

# SOFA – Search Optimized Flash Architecture

## Creating a searchable Flash Rich Internet Application

insight

Michael Scafidi | June 2007

razorfish.

Because for some sites, over 50% of traffic originates from search engines, one of the biggest challenges for Flash Web sites is their inability to be fully indexed. These sites also present a greater challenge for compliance with accessibility standards. We have been implementing **Search Optimized Flash Architecture (SOFA)** for our clients providing a solution to this challenge.

The solution provides:

- Searchable Flash Web site
- Accessible alternative Web site
- Links to content deep within the Flash site
- Enabled browser buttons for going back and forward within the Flash Web site.

### Problem:

Because Flash Web sites cannot be fully indexed by search sites such as Google and Yahoo, they can present challenges to accessibility.

### Solution:

As a corrective, develop the Flash Web site as a Flash application that receives its content from traditional XHTML.

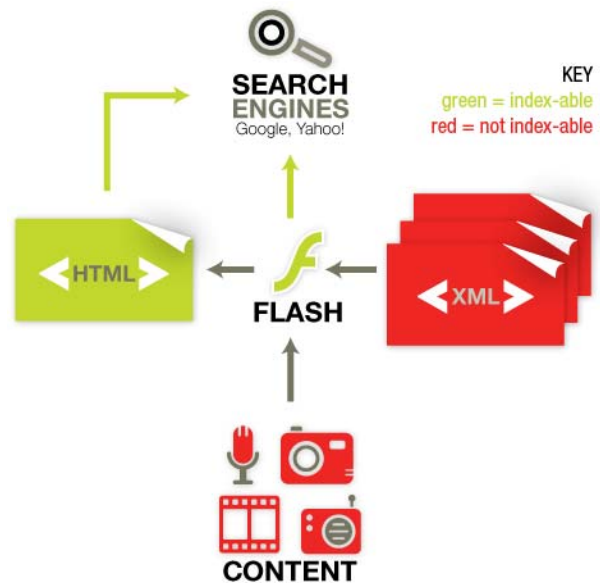
### Benefit:

With SOFA, the Flash Web site can be indexed by search sites and can support both accessibility standards and users without Flash.

## The Problem:

Officially a Flash file (SWF) is indexed by the latest search engines. The catch is that only content hard coded into the file is indexed and searched. Due to a modern site's content volume and maintenance needs, it is impractical to hard code all the content into the Flash file. Frequent content additions and more complex management processes add to the problem.

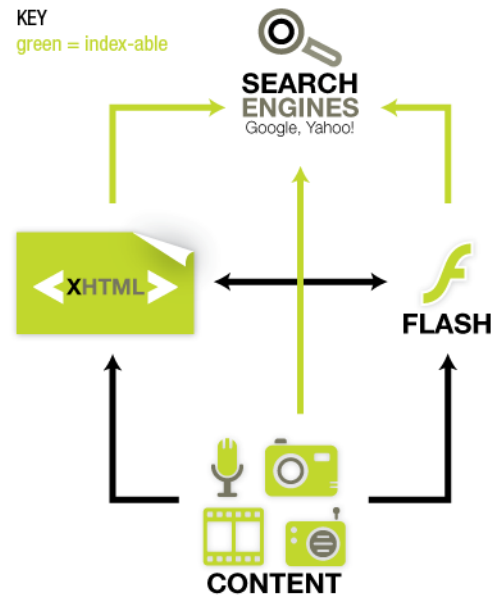
These needs have lead to externalizing the content so it is no longer hard coded in the Flash file. As more complex Flash Rich Internet Application (RIA) sites are created and integrated with content management systems, developers have been externalizing a site's content into XML files. Though this technique allows the content to be managed separately from the Flash application, the content in the XML file is not indexed by search engines. Because it is not indexed, any relevant content in these files will not show up in search engine results.



**Figure 1 - How typical Flash sites are indexed showing the exclusion of xml and media content**

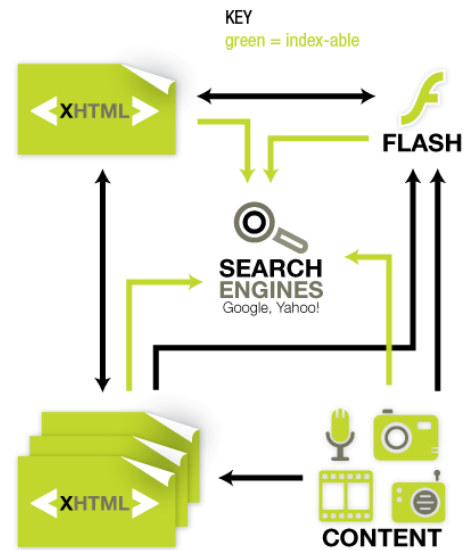
## The Solution:

In a standard Flash RIA implementation, the Flash file is placed on an HTML page, but Web standards have recently been moving HTML development into well-formed XHTML. This trend has occurred because traditional HTML is a type XML that is not well formed, meaning the syntax is not strict or rigorous. Poorly formed XML does not validate properly and can present errors when used. Now consider that an XHTML file is a well-formed XML file, the same type of file that developers are using to externalize content. The SOFA technique uses XHTML not only as the location of the Flash placement but as the Flash content source. A search engine indexes the content in the XHTML, thus indexing the content used in the Flash site. When a user finds the content in a search engine and links to the site, a check determines if the user has Flash and hides the XHTML content while displaying the Flash application with the corresponding content. If the user doesn't have Flash, the XHTML content can then be stylized using a cascading style sheet to present a non-Flash accessible page to the user.



**Figure 2 - SOFA facilitates the indexing of all content on a single page Flash site**

This process works for a single page RIA tool. For a Flash application with multiple deep pages, large quantities of content must be placed on the single XHTML page, which is impractical. The solution is to make multiple XHTML pages with an index page implementing a navigation list, including anchor links associating the additional content pages. As the search engine crawls the site, it will follow the linked content. When the Flash file incorporates the XHTML, the anchor links inform the Flash RIA which additional content is associated with which section of the Flash site. The navigation's lists are also incorporated by the RIA as the site's global navigation. By using semantic markup and other tags such as forms and their actions, lists and images, functionality and content can be identified for the Flash site to use. Additionally, IDs on tags within the XHTML and the idMap in Flash 8 ease the use of the incorporated content.



**Figure 3 - SOFA facilitates the indexing of all content on a multi-page Flash site**

Now content is available for indexing and searching. Once a user finds deeper content through a search engine and links to the deeper page, the Flash RIA loads and deep links to the associated deep content.

## Considerations:

The SOFA design pattern must be integrated into the fundamental code of the Flash RIA. Skinning, or changing the appearance of, the XHTML content for an HTML version of the site is optional and requires a more consistent XHTML implementation.

Additionally, search engine spamming or cloaking is a concern, a technique which occurs when Web sites present content to be indexed by Web pages but then does not show the content to the end user. The cloaking technique is used in an attempt to get a higher page ranking but is consistently condemned by Web guidelines, such as Google's Webmaster Guidelines, which states:

“Make pages for users, not for search engines. Don't deceive your users or present different content to search engines than you display to users, which is commonly referred to as ‘cloaking.’”

-Google Webmaster Guidelines

<http://www.google.com/support/webmasters/bin/answer.py?answer=35769#quality>

Yahoo! Search Content Quality Guidelines also condemns the cloaking practice:

<http://help.yahoo.com/help/us/ysearch/basics/basics-18.html>

Since the SOFA technique, after hiding the content on the page, does ultimately present it in the Flash RIA, it is not trying to deceive the end user and thus is **not cloaking** the content.

## Conclusion:

The SOFA technique is an evolution in Flash RIA development and represents a fundamental approach to solving the challenges in indexing these Web sites as well as making their content accessible. More limited solutions, such as Scalable Inman Flash Replacement (sIFR) developed by Shaun Inman, Mike Davidson and Thomas Jogin, allow for the indexing of some content in simple Flash pieces. Other approaches for consuming XHTML into Flash include Peter Hall's Ripple and Claus Wahlers' Search Engine Friendly Flash Site (SEFFS). The SOFA approach is a comprehensive solution for RIAs and exposes all content for indexing.

Razorfish has implemented the SOFA approach for Ford Motor Company's SYNC website ([www.syncmyride.com](http://www.syncmyride.com)), introducing the SYNC Voice-activated mobile phone and digital music system.

## About the Author

Michael Scafidi is the Senior Presentation Layer Architect and leads the presentation layer technology practice at Razorfish. He is located in the New York office and has been with the company since 1999. Michael's focus is to implement integrated Internet solutions by leveraging all aspects of the presentation layer including browser and server technologies to deliver aesthetic, robust, modular, maintainable, and measurable solutions. He graduated from the University of Rochester with a BS in Biology/Geology. Prior to joining Razorfish, Michael managed and developed the University of Rochester Medical School's online education system and classes in Web design and development.

Please send your comments and questions to [michael.scafidi@razorfish.com](mailto:michael.scafidi@razorfish.com)

## About Razorfish™

Razorfish is one of the largest interactive marketing and technology companies in the world, and also one of the largest buyers of digital advertising space. With a demonstrated commitment to innovation, Razorfish counsels its clients on how to leverage digital channels such as the Web, mobile devices, in-store technologies and other emerging media to engage people, build brand loyalty and provide excellent customer service. The company is increasingly advising marketers on Social Influence Marketing™, its approach for employing social media and social influencers to achieve the marketing and business needs of an organization. Its award-winning client teams provide solutions through their strategic counsel, digital advertising and content creation, media buying, analytics, technology and user experience. Razorfish has offices in markets across the United States, and in Australia, China, France, Germany, Japan and the United Kingdom. Clients--many of them served in multiple markets--include Carnival Cruise Lines, Coors Brewing Company, Levi's, McDonald's and Starwood Hotels. Visit <http://www.razorfish.com> for more information.

Razorfish  
821 2<sup>nd</sup> Avenue, Suite 1800  
Seattle, WA 98104  
Phone: 206.816.8800  
Fax: 206.816.8808