth through research collaboration or today's popular build-a-portal and they-will-come model.

This project leverages the Google Gadget and Google's infrastructure that hosts the gadgets, and utilizes this new Web 2.0 technology in making real-time satellite data available on the internet. While a gadget itself is not complex, its interaction with the backend processing system was designed to ensure real-time data delivery. In developing the gadgets, two Purdue undergraduate students were able to complete the gadget development and integration with the larger satellite data processing and management system under the guidance of researchers at the Purdue Rosen Center for Advanced Computing. Not only do the satellite data viewer gadgets help bring the highly complex data products to the average internet users in a simple, easy-to-understand rendition and interface, the development process has been a rich learning experience that exposed the students to a much larger context related to remote sensing data and processing.

2. BACKGROUND
As one of the eleven resource providers on the TeraGrid, Purdue University hosts several data collections in support of earth and environmental research. One of these data collections is the remote sensing satellite data from the Purdue Terrestrial Observatory (PTO) [1]. PTO is an interdisciplinary remote-sensing research facility that hosts two satellite ground stations capable of continuous ingestion of remote sensing data from the NOAA GOES-12, NASA MODIS Terra/Aqua, NOAA AVHRR, and Fengyun-1D MVISR sensor systems. Up to 30GB of raw data are received and processed each day. A large number of satellite data products for multiple disciplines of land and atmosphere studies can be generated from the raw remote sensing data. For example, the data from the MODIS sensor aboard the Terra and Aqua satellites can generate 14 different products. Each data product may contain up to 50 different variables. These data products, derived from the satellite observations, describe features of the land, oceans and atmosphere and are used to study processes and changing trends in real-time environmental remote sensing research and applications for both the local and global scales. Remote sensing data are widely used in both research and applications, including atmospheric science and climatology, hydrology, agriculture, ecology, transportation and logistics,

Lan Zhao, Shuang Wu, Rakesh Veeramacheneni, Carol X. Song, Larry Biehl, Delivering Real-time Satellite Data to a Broader Audience, International Workshop on Grid Computing Environments 2009, Nov. 2009, Portland, OR, USA.

aviation and air traffic safety, and disaster and emergency management.

To help users access the wide variety of satellite data products at PTO, we have developed PRESTIGE, a science gateway that allows users to subscribe to satellite data products of their interest [2]. Behind the scene, PRESTIGE automatically configures the data production back-end system to perform the appropriate data processing algorithms on the raw satellite data, generate the specified products, and deliver them to the user in a way he/she chooses. In this paper, we describe the recent development effort in enhancing the existing capabilities of the PRESTIGE portal using Google Gadget – a Web 2.0 technique - as a powerful tool to engage users and increase awareness of the data collections and services available at the Purdue TeraGrid site.

Web 2.0 represents a new set of technologies such as AJAX, Google gadgets and social networking applications. It drives a new generation of web applications with an emphasis on online collaboration and information sharing among users. For example, there are 40 Google gadgets that have more than 500,000 users, and more than 200 Facebook applications that have more than 500,000 active monthly users [3, 4]. As reported by the official Google gadgets API blog (http://googlegadgetsapi.blogspot.com/2007/10/its-been-long-time-coming.html) there are billions of gadget pageviews each week. Google gadgets provide an effective and powerful platform for rapid dissemination of scientific tools and services to a wider audience, including K-12 and college students.

Google gadgets are small self-contained web applications that are developed using XML, HTML, and JavaScript. They can be hosted by Google or on your own server. Users can dynamically insert gadgets into a web site such as iGoogle, Google Maps, Orkut, or any web page where people wish to view information for different services in one place. Google gadgets also support building a collaborative gadget, with very simple code changes, where users can share a gadget and collaborate on the same gadget data. Various levels of access can be set for recipients for view only, or view-and-edit. Furthermore, Google Gadget can be combined with OpenSocial [10] and quickly spread across different social networks and therefore reach a broader audience and connect people with scientific applications and information.

Gadgets have been primarily used for news feed aggregation or small non-scientific applications. Recently, gadgets have gained momentum as a new way of delivering scientific applications and datasets. The Open Grid Computing Environment (OGCE) has developed a number of gadgets for both the Open Life Science Gateway (OLSG) and the Social Informatics Data Grid (SIDGrid) [5,6,7]. In OLSG, gadgets are provided for running bioinformatics tools such as BLAST as well as for status tracking. In SIDGrid, OpenSocial gadgets can be dynamically created based on XML description files of SIDGrid applications defined by social and behavioral scientists. OGCE also developed GTLAB tag component libraries which can be used in gadget development.

In this project, we developed three satellite data viewer Google gadgets. They serve as an interface between end users and the data acquisition and processing backend connected to the satellite ground stations. The satellite data viewer gadgets deliver real time satellite data in a continuous animation of images through a compact web interface. A user may select from a list of products, e.g., cloud cover, water vapor, etc., and a time period (e.g., 2 hours, 24 hours). Behind the scene, a video consisting of an animation of the images for that period is generated and streamed to the user. It changes the way data is served to the users: the data is pushed in near real-time to the user's desktop or his favorite web page, instead of waiting for users to come to the data provider, such as the PRESTIGE science gateway for PTO satellite data. The gadgets are linked back to the PRESTIGE web site so that users may learn more about the various products and subscribe to new products through an easy-to-use web interface.

The rest of the paper is organized as follows: Section 3 describes the design of the backend and frontend systems. Section 4 describes in greater detail the design and implementation of the satellite data viewer gadgets. Section 5 describes future work and concludes this paper.

3. SYSTEM DESIGN

In order to better serve the needs from different users, we divide the users for the satellite remote sensing data into two groups: (1) Researchers and high level education users who need to define new data products and use them in their research or class. They typically need to assimilate the data into models and tools for analysis and visualization. The PRESTIGE portal and its backend services are designed to meet the needs of this group of users. The portal provides a user-oriented, interactive, and web-based platform for delivery of the PTO satellite data. It allows users to define new data products from a set of satellite data streams received by the ground station, and subscribe to the data products which will be delivered with minimal delay. (2) General public, such as school teachers and children, who are interested in seeing popular satellite data products in an easy-to-understand format. This group may not have extensive knowledge about the data products. A typical request is the display of the latest data directly which was not available in the PRESTIGE portal. Another limitation of our satellite data system is the lack of an effective way of publicizing the available data and services to the public.

The main goal of this project is to address the needs from general users and overcome the dissemination barrier using Google gadgets. It leverages the technology and infrastructure of Google gadgets for the promotion and delivery of satellite data to a large user community. In the meantime, these gadgets hide the functional complexities and present simple interfaces to the users. They are easy to use and meet the needs of users at the starting point of a learning curve.

The overall system architecture is shown in Figure 1. The main components of the system include satellite receiver, data processing cluster, data repository, web server, gadget container and the gadget directory hosted by Google. A workflow is created to dynamically chain the service pipelines that include data receiving, generation, delivery and processing. Externally the user-centered Web 2.0 enhanced gadget interface provides several options for data type and duration. The end-to-end workflow includes the following steps:

1. The receiving station at PTO acquires real-time data from satellites. For example the GOES-12 satellite sends down images almost continuously from its geostationary position at 75 degrees West over the equator.

2. The PTO data processing cluster connects directly to the satellite receiver at the ground station. The receiving station starts sending the image to the processing cluster soon after a

new image download starts. The system leverages an existing data production system responsible for processing the raw streaming satellite data and converting it to advanced data products. The PTO data processing system runs a proprietary TeraScan software system from SeaSpace [6]. TeraScan's daemon process monitors the incoming satellite data stream, scans the product configuration files, and creates batch jobs to generate the derived data products. The batch jobs are scheduled and executed using a 10-node Linux cluster managed by a PBS batching system. The batch job is configured to automatically push the output data product to a data repository at the end.

3. On the web server, a periodic task reads the latest data from the data repository, attaches a timestamp with each satellite image, and generates a video animation for the last N hours of data where N is 2, 12, 24 hours for GOES-12 data and 7 days for MODIS data.

4. When a user loads the satellite data viewer gadget, the browser sends a request to the gadget container which is responsible for rendering the gadget XML and JavaScript. After the user selects a satellite data product and duration, the pre-generated video is streamed to the user from the Apache server.

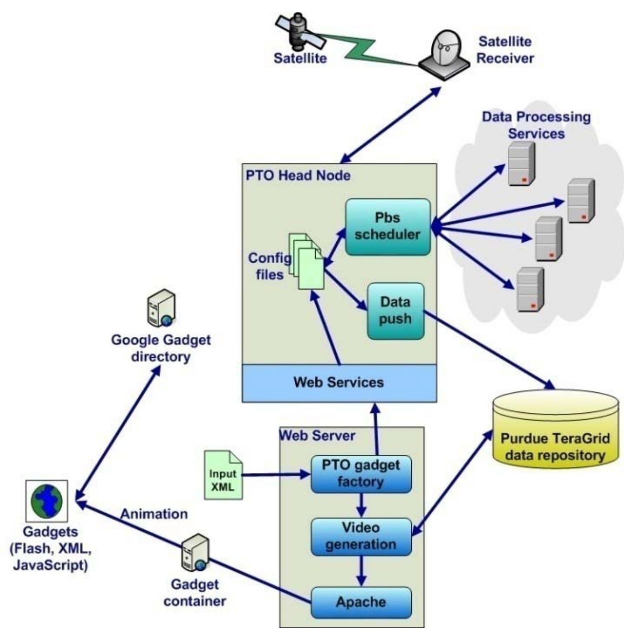

Figure 1. System architecture of satellite data viewer gadgets

Currently a small number of satellite data products are enabled by the PTO data processing system and accessible from the gadget viewers. To create an extensible system so that new data products can be easily added to the gadget, a gadget factory component is being developed to help automate the process. This component takes an XML input file which specifies information about the new data product to be enabled, including satellite name, sensor, product name, duration and repository path. The gadget factory component then invokes a web service interface running on the PTO data processing server which configures the backend system to automatically generate the specific data product and push it to the repository. It also configures the video processing component on the web server to automatically fetch the latest data and generate live animations.

Three gadgets have been published to the Google Gadget directory which users may search for by keywords. Apart from Google gadgets, the same web application system could easily be adapted to other Web 2.0 technologies. For example, similar web gadgets can be developed using the Facebook application framework and added to the Facebook application directory [4]. Our system architecture supports a wide range of mobile technologies that could put scientific data and visual representation into the hands of millions of people with a modest level of infrastructure support at the data provider.

4. Satellite Data Viewer Gadgets

As discussed in Section 3, a satellite data viewer gadget consists of two parts: On the server side, a cron job preprocesses the latest satellite images and generates a video animation for the latest data in a preconfigured time period for each of the data products supported by the gadgets. On the client side, the gadget allows users to select a product of interest and duration. A JavaScript dynamically determines the name of the video that matches the user's requirement and requests it from the web server. Since the animated videos are generated offline and the same video can be shared by multiple users, the system is resource-efficient and responds quickly to the users. The scalability of this solution is mainly determined by the performance of the web server to stream the videos. The videos are generally a few megabytes in size, in which case thousands of concurrent users can be supported by modern web servers such as Apache2 [8].

Because each gadget is small in size in order to fit into the gadget container, the data resolution and level of user control in web gadgets are limited. Nonetheless, this light-weight web application serves the needs of the majority of entry-level users ranging from personal interests on real-time weather related events to learning. These web gadgets can easily be integrated into course materials, learning and teaching modules, and school online resources.

All gadgets use Google Gadget Editor (GGE), and thus are hosted by Google. The gadget API serves as a medium or window for applications. We have developed two versions of the gadgets using different technologies. The first version is based on JavaScript and HTML. JavaScript along with majority of the code are inside the gadget xml file. The second version is based on a Flash application. The gadget code simply embeds the application, and communicates to it via JavaScript while the rest is done by Flash.

4.1 HTML based gadget

As for the user interface design, the main restriction is the limited screen real estate allocated for gadgets on an iGoogle page. As a result, we keep the gadget UI as compact as possible, e.g., using multiple tabs to group functions. As shown in Figure 2 three tabs were used for the gadget: Setting, Video, and Help. The *Help* tab is shown first to display instructions and contact information. The *Setting* tab keeps track of user selections and allows the user to toggle between datasets. This is done via *userprefs* in the gadget API. The *Video* tab contains play/pause, zoom, and frame-by-frame controls. The frame-by-frame feature is only possible natively with videos produced for Flash v5 or older versions. For each data product, the user can choose between two video resolutions: by default low resolution images are displayed in the video which fits in a two-column or three-column iGoogle layout. The user can also set the viewer to show the satellite data at higher resolution. In this case it is recommended to embed the gadget in a customized web site.

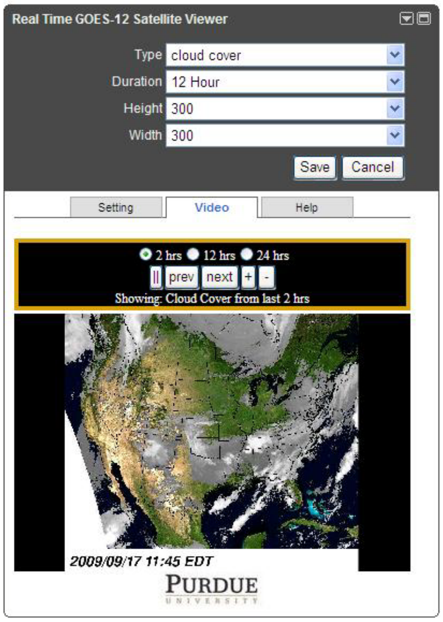

Figure 2. HTML based GOES-12 data viewer

4.2 Flash gadget

While developing the first version of the satellite data viewer gadget using HTML and JavaScript, which is typical for gadget development, we encountered several problems related to inconsistent support for the user interface across different browsers. As a result, we explored using Flash for the next version of the satellite data gadget. Flash provides a richer environment for both gadget developers and end users, which is the main reason for our choice. Flash supplies a richer set of tools for designing and developing more capable interfaces. It also helps reduce the complexity introduced by cross-browser compatibility issues, common for web page programming.

We developed the second version satellite viewer gadget using the open source Flex Builder [9]. Similar to the first version, the biggest factor affecting the UI design is the space constraint. Gadgets need to be designed to fit inside a three-column display in an *iGoogle* page. Similar to the previous version, tabs are available for the user to select the video or help (Figure 3). We achieved several improvements in the UI. A popup menu allows a user to switch between datasets, making it easy to toggle video sources within the existing application space. Another advantage is that Flash allows one to implement the Google gadget in a much simpler way, reducing errors and development time. The Flash-based GOES-12 gadget allows users to choose different satellite data products from three time periods: the last two hours, 12 hours, and 24 hours. Users can fast forward, rewind, pause, resume, and zoom while the video is steamed. The flash animation can be easily downloaded if the user has *RealPlayer* installed and enables the *web download and recording* feature.

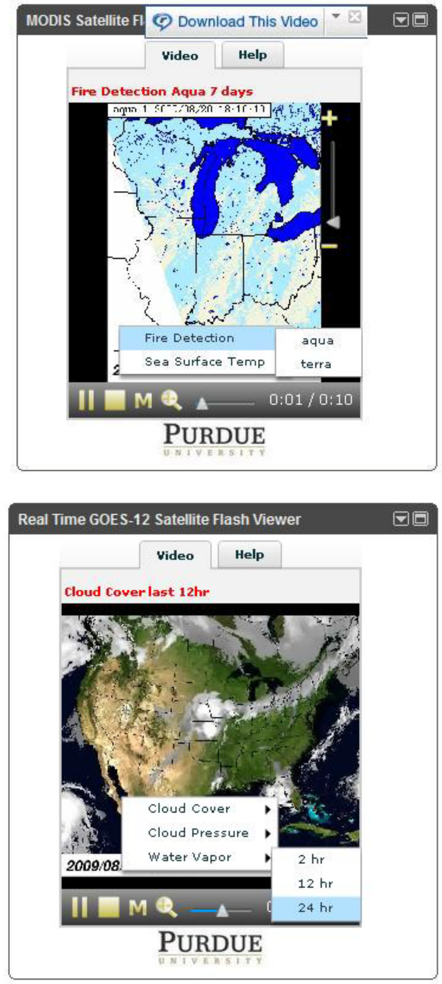

Figure 3. Flash based MODIS and GOES-12 data viewers

4.3 Usage Tracking

Gadget usage statistics can be easily tracked via Google Analytics. This feature tracks simple page views as well as specific interactions. However, analytics is not currently supported by the new gadget API and only the legacy API has built-in analytics functions.

Since we published the three satellite viewer gadgets for GOES-12 and MODIS in early July and August respectively, the gadgets have been visited 2665 times from 31 countries. A snapshot of the map view for the GOES-12 HTML viewer is shown in Figure 4. A summary of the usage for all three gadgets (HTML version for GOES-12, Flash version for GOES-12, and Flash version for MODIS) is listed in Table 1. As shown by the statistics, the average number of visits for the GOES-12 and MODIS gadgets are about 20 times per day.

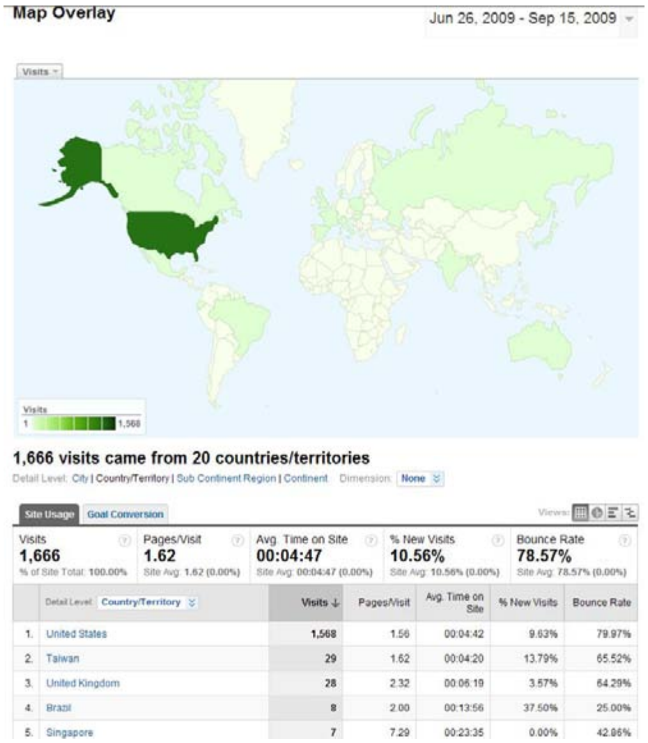

Figure 4. Usage statistics

Table 1. Statistics for the gadgets since they were released (As in 9/15/09)

Gadget Name	Release Date	Visits	PageViews	New Visits (%)	Avg. Visits Per Day
HTML GOES-12	6/26/09	1666	2697	10.51	22.3
Flash GOES-12	8/3/09	142	437	36.62	3.2
Flash MODIS	8/3/09	858	1120	5.94	19.5

4.4 Integration with PRESTIGE

As mentioned earlier, the PRESTIGE portal serves as a single access point for common functionalities related to the PTO remote sensing data. Along with the release of these viewer gadgets, we integrated them within the PRESTIGE portal interface. Since the gadgets can be found by any user searching the Google Gadget directory, e.g., by keywords such as *Satellite* or *TeraGrid*, they add a new discovery channel for the rich set of PTO satellite data products made available through the Purdue TeraGrid resource provider. On the other hand, user feedback through their interaction with the PRESTIGE portal, e.g., monitoring data products that are subscribed to most often, will help us determine the interest level from the user community and the next product to enable in a similar data viewer gadget.

5. CONCLUSIONS & FUTURE WORK

In this paper, we have described our effort in the design and implementation of three Google gadgets for streaming animations of the latest satellite data in near real-time. These gadgets are supported by a pipeline of services for satellite data receiving, data generation, transfer and image processing. The light-weight web applications enable interactive selection, generation, and delivery of the PTO satellite data products. Each of the gadgets is scalable to support multiple satellite data streams. It also provides flexibility to end users in selecting the products of their interest.

For future work, we plan to increase the number of data products available for the MODIS satellite. More specifically we plan to add the Normalized Difference Vegetation Index (NDVI) bi-weekly composite data product to the gadget which is commonly used to identify the presence of vegetation and its condition. Another product to be added is optical depth. As we gain experience in producing these gadgets through rapid development, we will automate the process of adding new data products to the gadget. We plan to explore the OpenSocial API to leverage the social network in data dissemination. Our experience with these data gadgets is also applicable to other projects, such as the visualization and access of NEXRAD Level II Doppler high resolution radar data as well as how to utilize gadget and similar technologies in scientific application development.

The focus of this paper is on the design of the system and its applicability to a real-world satellite data requirement, and to demonstrate a way for scientific data to reach to a very large user community with a modest level of infrastructure support from the university while leveraging the commercial infrastructure at Google. We plan to study the performance of this delivery system by investigating the effect of several factors including the number of users and products provided. We will also study ways to measure the effectiveness of these gadgets, for example, in bringing users to the PRESTIGE portal and the increase in access to the satellite data. We intend to publish the results of these studies in a future publication. As a TeraGrid data provider, this new satellite data delivery platform will encourage broader access by the general user community.

ACKNOWLEDGEMENT

This research is sponsored in part by the National Science Foundation under TeraGrid Resource Partners grant OCI-0503992.

REFERENCES

[1] PTO: Purdue Terrestrial Observatory. http://www.itap.purdue.edu/pto/

[2] R. Kalyanam, L. Zhao, T. Park, L. Biehl, and C. X. Song, "Enabling User-Oriented Data Access in a Satellite Data Portal", to appear in Proceedings of the 3rd International Workshop on Grid Computing Environments/Supercomputing'07, Reno, NV, November 2007.

[3] Google Gadgets. http://www.google.com/webmasters/gadgets/.

[4] FaceBook Applications: http://www.facebook.com/apps/

[5] W. Wu, K. Rajavenkateshwaran, T. Uram, M. Papka and R. Stevens, "A Web 2.0 Application Framework for Social

Informatics Data Grid", TeraGrid 2009 conference, Arlington, VA, June 2009.

[6] M. Pierce, S. Marru, W. Wu, G. Kandaswami, G. v. Laszewski, R. Dooley, M. Dahan, N. Wilkins-Diehr and M. Thomas, "Open Grid Computing Environment", TeraGrid 2009 conference, Arlington, VA, June 2009.

[7] W. Wu, R. Edwards, I. R. Judson, M. E. Papka, M. Thomas, R. Stevens, "TeraGrid Open Life Science Gateway", TeraGrid 2008 conference , June 9-13, 2008, Las Vegas

[8] Haddad, I., Butler, G., "Experimental studies of scalability in clustered web systems," *Parallel and Distributed Processing Symposium, 2004. Proceedings 18th International*, vol., no., pp. 185-, 26-30, April 2004.

[9] Adobe Flex: http://www.adobe.com/products/flex/.

[10] OpenSocial Specification, http://www.opensocial.org/.