Simplify

Version

0.1.1

SIMPLIFY USER MANUAL & CONFIGURATION GUIDE

SEEMANT RAJVANSHI

Simplify User Manual

VERSION 0.1.0

Simplify User Guide

© Seemant Rajvanshi
Seemant.rajvanshi@gmail.com

Table of Contents

Simplify

Simplify is a free to use advance note taking application with features unparallel to any other note taking software. It is a complete personal note management solution with features like information tagging, categorization, advance screen reader, advance searching (boost, proximity, fuzzy, wildcard etc), ability to find similar content, automatic report generation, external files and web resource management. At its heart it is a wiki coupled with a full text indexer which can index any file (with IFilter support). Currently indexes older and newer MS Office format (x based format), pdf, txt and html files.

Features

- Tag data at information level rather than file level. This allows an individual to generate a report out of the interesting information.
    - For example a student using the software to take notes can tag the important piece of information as 'Key points' and later he/she can generate a report out of these, allowing him to quickly revise import points without going to each individual file.
    - Manage all your web references, To-dos, addresses and any kind of unstructured information using tags and get a list whenever you wish.
    - Quickbit (tagged information which is not associated with article) is a quick and easy alternative to desktop notes with full text search support.
- Enjoy reading your notes with the help of advance screen reader which provides dynamic flow of content, switching from single to multi-page view and seamless zooming with pixel perfect graphics.
- Bring any article, quickbit or external file to the dashboard for even quicker retrieval.
- Wiki like categorization with support for hierarchical articles allowing one to take notes in book form.
- Attach associated PowerPoint, Word etc. files directly to your article and full text searching is automatically enabled over them allowing you to organize related content with ease.
- Find content similar to the current content. This enables in finding information which can't be found using simple searches.
- Simple and easy to use GUI with no clutter.
- Perform advance searches like proximity, fuzzy, boost etc.

Upcoming Features

These features mark the overall direction of Simplify development. Many of these will be a part of scheduled 0.2 release.

- A special kind of article called code which will enable one to manage source codes (over 14 programming languages supported out of the box).
- External and internal resource support. Internal resources will be copied to the Simplify database and will be managed internally. External resources will be loaded on demand. Possibility to load images from Flickr, Picasa etc.
- Live Media support (Streamed: Youtube & similar, Non-stream: audio and video files from hard disk) support allowing one to preview the information right from Simplify itself with opening the browser.
- Report generation functionality out of tagged data.
- Newer tag types like Geo tags, Picture tags, Timeline tags etc.
- Advance functionality like dynamically generated Timelines, charts based upon the tagged information.
- Research desk tool allowing one to gather and organize information faster in an intuitive manner.
- Portable version.
- Better printing support.
- System tray integration.
- More UI enhancements etc

Roadmap

2010-11-05: Simplify 0.2.0 release with a lot of enhancements
2010-12-15: Simplify 0.3.0 beta release
2011-03-15: Simplify RTM build
2010-04-15: Simplify 1.0.0 release

Supported Platforms

- Microsoft Windows XP, Vista, Windows 7

Wiki

https://bitbucket.org/Seemant/simplify/wiki/Home

Issue Tracker

https://bitbucket.org/Seemant/simplify/issues?status=new&status=open

RSS/Atom Feeds

https://bitbucket.org/Seemant/simplify/overview

Frequently asked Questions

I cannot index pdf and Office 2007 documents?

In order to index files Simplify relies on IFilter.

An **IFilter** is a plugin that allows the Windows Indexing Service and the newer Windows Desktop Search to index different file formats so that they become searchable. Without an appropriate IFilter, contents of a file cannot be parsed and indexed by the search engine. IFilters are also used by SQL Server as a component of the SQL Server Full Text Search service.

IFilter is designed to help users to index a large number of documents and then quickly find text within these documents. The documents can be files, email attachments or database records.

Windows Indexing Service, MSN Desktop Search, Internet Information Server, SharePoint Portal Server, Windows SharePoint Services (WSS), Site Server, Exchange Server, SQL Server and all other products based on Microsoft Search technology support indexing technology. Hereafter, we use term "search engine" to refer to these products.

Taken from: http://en.wikipedia.org/wiki/IFilter

You can get appropriate IFilters to your file types from the following sites. Please ensure that you use proper version (x86/x64) of IFilter.

Adobe PDF iFilter 9 for 64-bit platforms

http://www.adobe.com/support/downloads/detail.jsp?ftpID=4025

Adobe PDF iFilter 9 for 32-bit platforms

Already a part of Adobe Reader 9.

2007 Office System Converter: Microsoft Filter Pack (32 & 64 bit)
http://www.microsoft.com/downloads/details.aspx?FamilyId=60C92A37-719C-4077-B5C6-CAC34F4227CC&displaylang=en

Microsoft Office 2010 Filter Packs

http://www.microsoft.com/downloads/details.aspx?FamilyID=5cd4dcd7-d3e6-4970-875e-aba93459fbee&displaylang=en

You can get a lot of IFilters from the following sites:

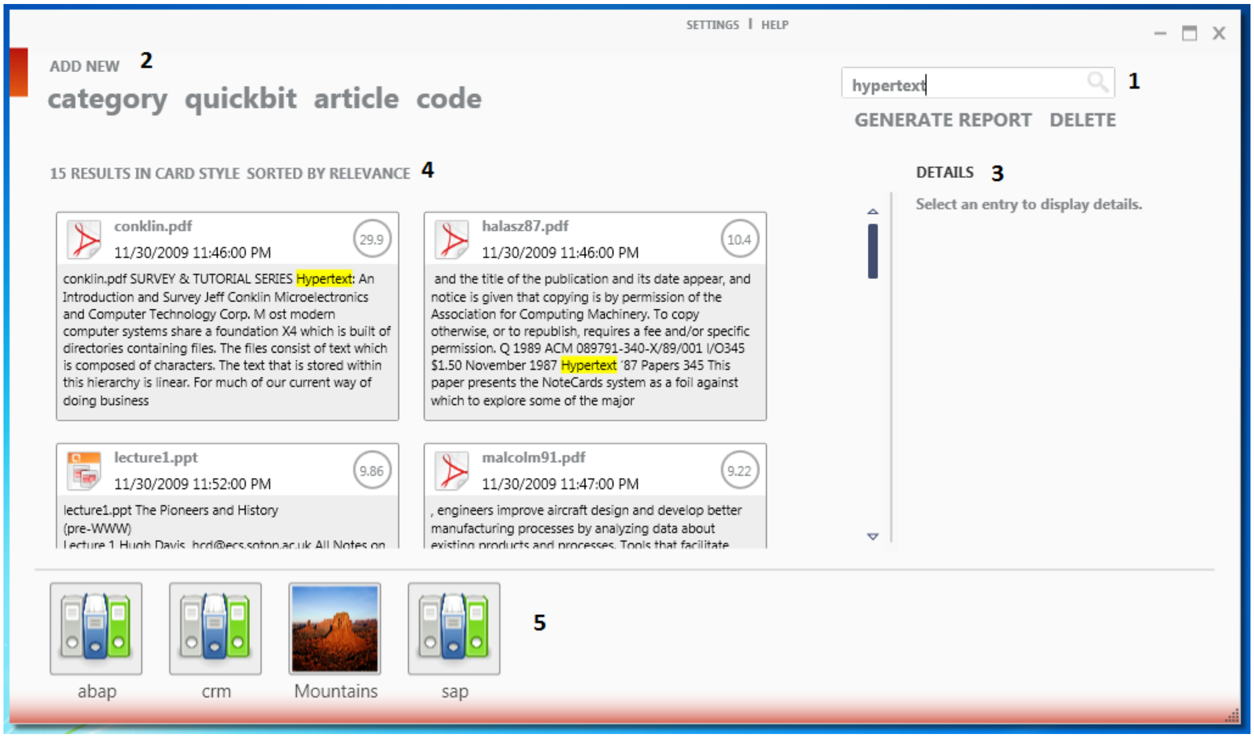
http://www.ifiltershop.com/

http://blogs.msdn.com/b/ifilter/default.aspx?wa=wsignin1.0/default.aspx

Getting Started Guide

Simplify User Interface

Simplify has a very easy going user interface which isn’t complex to understand. The tooltips provided in the system are very informative and will get you out of trouble most of the time.



conklin.pdf 29.9
11/30/2009 11:46:00 PM
conklin.pdf SURVEY & TUTORIAL SERIES Hypertext: An Introduction and Survey Jeff Conklin Microelectronics and Computer Technology Corp. M ost modern computer systems share a foundation X4 which is built of directories containing files. The files consist of text which is composed of characters. The text that is stored within this hierarchy is linear. For much of our current way of doing business

halasz87.pdf 10.4
11/30/2009 11:46:00 PM
and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission. Q 1989 ACM 089791-340-X/89/001 l/O345 $1.50 November 1987 Hypertext ‘87 Papers 345 This paper presents the NoteCards system as a foil against which to explore some of the major

lecture1.ppt 9.86
11/30/2009 11:52:00 PM
lecture1.ppt The Pioneers and History
(pre-WWW)
Lecture 1 Hugh Davis hcd@ecs.soton.ac.uk All Notes on

malcolm91.pdf 9.22
11/30/2009 11:47:00 PM
, engineers improve aircraft design and develop better manufacturing processes by analyzing data about existing products and processes. Tools that facilitate

Figure 1 Main View Simplify

1. Search Box

Search box can be used to search for any kind of data in the system. Many search styles are available for the user through the same interface:

- Wildcard Searches
- Fuzzy Searches
- Proximity Searches
- Range Searches
- Boosting a Term
- Boolean Operators

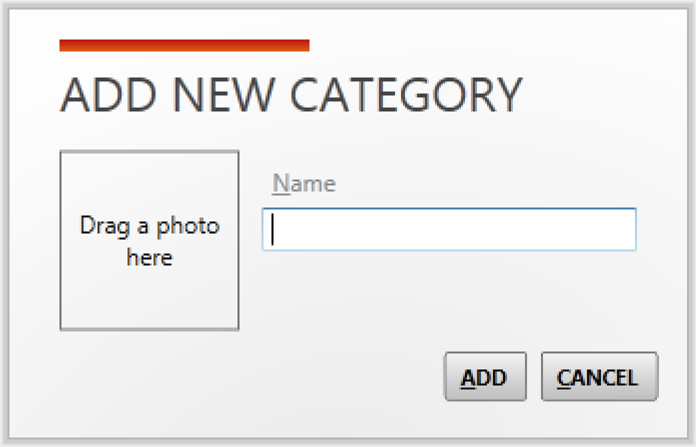
    -
    - [AND]
    - [+]
    - [NOT]
    - [-]
- [Grouping]
- [Field Grouping]

2. Add new pane

This pane can be used to add new content types to the Simplify.

Category

A user can add a new category to the system using the add category option. Category can be used to organize you content. It acts as a high level container to your content. Categories cannot be nested.


ADD NEW CATEGORY

Name

Figure 2 Add new Category Dialog Box

Clicking on add category opens a new dialog box with Name and image option. In order to associate an image with the category simply drag any image from your system over the Drag a photo here box. Duplicate category names are not allowed by the system.

QuickBit

Quickbit item can be used to store small piece of information which isn't really large enough to occupy an article. Quickbits are basically Desktop notes replacement but these are more powerful as one can generate a report out of them based on a certain criteria. You can attach Quickbit to dashboard for faster access.

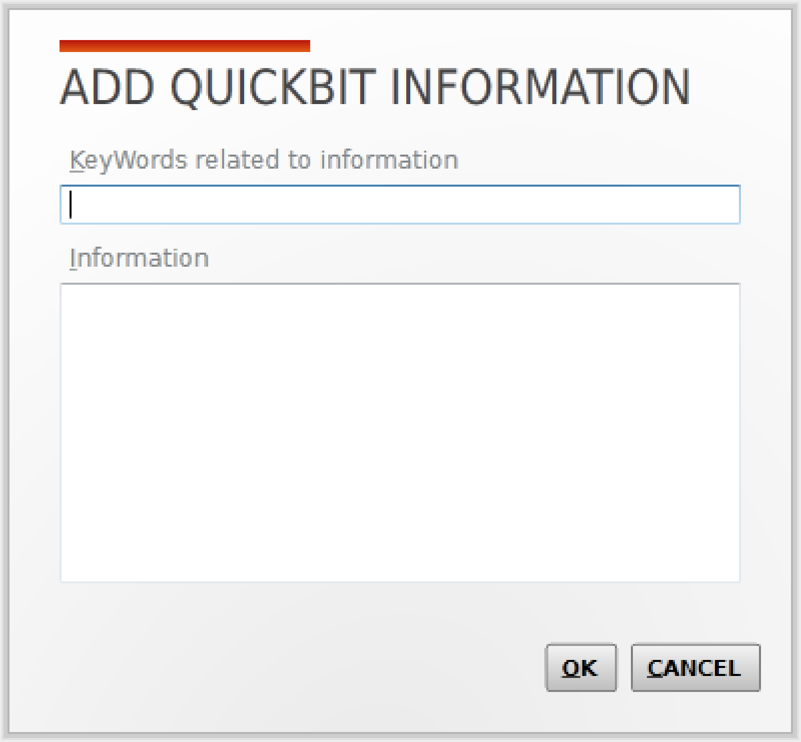


Figure 3 Add Quickbit Dialogbox

Article

User notes are organized in form of articles. This represents the most powerful information type available in the system. These can contain quickbits and multiple types of article.

We will deal with Article interface in Editor Section.

Organize your articles in hierarchrical form to create organized book or binder kind of structure. An article can have a single parent.

Code

Code is a special type of article which can be used to organize programming codes. It provides a syntax highlighted editor which makes it easy to edit codes. It supports more than 14 types of programming language.

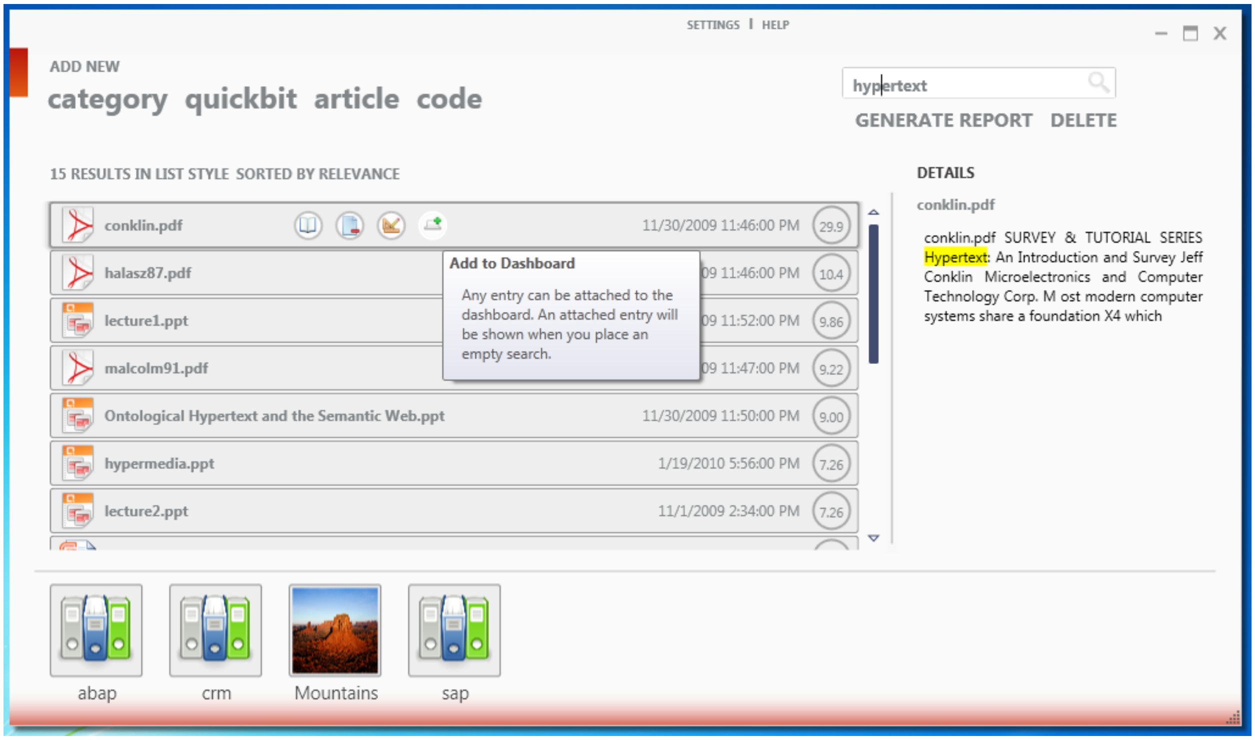
This feature isn't supported in version 0.1.0 but will be made a part of release 0.2.0

3. Details Pane

This pane is used to provide detailed information about the selected item.

4. Display View switcher and sorting mode changer

This pane is associated with search result area. Clicking on **# RESULTS IN #** toggles the view mode from Card style to list style. By default card style can show up to 50 results in the result area. This is done to conserve the system memory. In case you want to view more results then you should switch to list style.



conklin.pdf 11/30/2009 11:46:00 PM 29.9
halasz87.pdf 09 11:46:00 PM 10.4
lecture1.ppt 09 11:52:00 PM 9.86
malcolm91.pdf 09 11:47:00 PM 9.22
Ontological Hypertext and the Semantic Web.ppt 11/30/2009 11:50:00 PM 9.00
hypermedia.ppt 1/19/2010 5:56:00 PM 7.26
lecture2.ppt 11/1/2009 2:34:00 PM 7.26

Add to Dashboard

Any entry can be attached to the dashboard. An attached entry will be shown when you place an empty search.

conklin.pdf

conklin.pdf SURVEY & TUTORIAL SERIES Hypertext: An Introduction and Survey Jeff Conklin Microelectronics and Computer Technology Corp. M ost modern computer systems share a foundation X4 which

Figure 4 List style view

Clicking on **SORTED BY RELEVANCE** label sorts the displayed results according to:

- Sort by relevance (based on the search score)
- Sort by date modified
- Sort by title

The items are dynamically sorted on the screen in a visually appealing manner with each item flying to their new position on the screen.

5. Category display pane

This pane displays all the available categories in the system. User can remove/modify existing categories from this pane.

Modifying/Deleting categories

Hovering mouse over the category item brings a new menu which can be used to edit/delete a category.

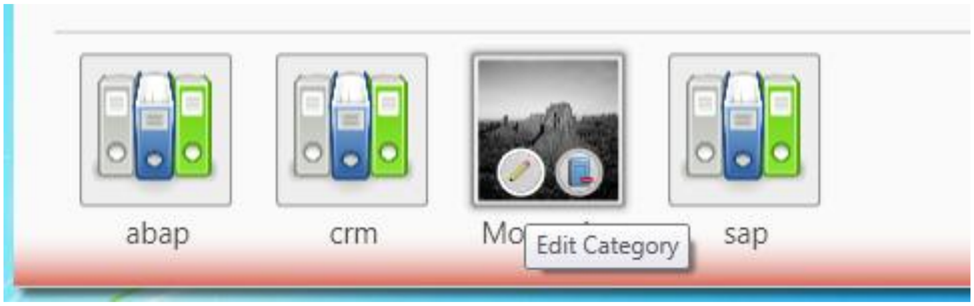
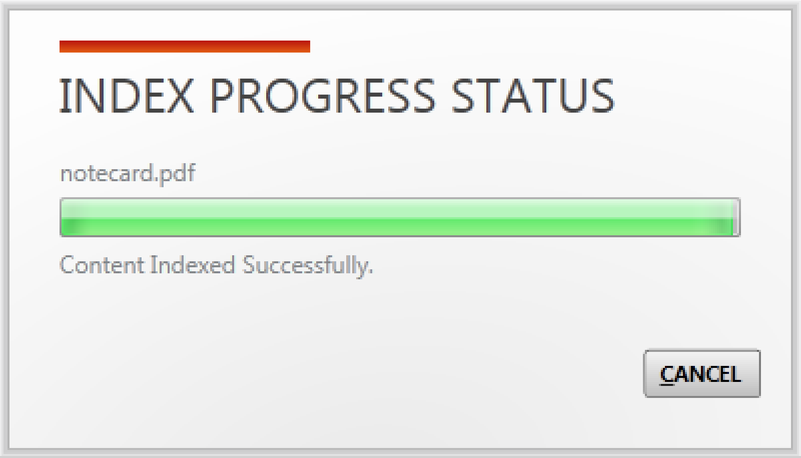


Figure 5 Edit/Delete Category

Adding Files to Simplify

Files can be added to Simplify by dragging and dropping the files over the Simplify Main window. User can drop both files and folder. Added folders are automatically added to the "**Folders to Index**" list which can be accessed from setting screen.

When a user drops a file or folder over the main search window a popup window is opened as shown below. This window shows the indexing progress and automatically closes down at the finish of indexing process.

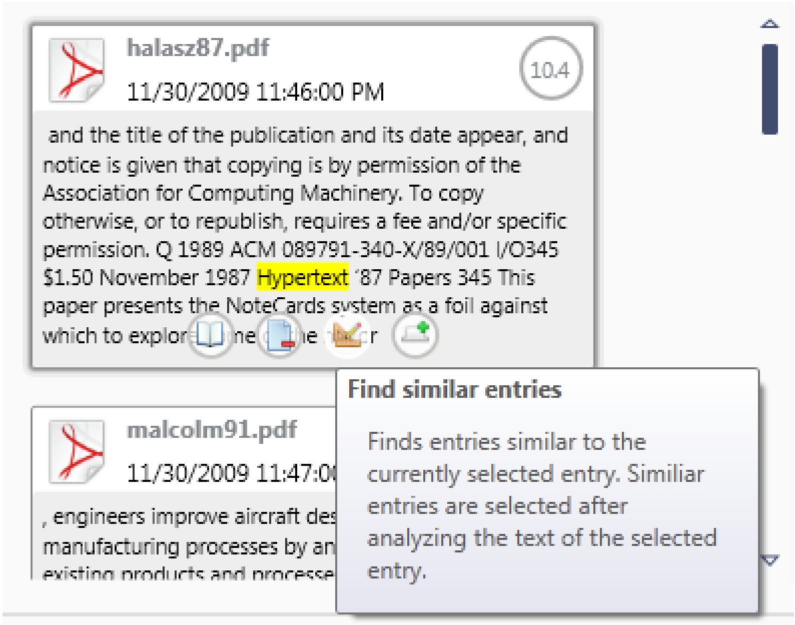

INDEX PROGRESS STATUS

notecard.pdf

Content Indexed Successfully.

Figure 6 Indexing progress status

Understanding Search Results



halasz87.pdf

11/30/2009 11:46:00 PM

10.4

and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission. Q 1989 ACM 089791-340-X/89/001 l/O345 $1.50 November 1987 Hypertext '87 Papers 345 This paper presents the NoteCards system as a foil against which to explor ne e r

Find similar entries

Finds entries similar to the currently selected entry. Similiar entries are selected after analyzing the text of the selected entry.

malcolm91.pdf

11/30/2009 11:47:0

, engineers improve aircraft des manufacturing processes by an existing products and processe

Figure 7 Search Results

Hovering mouse over the displayed results brings a new action menu. The items available in the action menu will depend upon the type of content. For example an external resource (in the example above) will only have an view item menu entry and no edit item menu entry as the resource cannot be edited inside the system.

Some of the most important options are:

- View item
- Edit item
- Delete item (external resources are only removed from Simplify and not deleted from the system)
- Find similar entries: This option can be used to find entries similar to the currently selected item. Simplify internally creates a number of searches based on the content of the current item and presents top 15 results which are similar to the current item.
- Attach to the dashboard: Any result can be attached to the dashboard. Dashboard is displayed when user places a blank search in the search box.

ⓘ	*Dashboard can be considered as a shortcut toolbar to your most used resources..*

- Remove from dashboard menu item is available in Dashboard mode only.

Most of the labels are self explanatory apart from the number in circle on the right. This number represents the relevance of the current item as per user search. This can also be considered as a search score. Higher is better.

Generating Reports

Based on user’s Quickbit or tag search results a report can be auto-generated. This report will have user’s tags and values in a report format organized according to the category. Please refer: Editor View for more information about tagging.

Editor View

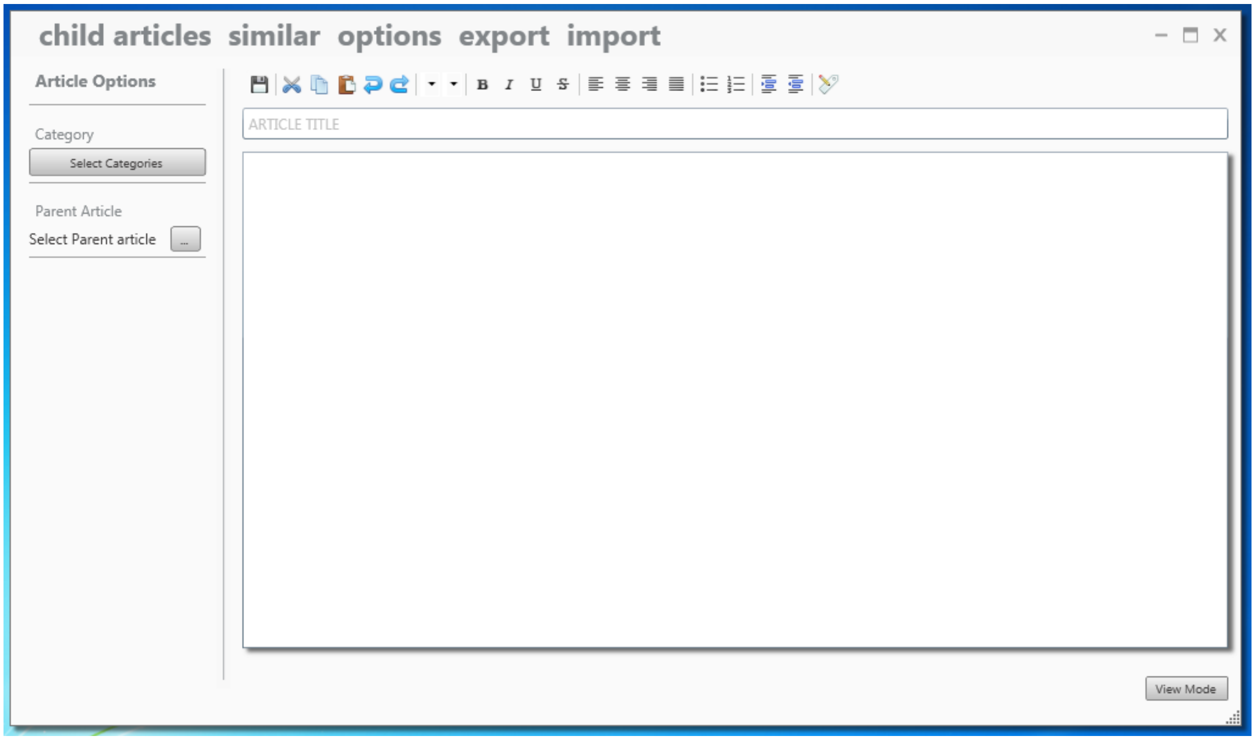


Figure 8 Editor Window

Editor Window is used to edit/enter Article content. User can create article to category associations directly from this window.

An article can be associated to multiple categories and a single parent article.

Child Articles: Gives the child articles of the current article.

Similar Articles: Same as the option in dashboard.

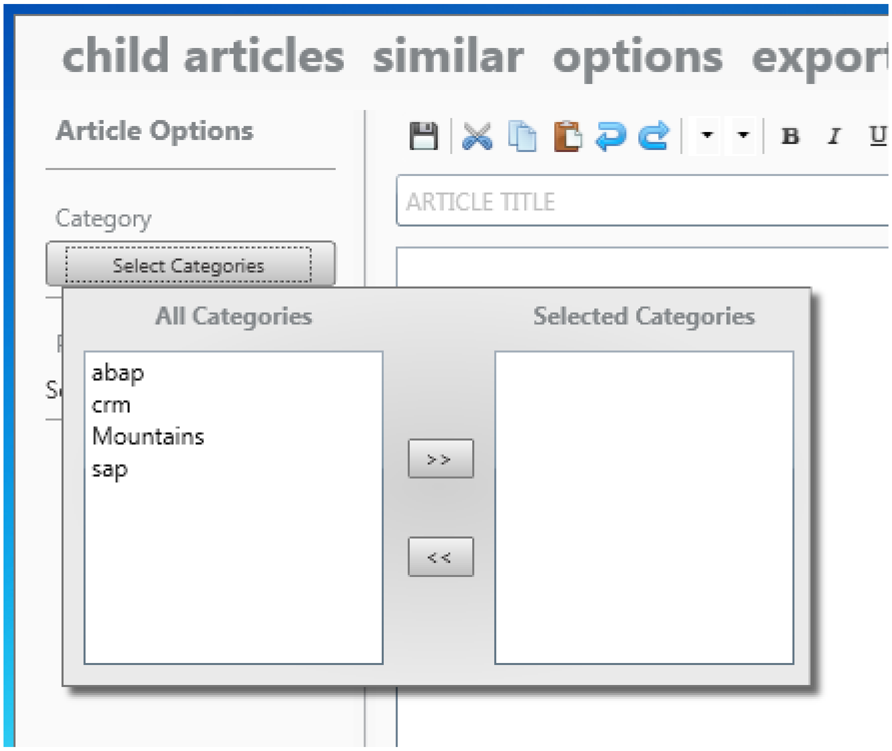


Figure 9 Associate article to categories

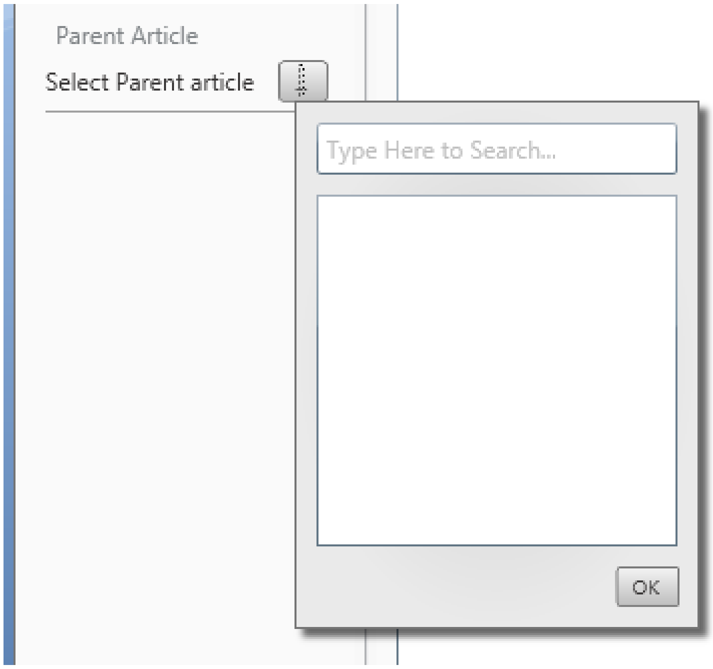


Figure 10 Set Parent Article

Tagging at information level

People are use to tagging at file level in Wiki's and other note management software. Tagging directly at information level gives you distinct advantage.

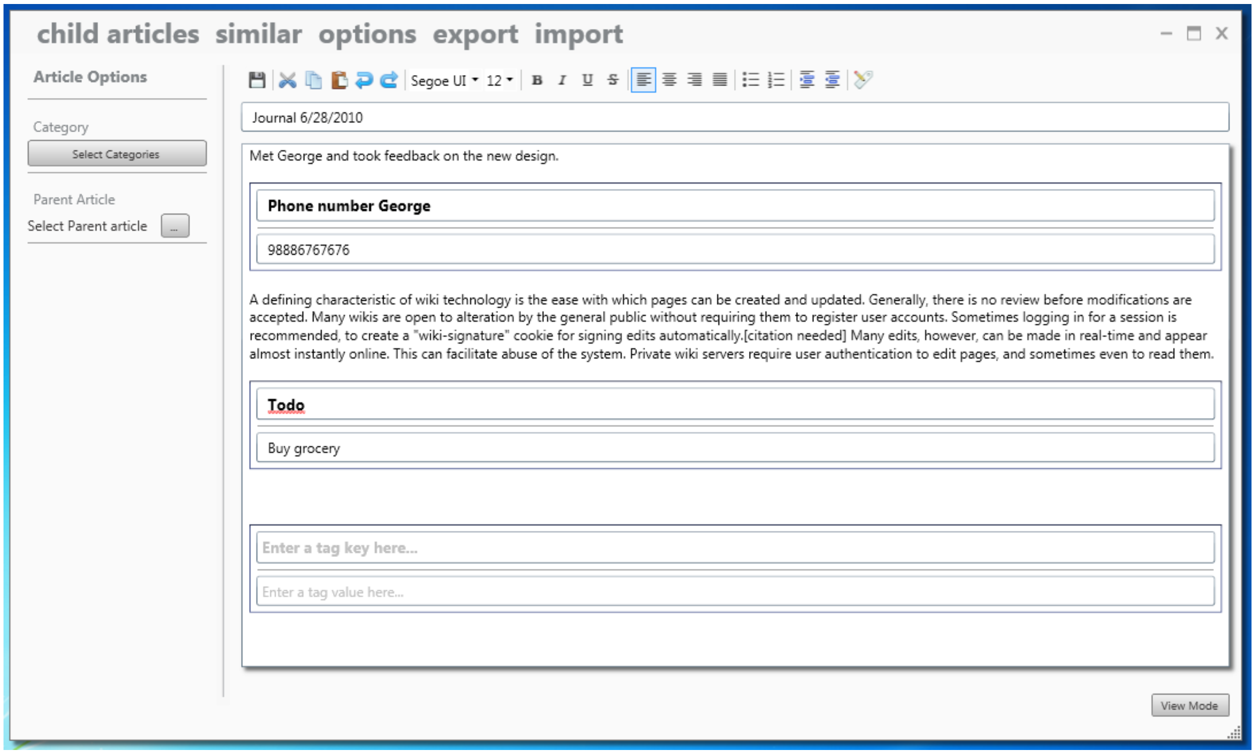
- The system can know what exactly the tagged content means by looking at the key of the tag.
- Since information is tagged, it can be retrieved faster than regular searching. Searching across many files may give you results that a particular term is present in a file x number of times but it may not exactly retrieve the information for you from all the files.
- You can auto-generate reports based on your tagged data.

Tagging allows you to keep all kind of data in a single file without caring about their organization. For example you can create a new article for daily activities and mark down all your days work right in their.

Suppose you have a todo item to complete put it in the same file and tag it as todo.

Suppose you got a phone number, instead of finding a place to save it and ultimately losing it forever. Put it right there in your daily journal article and tag it as pnone no <person name>.

You can retrieve all this using a simple search and if you want you can generate a report out of the tags.



Journal 6/28/2010

Met George and took feedback on the new design.

Phone number George

98886767676

A defining characteristic of wiki technology is the ease with which pages can be created and updated. Generally, there is no review before modifications are accepted. Many wikis are open to alteration by the general public without requiring them to register user accounts. Sometimes logging in for a session is recommended, to create a "wiki-signature" cookie for signing edits automatically.[citation needed] Many edits, however, can be made in real-time and appear almost instantly online. This can facilitate abuse of the system. Private wiki servers require user authentication to edit pages, and sometimes even to read them.

Todo

Buy grocery

Enter a tag key here...

Enter a tag value here...

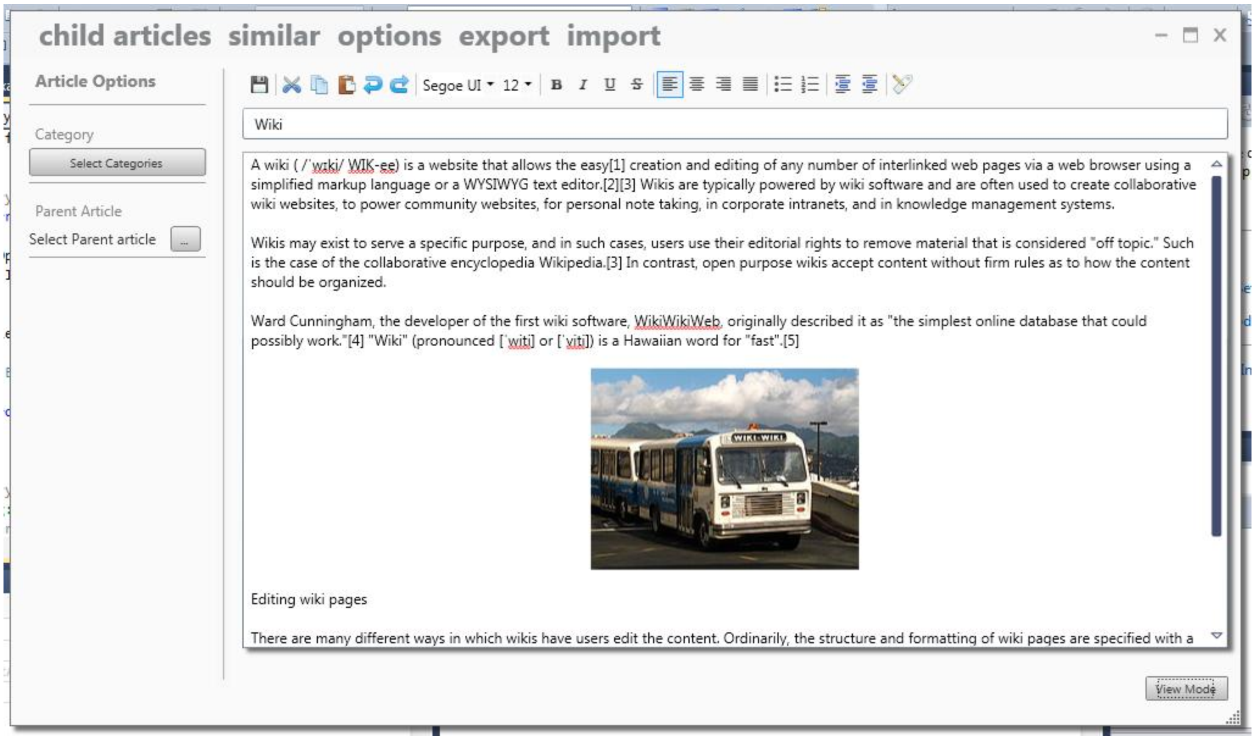
Figure 11 Adding tags alongwith normal content

View Mode

Editor Window has a special screen reader which can be invoked by clicking on to the view mode button at the right hand corner of the screen.

This creates a special screen viewer window which allows the user to read the article in dynamic content layout mode. Here the user can increase or decrease font size, switch to single and multipage views without affecting the original article layout.

Below figure shows an article in editing mode. On the next page the same article is shown in the view mode.



Wiki

A wiki (/ˈwɪki/ WIK-ee) is a website that allows the easy[1] creation and editing of any number of interlinked web pages via a web browser using a simplified markup language or a WYSIWYG text editor.[2][3] Wikis are typically powered by wiki software and are often used to create collaborative wiki websites, to power community websites, for personal note taking, in corporate intranets, and in knowledge management systems.

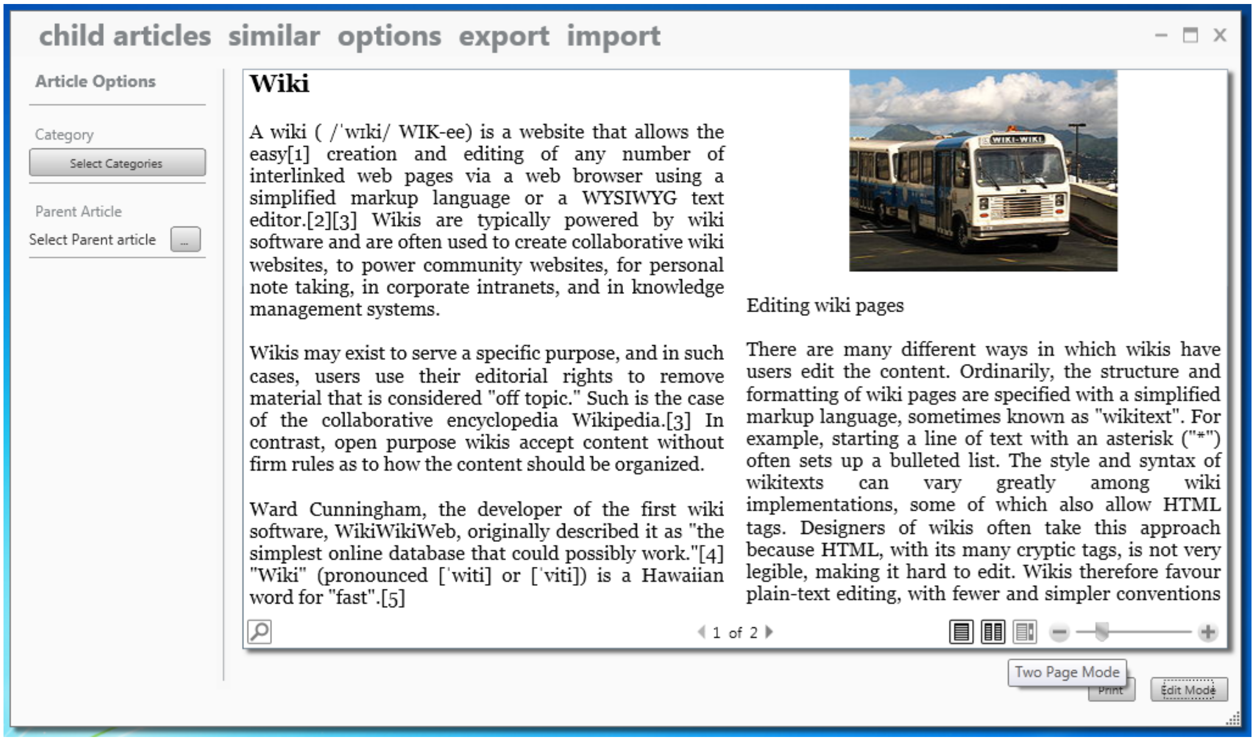
Wikis may exist to serve a specific purpose, and in such cases, users use their editorial rights to remove material that is considered "off topic." Such is the case of the collaborative encyclopedia Wikipedia.[3] In contrast, open purpose wikis accept content without firm rules as to how the content should be organized.

Ward Cunningham, the developer of the first wiki software, WikiWikiWeb, originally described it as "the simplest online database that could possibly work."[4] "Wiki" (pronounced [ˈwiti] or [ˈviti]) is a Hawaiian word for "fast".[5]

Editing wiki pages

There are many different ways in which wikis have users edit the content. Ordinarily, the structure and formatting of wiki pages are specified with a

Figure 12 Wiki article in editing mode



Wiki

A wiki (/ˈwɪki/ WIK-ee) is a website that allows the easy[1] creation and editing of any number of interlinked web pages via a web browser using a simplified markup language or a WYSIWYG text editor.[2][3] Wikis are typically powered by wiki software and are often used to create collaborative wiki websites, to power community websites, for personal note taking, in corporate intranets, and in knowledge management systems.

Wikis may exist to serve a specific purpose, and in such cases, users use their editorial rights to remove material that is considered "off topic." Such is the case of the collaborative encyclopedia Wikipedia.[3] In contrast, open purpose wikis accept content without firm rules as to how the content should be organized.

Ward Cunningham, the developer of the first wiki software, WikiWikiWeb, originally described it as "the simplest online database that could possibly work."[4] "Wiki" (pronounced [ˈwiti] or [ˈviti]) is a Hawaiian word for "fast".[5]

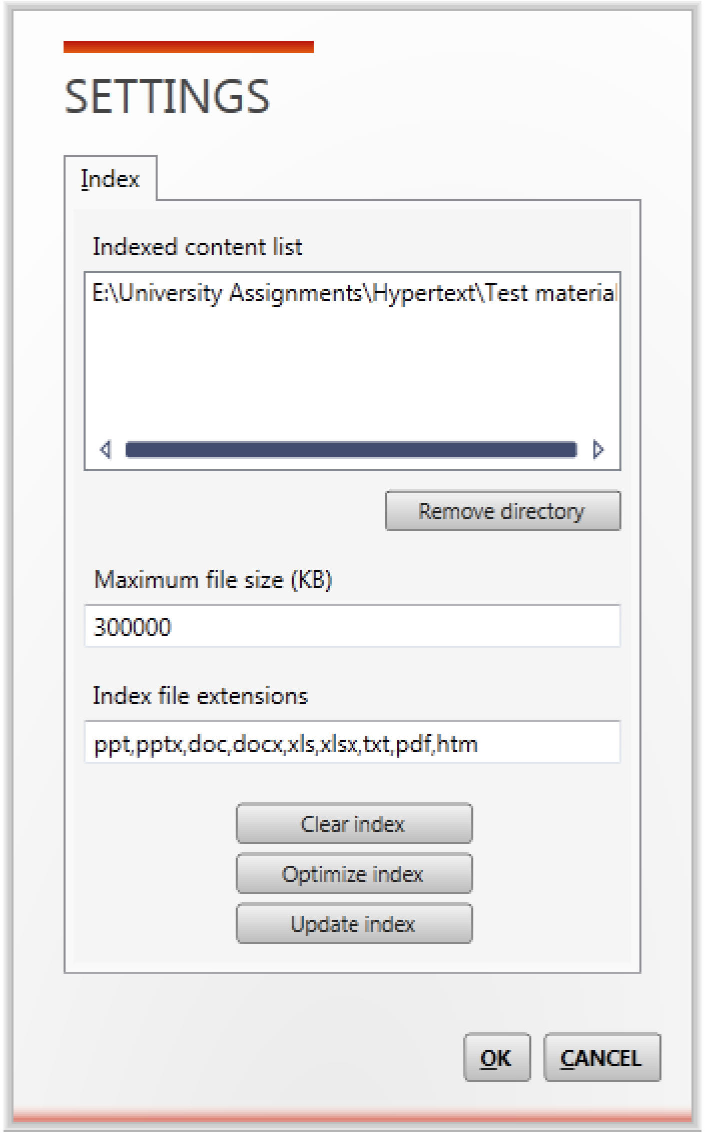
Editing wiki pages

There are many different ways in which wikis have users edit the content. Ordinarily, the structure and formatting of wiki pages are specified with a simplified markup language, sometimes known as "wikitext". For example, starting a line of text with an asterisk ("*") often sets up a bulleted list. The style and syntax of wikitexts can vary greatly among wiki implementations, some of which also allow HTML tags. Designers of wikis often take this approach because HTML, with its many cryptic tags, is not very legible, making it hard to edit. Wikis therefore favour plain-text editing, with fewer and simpler conventions

Figure 13 Wiki article in view mode.

Observe that the content can be dynamically laid out with multiple viewing options.

Setting Panel



SETTINGS

Indexed content list

E:\University Assignments\Hypertext\Test materia

Maximum file size (KB)

300000

Index file extensions

ppt,pptx,doc,docx,xls,xlsx,txt,pdf,htm

Setting panel can be used to configure various advance settings.

- Indexed content list: Gives the list of folders that are indexed.
- Maximum file size: The maximum amount of content that will be indexed per file
- Index files extension: The extension of the files to index.
- Clear index: Removes all the external documents from the system. Articles and tags are not removed.

- Optimize index: Optimizes index for better search performance. Can take up to several minutes to perform.

- Update index: Updates the internal file index.

Advance Search Types

Taken from: http://lucene.apache.org/java/2_4_0/queryparsersyntax.html

Wildcard Searches

Lucene supports single and multiple character wildcard searches within single terms (not within phrase queries).

To perform a single character wildcard search use the "?" symbol.

To perform a multiple character wildcard search use the "*" symbol.

The single character wildcard search looks for terms that match that with the single character replaced. For example, to search for "text" or "test" you can use the search:

```
te?t
```

Multiple character wildcard searches looks for 0 or more characters. For example, to search for test, tests or tester, you can use the search:

```
test*
```

You can also use the wildcard searches in the middle of a term.

```
te*t
```

Note: You cannot use a * or ? symbol as the first character of a search.

Fuzzy Searches

Lucene supports fuzzy searches based on the Levenshtein Distance, or Edit Distance algorithm. To do a fuzzy search use the tilde, "~", symbol at the end of a Single word Term. For example to search for a term similar in spelling to "roam" use the fuzzy search:

```
roam~
```

This search will find terms like foam and roams.

Starting with Lucene 1.9 an additional (optional) parameter can specify the required similarity. The value is between 0 and 1, with a value closer to 1 only terms with a higher similarity will be matched. For example:

```
roam~0.8
```

The default that is used if the parameter is not given is 0.5.

Proximity Searches

Lucene supports finding words are a within a specific distance away. To do a proximity search use the tilde, "~", symbol at the end of a Phrase. For example to search for a "apache" and "jakarta" within 10 words of each other in a document use the search:

```
"jakarta apache"~10
```

Range Searches

Range Queries allow one to match documents whose field(s) values are between the lower and upper bound specified by the Range Query. Range Queries can be inclusive or exclusive of the upper and lower bounds. Sorting is done lexicographically.

```
mod_date:[20020101 TO 20030101]
```

This will find documents whose mod_date fields have values between 20020101 and 20030101, inclusive. Note that Range Queries are not reserved for date fields. You could also use range queries with non-date fields:

```
title:{Aida TO Carmen}
```

This will find all documents whose titles are between Aida and Carmen, but not including Aida and Carmen.

Inclusive range queries are denoted by square brackets. Exclusive range queries are denoted by curly brackets.

Boosting a Term

Lucene provides the relevance level of matching documents based on the terms found. To boost a term use the caret, "^", symbol with a boost factor (a number) at the end of the term you are searching. The higher the boost factor, the more relevant the term will be.

Boosting allows you to control the relevance of a document by boosting its term. For example, if you are searching for

```
jakarta apache
```

and you want the term "jakarta" to be more relevant boost it using the ^ symbol along with the boost factor next to the term. You would type:

```
jakarta^4 apache
```

This will make documents with the term jakarta appear more relevant. You can also boost Phrase Terms as in the example:

```
"jakarta apache"^4 "Apache Lucene"
```

By default, the boost factor is 1. Although the boost factor must be positive, it can be less than 1 (e.g. 0.2)

Boolean Operators

Boolean operators allow terms to be combined through logic operators. Lucene supports AND, "+", OR, NOT and "-" as Boolean operators(Note: Boolean operators must be ALL CAPS).

The OR operator is the default conjunction operator. This means that if there is no Boolean operator between two terms, the OR operator is used. The OR operator links two terms and finds a matching document if either of the terms exist in a document. This is equivalent to a union using sets. The symbol || can be used in place of the word OR.

To search for documents that contain either "jakarta apache" or just "jakarta" use the query:

```
"jakarta apache" jakarta
```

or

```
"jakarta apache" OR jakarta
```

AND

The AND operator matches documents where both terms exist anywhere in the text of a single document. This is equivalent to an intersection using sets. The symbol && can be used in place of the word AND.

To search for documents that contain "jakarta apache" and "Apache Lucene" use the query:

```
"jakarta apache" AND "Apache Lucene"
```

+

The "+" or required operator requires that the term after the "+" symbol exist somewhere in a the field of a single document.

To search for documents that must contain "jakarta" and may contain "lucene" use the query:

```
+jakarta lucene
```

NOT

The NOT operator excludes documents that contain the term after NOT. This is equivalent to a difference using sets. The symbol ! can be used in place of the word NOT.

To search for documents that contain "jakarta apache" but not "Apache Lucene" use the query:

```
"jakarta apache" NOT "Apache Lucene"
```

Note: The NOT operator cannot be used with just one term. For example, the following search will return no results:

```
NOT "jakarta apache"
```

-

The "-" or prohibit operator excludes documents that contain the term after the "-" symbol.

To search for documents that contain "jakarta apache" but not "Apache Lucene" use the query:

```
"jakarta apache" -"Apache Lucene"
```

Grouping

Lucene supports using parentheses to group clauses to form sub queries. This can be very useful if you want to control the boolean logic for a query.

To search for either "jakarta" or "apache" and "website" use the query:

```
(jakarta OR apache) AND website
```

This eliminates any confusion and makes sure you that website must exist and either term jakarta or apache may exist.