1.2.2 What Lucene can do for you

Lucene allows you to add search capabilities to yourapplication. Lucene can index and make searchable

any data that you can extract text from. As you can see in figure 1.4, Lucene doesn't care about the source

of the data, its format, or even its language, as longas you can derive text from it. This means you can

index and search data stored in files: web pages on remote web servers, documents stored in local file

systems, simple text files, Microsoft Word documents, XML or HTMLor PDFfiles, or any other format

from which you can extract textual information.

Similarly, with Lucene's help you can index data stored in your databases, giving your users full-text

search capabilities that many databases don't provide. Once you integrate Lucene, users of your

applications can perform searches by entering queries like +George +Rice -eat -pudding, Apple -pie +Tiger, animal:monkey AND food:banana, and so on. With Lucene, you can index and search

email messages, mailing-list archives, instant messenger chats, your Wiki pages … the list goes on. Let's

recap Lucene's history now.

1.2.3 History of Lucene

Lucene was originally written by Doug Cutting;

2

it was initially available for download from its home at the

SourceForge web site. It joined the Apache SoftwareFoundation's Jakarta family of high-quality open

source Java products in September 2001 and became its own top-level Apache project in February 2005.

It now has a number of sub-projects, which you can see at http://lucene.apache.org. With each release,

the project has enjoyed increased visibility, attracting more users and developers. As of March 2009, Lucene

version 2.4.1 has been released. Table 1.1 shows Lucene's release history.