

KDE desktop search: Kat, Strigi and Tenor



Published June 17, 2006

Citation

Willighagen, E. (2006, June 17). KDE desktop search: Kat, Strigi and Tenor. *Chem-bla-ics*. <https://doi.org/10.59350/9n9m7-y4v29>

Keywords

Kde, Strigi, Kalzium, Linux, Technorati

Abstract

Desktop searching has become a hot topic (some earlier blogs), now that years of data accumulated on ones hard disk: PDFs, OpenOffice.org documents, Latex manuscripts, old Java source code, digitized music, and a lot of chemical files. Well, on my hard disk that is. Unlike piles of paper, a computer could search this data, but due to the size an index is required. What's KDE4 going to offer?

Copyright

Copyright © None 2006. Distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Desktop searching has become a hot topic (some [earlier blogs](#)), now that years of data accumulated on ones hard disk: PDFs, OpenOffice.org documents, Latex manuscripts, old Java source code, digitized music, and a lot of chemical files. Well, on my hard disk that is. Unlike piles of paper, a computer could search this data, but due to the size an index is required. What's KDE4 going to offer?

For the KDE desktop [Kat](#) has for more than a year offered this, and latter [Kerry](#) came along as frontend to [Beagle(http://beaglewiki.org/Main_Page)], though this does not have the nice integration with KDE [kfile plugins](#). Since then, Kat developed has come to a stop (unfortunately), and attempts to reach the main author ([Roberto](#)) have been unsuccesfull. Last thing happening was a rewrite of the database backend.

Additionally, [Scott Wheeler proposed Tenor](#) on FOSDEM 2005: "*KDE 4: Beyond Hierarchical Data, The Desktop as a Searchable Web of Context*". A semantic desktop; potentially cool, but I have heard [little from it lately](#), except for some rumours that [Scott has some actual code at home](#).

Now, [Strigi](#) ([download](#)) has come along, with a fast indexing engine, just the thing where the Kat developed seemed to have stopped. The design is different from that of Kat, but it does not seem unlikely that Kat code can be ported. No support for PDF or OpenOffice.org documents yet, but that's really the easy part, and kfile is on its way.

Getting back to Tenor, one might wonder how Strigi could implement Tenor concepts. A simple approach is at least to allow users to tag files, just like we have become used to with blogs (e.g. [Technorati.com](#)) and websites (e.g. [Connotea](#)). This could be easily implemented using [extended attributes](#) (xattr), [already used by Beagle](#) :

```
# file: home/egonw/1CRN.jpg
user.Tenor.Keywords="crambin"
user.Tenor.Comment="Used in my ontologies presentation."
```

Obviously, this example shows not just these tags, but a user comment too. The idea, here, is that Strigi mines these attributes in addition to the file itself, so that search on tags can be done too. BTW, my argument to use this, instead of putting these things in the Strigi database itself, is persistence: data and metadata are kept together. KDE's file properties dialog would be extended with an extra tab that allows editing these fields.

Strigi itself can be embedded in KDE applications to search specific information (e.g. search molecular data within [Kalzium](#) using the [InChI](#)), and even in the FileOpen dialog. We need patches for KDE4 that allows this, soon.