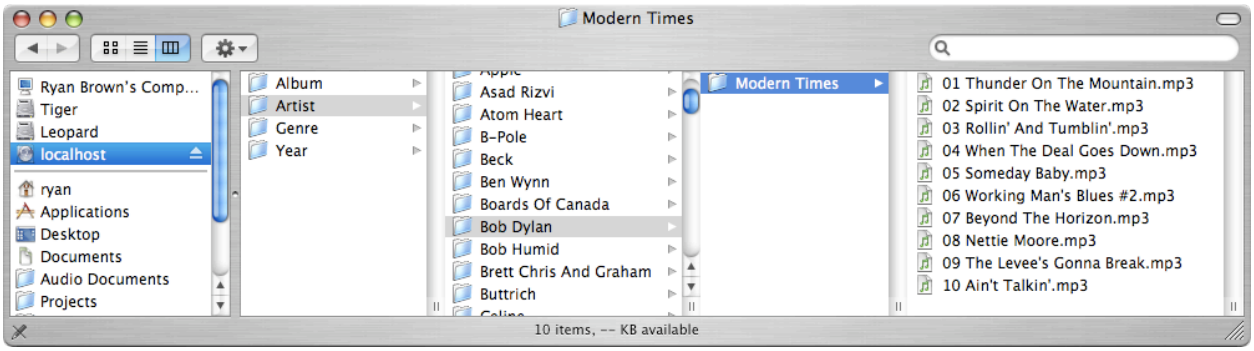
Metaphor Abstract

Ryan Brown

File metadata is a hot topic in computer science right now. Apple recently introduced a metadata search engine in Mac OS X 10.4 and Microsoft has been working on metadata tagging and search for Vista. An open source project named Beagle brings such features to Linux. These metadata systems all allow a user to tag files with arbitrary metadata. Pictures, for example, can be given a tag indicating where they were taken, and this information can later be used in searches. These systems also automatically extract metadata where ever possible (using ID3 and EXIF tags for instance). Metaphor addresses two problems common to these systems. First, it provides an interface for browsing metadata. Second, it exports metadata in way that is accessible through traditional file system views and utilities.

To illustrate why browsing is desirable over searching in some circumstances, assume a user wants to find all the rock music on his system. This could be accomplished by doing a search, but the search results would be presented as a flat list. Even with several levels of groupings, one artist could take up several pages of results, preventing a quick glance at what rock artists are available (or what rock albums are available, etc). An interface to browse metadata, on the other hand, would present the available rock music in a hierarchy. The first level could contain all artists, the next all albums for that artist, etc.

The interface Metaphor uses for browsing such hierarchies is the traditional file system. Metaphor creates a WebDAV server (a ubiquitous network filesystem based on HTTP) which contains a virtual file system of metadata hierarchies. Many common hierarchies are predefined, like the hierarchy found on an iPod for browsing music and something along those lines for movies. By mounting the WebDAV server locally, users can access the metadata hierarchy just like the regular directory hierarchy.



Users that aren't satisfied with the defaults can add their own hierarchies using a simple UI that interacts with Metaphor using a CGI interface on Metaphor's HTTP server. Rather than write Metaphor's server from scratch, a lightweight, high performance HTTP server named lighttpd was extended via it's plugin API. lighttpd is a mature project that has facilitated the development for both Linux and OS X.

In this work I also address the issue of performance. In order to avoid making costly queries every time the user requests a new view, Metaphor maintains its own optimized copy of the metadata, which is updated based on kernel file-change notifications.

Operating systems seem to be moving away from the traditional file system and towards a system consisting only of metadata. In a pure metadata system rather than saving files to folders to maintain organization, users would only have to assign tags. As this transition takes place, a means to allow legacy systems to access pure metadata systems is necessary, and this is what Metaphor provides.