

# Open issues for the next release

## Table of contents

1 [FOR-591] MaxMemory needs increasing for large document sets: Memory Leak with XMLFileModule.....	2
2 [FOR-867] need doc to explain status of skins and dispatcher.....	2
3 [FOR-868] add relevant notes to the "Upgrading" xdoc.....	2
4 [FOR-865] Add missing entries to status.xml to generate the changes list.....	2
5 [FOR-855] verify the license situation prior to each release.....	2
6 [FOR-572] run a memory profiler while forrest is operating.....	3
7 [FOR-533] Auto Generate plugins.xml entry.....	3
8 [FOR-639] define terminology for the various aspects of Dispatcher.....	3
9 [FOR-735] Plugins are not correctly deployed in webapp mode.....	3
10 [FOR-711] Cache results from the Locationmap.....	3
11 [FOR-742] trouble accessing unversioned plugin for a released version of Forrest, e.g. projectInfo.....	4
12 [FOR-241] character entities (e.g. ampersand) are expanded again for href or src attributes.....	4
13 [FOR-210] whole-site html and pdf: broken link faq, broken image links.....	4
14 [FOR-211] whole-site html and pdf: broken ext links.....	4
15 [FOR-200] Locationmap for Forrest and Users.....	5
16 [FOR-546] Sitemap reference doc should be updated to reflect plugin architecture.....	5
17 [FOR-560] Remove duplicate jars from eclipse plugins.....	5
18 [FOR-644] code-style cleanup for xml files.....	5
19 [FOR-666] clarify the sitemap matches etc. in FAQ about non-skinned html.....	6
20 [FOR-726] use locationmap in all of the plugins.....	6
21 [FOR-721] entries without labels in site.xml are now being crawled and generated.....	6
22 [FOR-677] leading slash in gathered URIs causes double the number of links to be processed....	6
23 [FOR-765] forrest war gets NoSuchMethodError for some core transformer.....	6
24 [FOR-772] INvalid HTML as source causes unexpected behaviour.....	7
25 [FOR-776] rationalise the pluginTemplate and current plugins to have minimal configuration files, etc.....	7

**Note:**

These are the top 25 open issues for the next release that are scheduled in our [issue tracking system](#) (see [all](#) open issues). The listing below is regenerated on each Forrest run.

## 1. [FOR-591] MaxMemory needs increasing for large document sets: Memory Leak with XMLFileModule

<http://issues.apache.org/jira/browse/FOR-591>

Since the docs restructuring for the 0.7 release it has become necessary to increase the maxmemory to be able to build the Forrest site. We gained three copies of the documents so suddenly have a large document set to trigger memory leakage issues. <br/> <br/> Possibly <a href="http://issues.apache.org/jira/browse/COCOON-1574" title="Memory Leak with XMLFileModule">COCOON-1574</a> &quot;Memory Leak with XMLFileModule&quot;; <br/> <br/> Does someone have the tools to run some diagnostics? <br/> <br/> (NB maxmemory has been increased in our site-author/forrest.properties, if we resolve this issue it should be reduced again)

## 2. [FOR-867] need doc to explain status of skins and dispatcher

<http://issues.apache.org/jira/browse/FOR-867>

Need doc to explain status of skins and dispatcher. Link to the doc from warnings of dispatchers docs, from skins.html and from upgrading\_08.html <br/> <br/> I have already commenced this xdoc.

## 3. [FOR-868] add relevant notes to the "Upgrading" xdoc

<http://issues.apache.org/jira/browse/FOR-868>

We need to add some notes to the upgrading\_0\*.html doc for the upcoming release. This would most easily be done after attending to <a href="http://issues.apache.org/jira/browse/FOR-865" title="Add missing entries to status.xml to generate the changes list">FOR-865</a> &quot;Add missing entries to status.xml to generate the changes list&quot;;

## 4. [FOR-865] Add missing entries to status.xml to generate the changes list

<http://issues.apache.org/jira/browse/FOR-865>

There are insufficient entries in our site-author/status.xml to generate the changes.html for the upcoming release. <br/> <br/> Besides informing users of the major changes and providing cross-links to the issue tracker, this file also provides recognition of contributions ... at the moment there are too few. It also has the importance=high attribute, which generates our release announcement ... again, too few. <br/> <br/> Re: better use of changes.html (Was: Community health) <br/> <a href="http://marc.theaimsgroup.com/?t=114274836600001">http://marc.theaimsgroup.com/?t=114274836600001

## 5. [FOR-855] verify the license situation prior to each release

<http://issues.apache.org/jira/browse/FOR-855>

This should be continually happening anyway, but immediately prior to each release we need to verify that our license situation is in order. This issue should not ever be closed, rather just move the &quot;Fix for Version&quot; on to the next release. <br/> <br/> Here are some of the tasks: <br/> <br/> A) Ensure that all supporting libraries have a corresponding license. Basically every jar file or other external package needs to have a \*.license.txt file. Ensure that any license conditions are met, e.g. for some we must add an entry to NOTICE.txt, while for some others we must not. Remember to abide by the ASF guidelines (e.g. nothing more restrictive than the Apache License). <br/> <br/> B) Scan the whole trunk repository to add missing ASF license headers to source files and to ensure that the ASF license headers have not been accidentally added to external files. See etc/relicense.txt <br/>

## 6. [FOR-572] run a memory profiler while forrest is operating

<http://issues.apache.org/jira/browse/FOR-572>

We need to run a memory profiler while forrest is operating.

## 7. [FOR-533] Auto Generate plugins.xml entry

<http://issues.apache.org/jira/browse/FOR-533>

The information in the plugins.xml file would be better kept in the plugin directory and added to the plugins.xml file when deployed. This would reduce the amount of duplication in the plugins config files. <br/> <br/> All the necessary values are now in the plugin build.xml file. <br/> <br/> This change will require that the plugins.xml file be retrieved from the build directory rather than the plugins directory when building the plugin documentation pages. We will therefore need a fall back to retrieve this file from the network if it is not currently available - this can be done with the locationmap

## 8. [FOR-639] define terminology for the various aspects of Dispatcher

<http://issues.apache.org/jira/browse/FOR-639>

See the email thread: <br/> <a

href="http://marc.theaimsgroup.com/?t=112276643700001">http://marc.theaimsgroup.com/?t=112276643700001<br/>

## 9. [FOR-735] Plugins are not correctly deployed in webapp mode

<http://issues.apache.org/jira/browse/FOR-735>

(At this point I didn't try this scenario step by step any more; basically it is what we did to setup the site, except that we copied our project data from forrest 0.6 before trying the pdf links) <br/> <br/> - run forrest webapp to create an empty webapp <br/> - configure Tomcat to run the webapp (we did it by creating a context descriptor and put it onder Tomcat's config directory) <br/> - the pdf links give an error &quot;Resource Not Found&quot;;

## 10. [FOR-711] Cache results from the Locationmap

<http://issues.apache.org/jira/browse/FOR-711>

Now that we are using the locationmap extensively it is showing up just how innefficient it is. The problem is that for the majority of requests there are multiple requeests to the locationmap. We can make things much faster (especially on the first page request) by caching results in the locationmap. <br/> <br/> I think a simple cache will suffice, lets just provide a static hashmap using the hint as a key and, of course, the location as the value. <br/> <br/> If we test all locationmaps and find no result we should record that tere is no result in this hashmap. This will also be a good place to throw an exception so that Cocoon can better report such errors (see <a href="http://issues.apache.org/jira/browse/FOR-701" title="Missing locationmap entry gives poor error">FOR-701</a>)

## 11. [FOR-742] trouble accessing unversioned plugin for a released version of Forrest, e.g. projectInfo

<http://issues.apache.org/jira/browse/FOR-742>

The plugin retrieval and deployment process are not quite correct. Recently the projectInfo plugin had its version number incremented and the &quot;unversioned&quot; plugin now relates specifically to 0.8-dev version. That prevents 0.7 from accessing the relevant plugin. <br/> <br/> The solution is discussed here: <br/> <a href="http://marc.theaimsgroup.com/?t=113176328300002">http://marc.theaimsgroup.com/?t=113176328300002

## 12. [FOR-241] character entities (e.g. ampersand) are expanded again for href or src attributes

<http://issues.apache.org/jira/browse/FOR-241>

If I have a url like this inside the XML source: <br/> <br/> <a href="http://sourceforge.net/sflogo.php?group\_id=comics-grabber&type=1">http://sourceforge.net/sflogo.php?g  
<br/> <br/> validate-xdocs fails because it wants the literal &amp; to be typed out as <br/> &amp; instead. So I correct it to become: <br/> <br/> <a href="http://sourceforge.net/sflogo.php?group\_id=comics-grabber&type=1">http://sourceforge.net/sflogo.php?g  
<br/> <br/> However, the generated HTML has the entity resolution UNDONE!: <br/> <br/> <a href="http://sourceforge.net/sflogo.php?group\_id=comics-grabber&type=1">http://sourceforge.net/sflogo.php?g  
<br/> <br/> This HTML fails validation as 4.01 transitional. <br/> <br/> David Crossley adds: <br/> There is a demonstration of the problem in the document <br/> src/documentation/content/xdocs/docs/dreams.xml <br/> The ampersand in the link element text is properly handled but <br/> the ampersand in the link href attribute is not.

## 13. [FOR-210] whole-site html and pdf: broken link faq, broken image links

<http://issues.apache.org/jira/browse/FOR-210>

The &quot;fresh-site&quot; build from 'forrest seed site' reports some failures for faq.html and various missing images. Wonder if sitemap issue?<br/> <br/>

## 14. [FOR-211] whole-site html and pdf: broken ext links

<http://issues.apache.org/jira/browse/FOR-211>

In the the generated site.html all of the external links are broken (i.e. the href attributes are like ... error:#ext:forrest).

## 15. [FOR-200] Locationmap for Forrest and Users

<http://issues.apache.org/jira/browse/FOR-200>

The locationmap gives us the ability to specify where sources are, both for Forrest and for the users. Beware that it will not work for raw files that are not linked, as this "feature" currently uses a fixed dir being copied by Ant.

## 16. [FOR-546] Sitemap reference doc should be updated to reflect plugin architecture

<http://issues.apache.org/jira/browse/FOR-546>

The example at <http://forrest.apache.org/0.7/docs/sitemap-ref.html#pdf> uses a fo2pdf serializer. This probably worked in versions < 0.7, because the PDF output feature was in the core, and the serializer was very likely declared there. With version 0.7, PDF output is moved to a plugin, the fo2pdf serializer seems to be no longer there. The serializer should probably be added to the example, for completeness.

```
<map:components>  
  <map:serializers>  
    <map:serializer name="fo2pdf"/>  
  </map:serializers>  
</map:components>
```

I found out about this because my sitemap uses the fo2pdf too (docbook to PDF), and had to add the serializer.

## 17. [FOR-560] Remove duplicate jars from eclipse plugins

<http://issues.apache.org/jira/browse/FOR-560>

tools/eclipse/plugins/org.apache.forrest.eclipse.servletEngine/lib contains some duplicate jars to those in the main Forrest trunk. We should find a way of reusing the jars from their existing location.

## 18. [FOR-644] code-style cleanup for xml files

<http://issues.apache.org/jira/browse/FOR-644>

We have much inconsistent whitespace in all of our files. This is known to cause trouble in a collaborative environment. See discussion at: <http://marc.theaimsgroup.com/?t=112450901100001> and <http://marc.theaimsgroup.com/?t=112495618800002>

and various linked discussions. <br/>

## 19. [FOR-666] clarify the sitemap matches etc. in FAQ about non-skinned html

<http://issues.apache.org/jira/browse/FOR-666>

The pattern matching example is quite misleading. There should be a note where exactly to copy your old site (in a standard seeded project) to make the pattern matching work. <br/> <br/> At least for me it worked only after I used a \*.html directive. <br/> <br/> <code>map:match pattern=&quot;old\_site/\*.html&quot;&gt;

## 20. [FOR-726] use locationmap in all of the plugins

<http://issues.apache.org/jira/browse/FOR-726>

Each plugin has sitemaps that need to use the locationmap. <br/> <br/> (Perhaps we need Jira sub-tasks for each plugin.)

## 21. [FOR-721] entries without labels in site.xml are now being crawled and generated

<http://issues.apache.org/jira/browse/FOR-721>

In our forrest/site-author/content/xdocs/site.xml there are two entries without &quot;label&quot; attributes. Previously these documents were not being generated. Not sure if this change in behaviour is good or bad. Needs investigation. <br/> <br/> This is most likely a side-effect of the workaround for issue <a href="http://issues.apache.org/jira/browse/FOR-675" title="upgrading to commons-jxpath-1.2.jar causes failures with linkrewriter protocols site: etc.">FOR-675</a>

## 22. [FOR-677] leading slash in gathered URIs causes double the number of links to be processed

<http://issues.apache.org/jira/browse/FOR-677>

Doing 'forrest' starts at the virtual document called linkmap.html where the Cocoon crawler gathers the initial set of links, then starts crawling and generating pages. Any new links are pushed onto the linkmap. However, for some sites, such as our own &quot;seed-sample&quot; and our &quot;site-author&quot;, there is a sudden jump in the number of URIs remaining to be processed. <br/> <br/> This is due to a URI with a leading slash (e.g. /samples/faq.html). When that URI is processed, it gains a whole new set of links all with leading slashes, and so the list of URIs is potentially doubled. <br/> <br/> This issue could be due to a user error, i.e. adding a link that deliberately begins with a slash. Sometimes, that is unavoidable. <br/> <br/> However, we do have a sitemap transformer to &quot;relativize&quot; and &quot;absolutize&quot; the links. Should it always trim the leading slash? Or are there cases where that should not happen, so cannot generalise?

## 23. [FOR-765] forrest war gets NoSuchMethodError for some core transformer

<http://issues.apache.org/jira/browse/FOR-765>

Until recently we could run forrest as a WAR. With today's SVN r356945 it gets past the Cocoon startup phase, then on the first client request it suffers some error which causes a NoSuchMethodError. See attached log ... no other clues. This happens for any site, e.g. 'forrest seed-sample war'. All is okay for 'forrest seed-sample run'.

## 24. [FOR-772] INvalid HTML as source causes unexpected behaviour

<http://issues.apache.org/jira/browse/FOR-772>

Reported via user list:   
 Using HTML as a source:   
 > I manage to cut out one '>' too much. The page still manages to render. There   
 > is no error message, and...I'll be damned...the entire title line is gone,   
 > even when I "view source".   
 I guess this is a problem of using HTML as the source and would not occur if you were using XDoc.   
 We use JTidy to preprocess the HTML (we must have valid XML to work with). I guess JTidy is just removing the invalid content. Can we configure JTidy to fail if such an error exists.   
 Does such an error get spotted during the validation stage of a "forrest site" command (execute directly with "forrest validate")

## 25. [FOR-776] rationalise the pluginTemplate and current plugins to have minimal configuration files, etc.

<http://issues.apache.org/jira/browse/FOR-776>

forrest.properties can have minimal entries and rely on the default.forrest.properties   
 Remove PDF from each "project.required.plugins"   
 skinconf.xml needs to be ready-to-go for Apache Forrest hosted plugins. Helps with keeping the plugin docs consistent. (The need for skinconf.xml will probably go away with the new Dispatcher.)   
 Fix whitespace in pluginTemplate files and existing plugins. Helps to use diff to keep consistent config.   
 xdocs/images should be in documentation/resources/images   
 site.xml needs to be consistent and remove any unnecessary files for xdocs.   
 Tidy up the pluginTempate/status.xml