

## Is Your Content Alive?

The first article in the Apache Lenya series is of rather general nature. Before digging into technical concepts, I'll try to give you some motivation to read on and to provide you with a basic understanding what's so special about the content storage in Lenya.

If you'd like to discuss the contents of the article, I'd like to invite you to the Apache Lenya user mailing list (<http://lenya.apache.org/community/ml.html>).

I'd like to start with a little report about a project we (the BeCompany) finished recently. Our customer runs quite a lot of Apache Lenya publications. Most of them are powered by a custom version of Lenya 1.2.4, but one publication, a news magazine, ran on an older Lenya version, using an outdated layout. The goal of the project was to update the news magazine to the new Lenya version of the custom CMS, thus providing better integration with the rest of the website.

During one of the first meetings, the customer stated that they want to export the existing content (about 1600 scientific articles and announcements) of the news magazine as plain HTML pages to keep them available without having to run the outdated CMS. I was surprised, even shocked – after all, one of the major reasons they had chosen Lenya was that the content can be kept alive, even when migrating to another system, let alone another Lenya version! Fortunately they decided to invest three days into the content migration. It took me about half a day to setup an Ant buildfile and some XSLT stylesheets, providing me with an automatic migration script. Everyone who does CMS projects knows that the content model is bound to change as soon as people start working with the system, therefore it was crucial to have an automated migration script which allows to migrate the content without any manual adaptations everytime the target structure changes. This had to be done three or four times; after updating the XSLT stylesheets the actual migration was a matter of minutes.

By investing only three days, our customer was able to keep their valuable content in a state which allows to edit, process, search, and integrate it into other websites in the same manner as the new content. The old news magazine also contained ancient HTML pages, due to the unstructured nature of this content format it was simply out of the question to migrate it as well.

So, what does it mean if your content is alive? Your content probably isn't alive if you can answer at least one of the following questions with "yes":

- Once my content is published, it's very hard to change it.
- If I want to change content, I have to ask especially skilled people or even an external agency to apply the changes.
- Sometimes I want to include a piece of content in another web page, but I don't think this is possible.
- My content team frequently uses copy & paste to transfer content to other pages. Unfortunately, if the original changes, the copy is not updated.
- I'd like to add some functionality to the website, e.g. an option to view a list of all employees with a particular skill. The employees maintain their personal web pages on the intranet, but I don't know how to access the skill information.
- I'm not pleased with our CMS anymore, but I'm afraid of the migration to another system because I don't want to lose our precious content, or invest thousands of [insert local currency] into the content migration.
- The corporate website layout guidelines have changed, but some layout information is intermingled with the content so I can't simply apply a new style.
- I'd like to output our content in different formats, e.g. as RSS feed, PDF, or a special version containing only the important bits.
- I'd like to access my content from other applications, but I don't know how to accomplish this.
- We have lots of office documents with interesting content which we'd like to make available on the intranet. Unfortunately we don't have the time to categorize them, so nobody will find what she or he is looking for.
- Some years ago we produced some content in plain text files. At the time we thought it wouldn't be important, but now I wish I could easily make it available on the intranet.

If your content is alive, you might also say free, it means that there are no limits to leverage it in any way:

- Publish content in any formats or media.
- Easily change content after it has been published.
- Find and extract information, manually or automatically.
- Process content automatically, preferably using standard tools.
- Combine, filter and transform content items.
- Export or migrate content to other systems, even after many years.

Your content is one of the most valuable assets of your company. Many others use the same content management system, but no other company has the same content! Your employees invested countless hours in creating and maintaining your content. By acquiring and maintaining it in a form that keeps it alive, you will be able to use it multiple times, and in other formats and applications than the original one it was created in. Just remember your childhood times – hasn't the building set or doll's house been much more fun than the already assembled knight's castle or pony farm you could only look at, but not disassemble and combine in new, interesting ways?