

Malicious content in enterprise portals

Introduction

In 2005, enterprise portals rank in the top 10 of CIO technology focus areas in many surveys¹. The main drivers of the portal business growth are the horizontal portal suites, which provide content management capabilities, application integration tools, and specific solutions for collaboration and knowledge management.

This paper will address the security problems an enterprise may have due to the various content management abilities in a typical portal implementation, and will focus on cross site scripting attacks.

Common characteristics of enterprise portals

An enterprise portal can either manage its own users, or import users from an existing user directory, like LDAP or Active directory.

The login procedure is usually via a form-based login screen, and the portal manages the session by standard means of cookies or URL session strings.

Content is added to a portal by content-entry templates and forms, by uploading files, by linking to external web content, and by crawling and indexing mechanisms.

Content entry templates

Similar to the content management capabilities found in content applications like PostNuke, PHPNuke, JetSpeed, Mambo, Plone and others, the portal content entry templates may be vulnerable to cross site scripting attacks, and therefore should be tested for XSS vulnerabilities. Most XSS problems reported so far, originate in this area of content management.

Uploaded files

All enterprise portals provide users with the ability to upload files and documents into a content area (depending on the users' permissions). These files include MS-Office documents, Acrobat documents, text files, HTML files, graphic files, and various binary files. Some of these files may present a security threat regardless of the portal environment. An uploaded file can usually be described by specifying meta-data properties such as the title, subject, description, author, keywords, and any other properties the enterprise decides upon. The uploaded files can then be searched online, either by meta-data search or by full-text search. The portal usually also provides a

¹ Gartner report, http://www.microsoft.com/office/sharepoint/prodinfo/gartner_may05.msp

way to browse the file's meta-data properties. If the file properties can be manually input during the upload process, then this provides an option to insert malicious JavaScript into the properties' values.

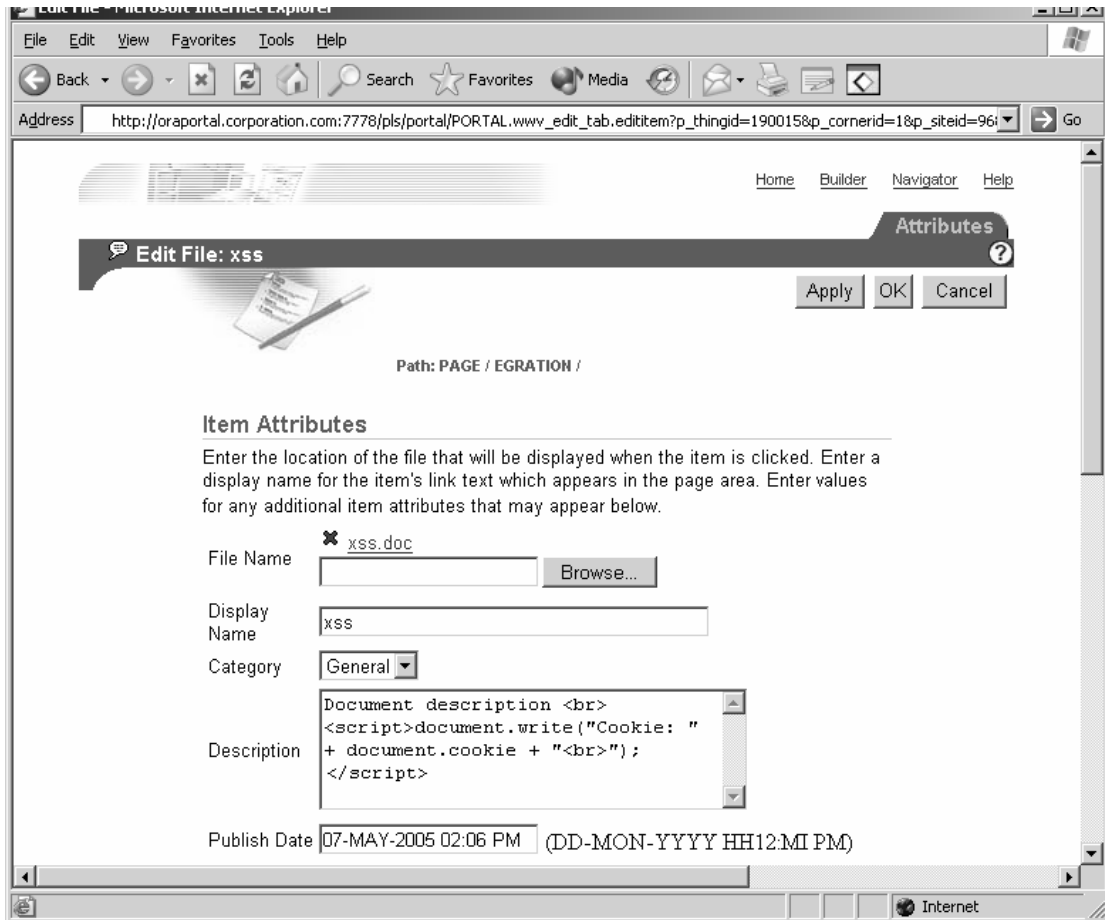


Figure 1: Oracle Portal - document description script insertion

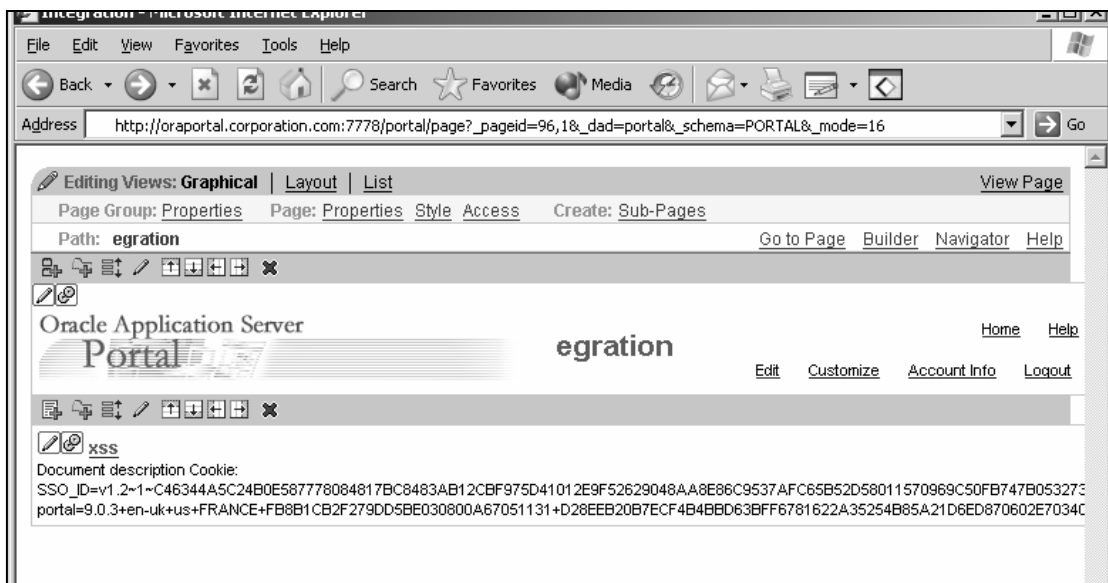


Figure 2: Oracle portal - document description script execution

Alternatively, some portals can automatically define some document attributes based on the existing document properties. In such a case, consider the following MS-Word document properties:

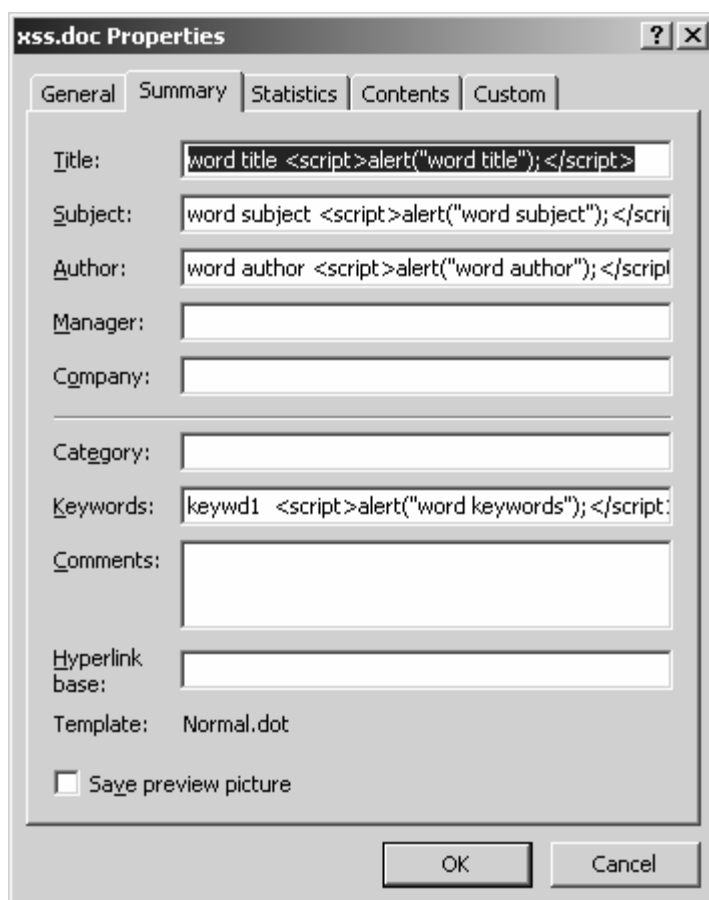


Figure 3: Poisoned word document properties

Or the following PDF properties:

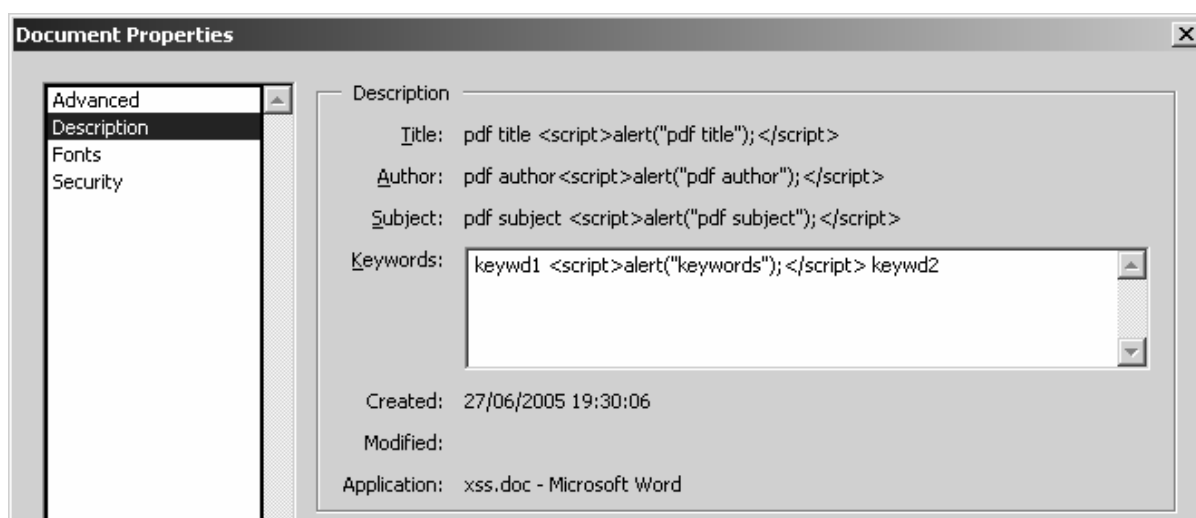


Figure 4: Poisoned Acrobat document properties

If the portal is aware of the built-in attributes of MS-Office and Acrobat, then it must be checked for cross site scripting attacks on the display of these attributes in the document's property sheet.

A similar attack can be directed at the document text itself. Some portals compile a document extract and display it as the description of the uploaded file. A document can be crafted to contain malicious JavaScript code that is obscured from reading or printing by setting the text color to be similar to the background color, or by hiding the code behind white rectangles². By proper placement of text, the JavaScript code will be displayed in a portal web page in the portal's context.

The most dangerous uploaded documents are of course HTML files. The entire document is displayed by the portal, and since the attacker is not limited to a short description text, the attack vectors are limited only by the browser client capabilities (they do have limits, don't they?). A cross site scripting attack in such a case is extremely easy.



Figure 5: Plumtree portal - Uploaded HTML file with malicious script

External web content

The building blocks of the enterprise portals' user interface are called portlets. The user sees a web page that contains one or more rectangles with content. Some of this content may come from sources external to the portal, either from the Intranet or from the World Wide Web.

When the external content portlet is based on an iframe, the attack possibilities are limited to those provided courtesy of the browser manufacturer.

However, some portals support gateways to external content. These gateways act in effect like reverse proxies, hiding the actual data source and providing standard portal context to the browser. On the one hand, this makes a lot of sense, for example when the portal is

² <http://news.zdnet.co.uk/internet/0,39020369,39197313,00.htm>

used to consolidate access to internal web sites, or when the portal contains caching mechanisms for remote sites. On the other hand, the gatewayed web sites are likely not to be within the control scope of the portal manager, and malicious XSS code is likely to be gatewayed as well. In figures 6 and 7 we see a document that originates in server prodauth02, but is served to the user via a gateway from server portal. In this case it is a PowerPoint presentation, but it could have been an html document with embedded javascript.

General Document Properties	
Property	Value
Name	December 15, 2005 New Product Roadmap Vision for BEAs Portal.ppt
Description	Presentation
Card Created	Dec 20, 2005 12:49:09 PM
Card LastModified	Mar 30, 2006 2:18:03 PM
Open Document URL	http://portal.plumtree.com/portal/server.pt/gatewa y/PTARGS_0_2_2805427_0_0_18/December%2015,%202005% 20New%20Product%20Roadmap%20Vision%20for%20BEAs%20 Portal.ppt
Plumtree Document Type ID	585
Card Content Language	English

Figure 6: A document URL in the user's eye

Keywords	BEA_BIP_Portal_Overview_simple_v3, BEA_BIP_Portal_Overview_simpl BEA_BIP_Portal_Overview_simple_v12, BEA_BIP_Portal_Overview_simp
Security	Employees-Only
Subject	BEA BIP Portal Overview Simple
URI	http://prodauth02.plumtree.com/ntcws/docfetch/DocF etch.aspx?path=%5C%5Cifile%5Cportal%5CEvents%5C200 5+Presentations%5CDecember+15%2C+2005+New+Product+ Roadmap+Vision+for+BEAs+Portal.ppt&signature=63270 2527040150000&locale=

Figure 7: Same document's real origin

Crawling and indexing

Some of the most hyped features of enterprise portals are the crawl and retrieve capabilities. An enterprise portal can act as an enterprise search engine. It can crawl and create indexes of web sites, NT and Novell file servers, databases, and email archives. Some portals have federated search abilities as well, allowing them to interact with applications that have their own internal search features. Like it does for uploaded content, the portal crawler will create an index for the crawled document, and will create a property sheet that may be poisoned by XSS code. Depending on the portal and the data source, the indexed documents may be presented to the user via a gateway.

The potential hazards are similar to those of presenting content from an external data source through a portlet.

Portal searching

All enterprise portals have a search feature that searches all of the content that exists in the portal application, including internal content, uploaded files, and crawled data sources.

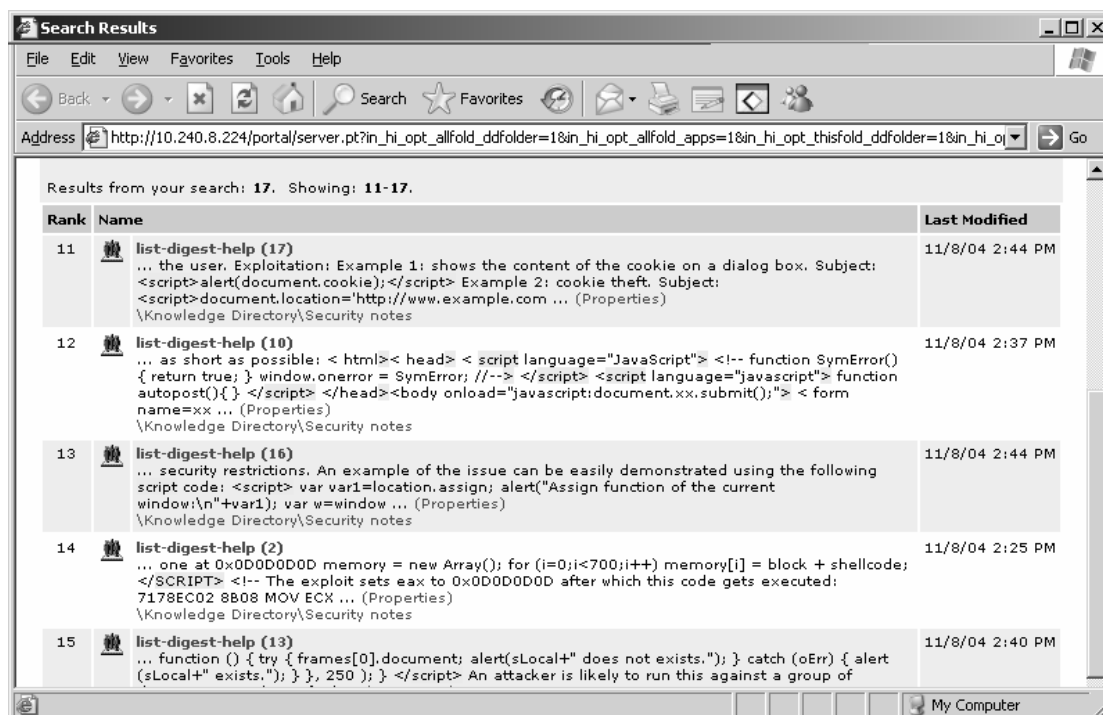


Figure 8: Search results

The search results page is yet another place to check for cross site scripting that is due to malicious content.

Summary

Enterprise portals have a lot of content sources.

Content templates provide functionality similar to the functionality of the ubiquitous open source content management applications, and must be checked in a similar manner.

Files uploaded to a portal are not screened by a content application, and may contain malicious XSS code in surprising places. Moreover, they are served to the portal users in the portal's session, providing a convenient cross site scripting option.

The portal can mirror through its internal gateway external content, giving the owners of the external content an option to insert unwanted XSS code. Portal crawlers can enable the publishing of otherwise inaccessible content, but again – the content owners can abuse it to attack the portal user's session by cross site scripting.