

IRIS – Information retrieval and Integrated services

(for low bandwidth networks)

Group Name – SHAs
Members: Manjusha Nair & Jisha R C
Amrita VishwaVidyaPeetam,
Amritapuri, Kollam, Kerala, India.

Recap

- An information retrieval system intended to support efficient web search over intermittent networks.
- It supports an expanded search query interface which allows a user to specify additional query terms to maximize the utility of the results returned by a search query.
- It performs several optimizations to pre fetch pages to best satisfy a user based on their search preferences

Objective

- Search Optimisation: provide a more effective web search over low bandwidth networks.
- provide a more user friendly design to an internet searching.
- Fast & efficient content retrieval: provide quickness for the retrieval of documents.
- reduce the greater search time for the user – by E book creation.
- economical bandwidth usage.

Modules

I . Local Search

- Database lookup
- Content retrieval

II . Contextual Caching

- Focused crawling
- Link filtering and publishing
- Classification
- Compression / Content adaptation
- Caching and Indexing

- Algorithm –

Input: keyword from user

Output: retrieved Page to the browser / or to the local store

BEGIN

Input Keyword from the user

If(Input present in the local database of keywords)

 localsearch()

Else

{

 crawl();

 publishurl();

 contentadapt();

 localstore();

}

END

Module I - Local Search

- Algorithm – Local Search

Input: keyword from user

Output: retrieved Page to the browser / or to the local store

BEGIN

Input Keyword from the user

If(Input present in the local database of keywords)

{

 read user preference;

 retrieve link from local store based on the preference;

 display result in browser;

}

Else

{

 contextual caching();

 display result in browser;

}

END

Local Search

- Status
 - Integrated local search() and localstore();
 - When the user types his keyword, the module searched the keyword in the local database. If it is present, the pages are retrieved from the local store.
 - The pages are stored in the local store flexible for content adaption.
- Todo
 - User preferences (eg, bandwidth – low, medium, high) need to be taken through the user interface
 - The pages need to be retrieved based on the user preferences.
 - If the keyword is not present in the database, the crawler() should be invoked. (presently the crawler module is not integrated.)
 - Sample data base ‘Microsoft access’ is used. Need to select a DBMS considering the performance penalties.

Local Search - keyword present in the database



IRIS

asp

SEARCH

[REFRESH]

#	Query/Page Request	Status/ETA	Options
1.	\root\text\com\4guysfromrolla\aspnet\articles\012203-1.2.aspx	Unknown	[REMOVE]
2.	\root\text\com\devdirect\ALL\Imagingmanipulation_PCAT_1862.aspx	Unknown	[REMOVE]
3.	\root\text\com\vwd-cms\open-source\source-code-viewer.aspx	Unknown	[REMOVE]
4.	\root\text\com\15seconds\issue\050526.htm	Unknown	[REMOVE]

[REFRESH]

Local Search - keyword present in the database



RuralCafe Homepage x +
http://www.iris.net/

IRIS

[ASPFAQs.com](#)
[Message Board](#)
[Related Web Technologies](#)
[User Tips!](#)
[Coding Tips](#)
[Search](#)

True Image Resizing, Part 2

By [Scott Mitchell](#)

- Read [Part 1](#)

In [Part 1](#) we examined how to create an ASP.NET Web page, `ShowImage.aspx`, that would display an image. This ASP.NET Web page could then be referenced via an HTML `` tag. In this part we'll look at how to extend the `ShowImage.aspx` so that it can produce thumbnail images; then, we'll see how to tie this lesson together with the lessons learned from the article [Displaying a List of Scaled Images](#).

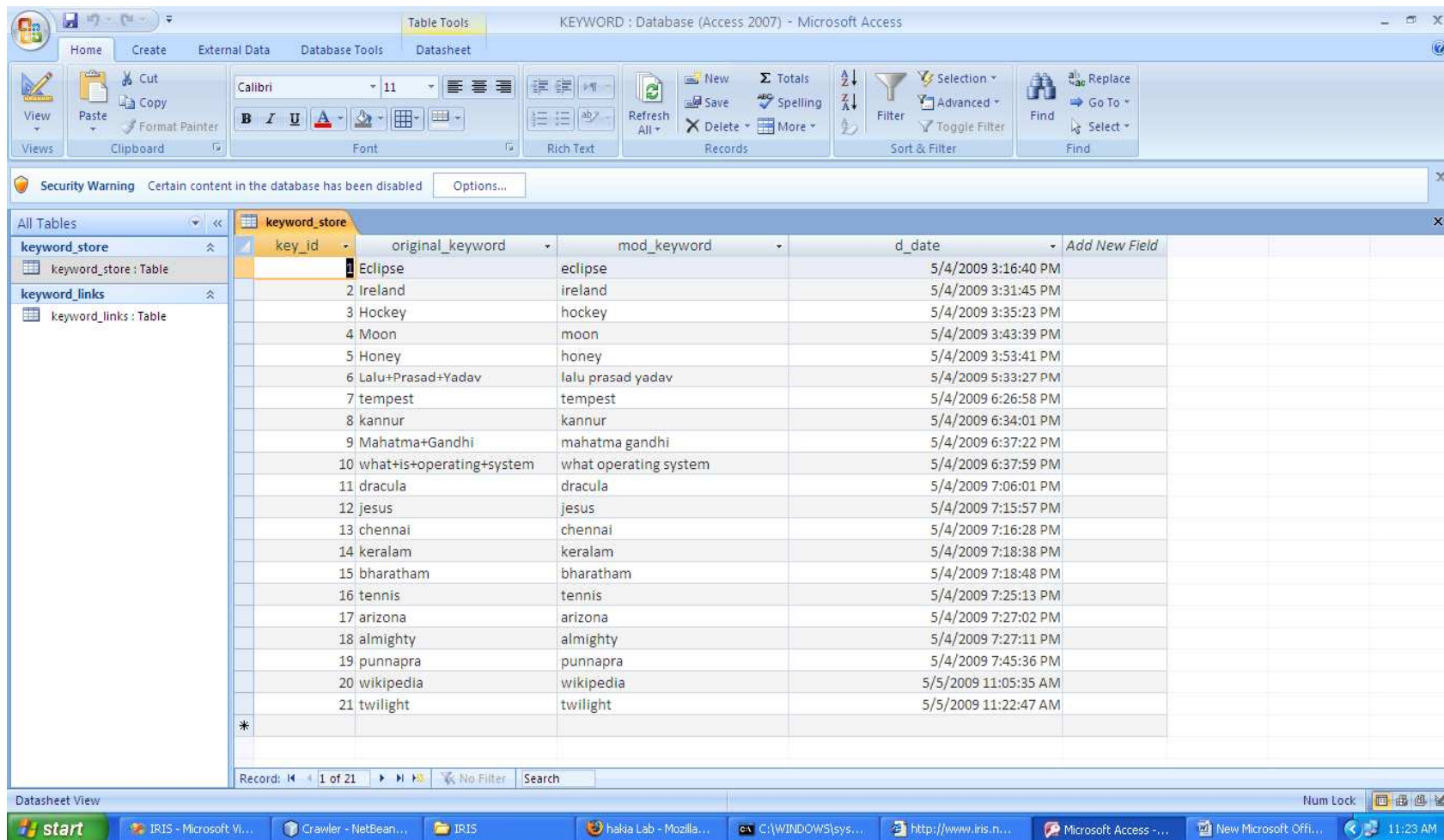
[- continued -](#)

Sections:
[Book Reviews](#)
• [Sample Chapters](#)
[Commonly Asked Message Board Questions](#)
[Headlines from ASPWire.com](#)

Altering ShowImage.aspx to Create a Thumbnail
Earlier we saw that we could create an ASP.NET Web page called `ShowImage.aspx` that could be passed in an image file name and display the image. In addition to accepting the file name, let's augment this ASP.NET Web page so that we can pass in an optional height and width. When a height and width are passed in, the image displayed is first dynamically resized to the specified height and width. The following code implements the needed changes in `ShowImage.aspx`.

```
<%@Import Namespace="System.Drawing.Imaging" %>  
<script language="VB" runat="server">  
    Function ThumbnailCallback() as Boolean  
        Return False
```

Local Search – Database of keywords



Security Warning: Certain content in the database has been disabled. Options...

key_id	original_keyword	mod_keyword	d_date
1	Eclipse	eclipse	5/4/2009 3:16:40 PM
2	Ireland	ireland	5/4/2009 3:31:45 PM
3	Hockey	hockey	5/4/2009 3:35:23 PM
4	Moon	moon	5/4/2009 3:43:39 PM
5	Honey	honey	5/4/2009 3:53:41 PM
6	Lalu+Prasad+Yadav	lalu prasad yadav	5/4/2009 5:33:27 PM
7	tempest	tempest	5/4/2009 6:26:58 PM
8	kannur	kannur	5/4/2009 6:34:01 PM
9	Mahatma+Gandhi	mahatma gandhi	5/4/2009 6:37:22 PM
10	what+is+operating+system	what operating system	5/4/2009 6:37:59 PM
11	dracula	dracula	5/4/2009 7:06:01 PM
12	jesus	jesus	5/4/2009 7:15:57 PM
13	chennai	chennai	5/4/2009 7:16:28 PM
14	keralam	keralam	5/4/2009 7:18:38 PM
15	bharatham	bharatham	5/4/2009 7:18:48 PM
16	tennis	tennis	5/4/2009 7:25:13 PM
17	arizona	arizona	5/4/2009 7:27:02 PM
18	almighty	almighty	5/4/2009 7:27:11 PM
19	punnapra	punnapra	5/4/2009 7:45:36 PM
20	wikipedia	wikipedia	5/5/2009 11:05:35 AM
21	twilight	twilight	5/5/2009 11:22:47 AM

Record: 1 of 21 | No Filter | Search

Module II – Contextual Caching

- Algorithm

Input: keyword from localsearch()

Output: content stored in local store

BEGIN

Input = keyword from localsearch();

urllist = Crawl(input);

Publish(urllist);

Contentadapt();

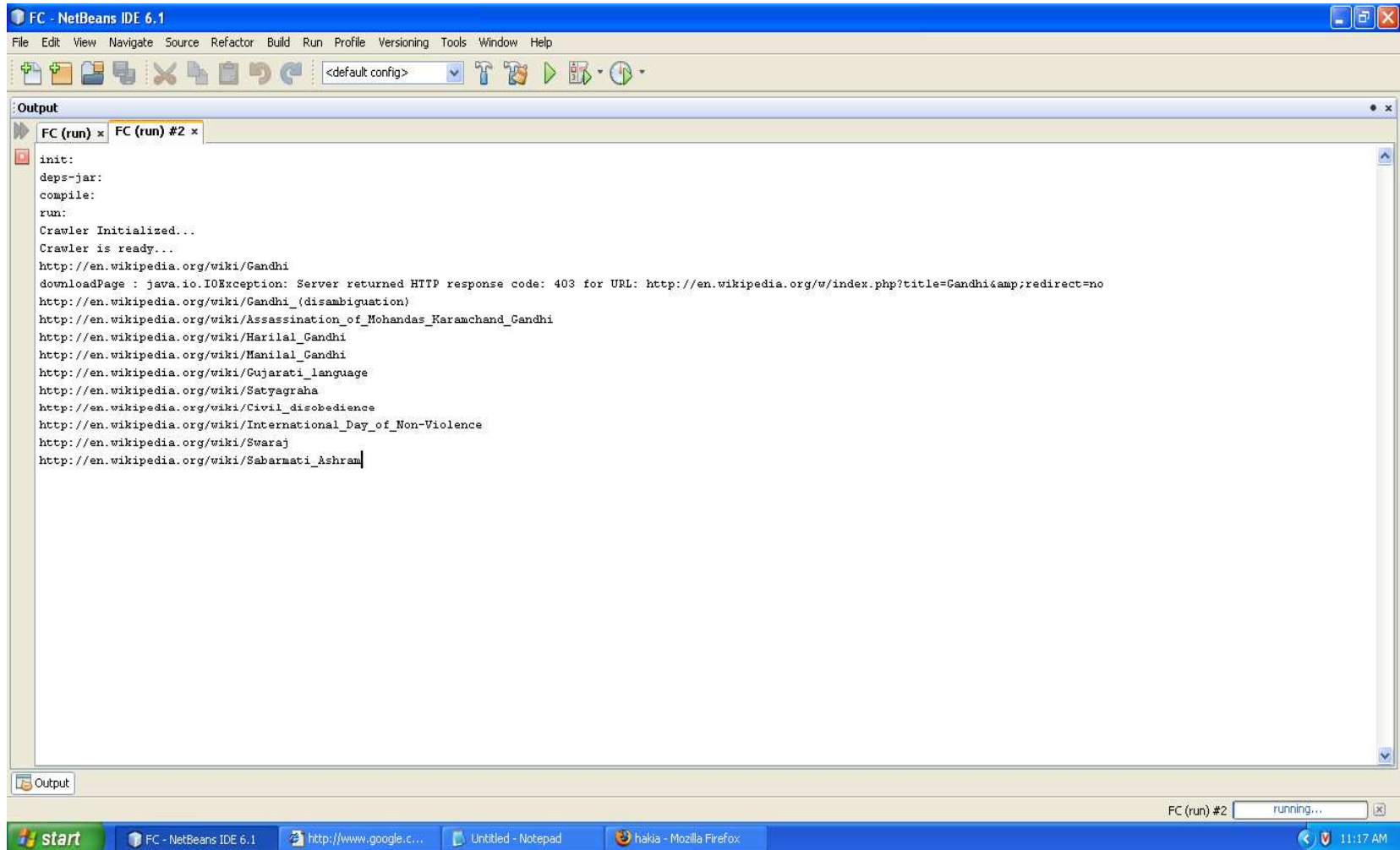
Localstore();

END;

Module – Contextual Caching

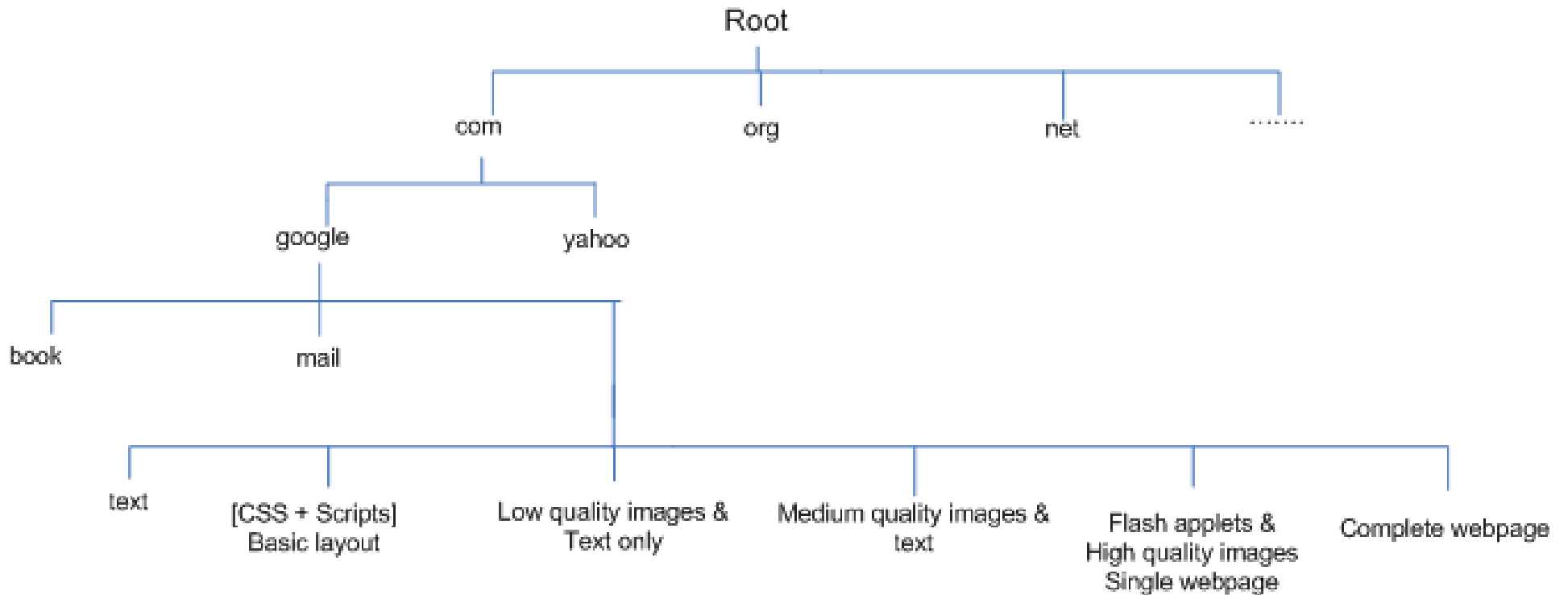
- Status
 - Crawler(Lucene.Net) takes keyword and outputs a list of URLs.
 - Authentic pages only – crawling starts from Wikipedia and hokia.(Semantic search engine – credible for health topics)
 - Pages are downloaded.
 - Content is adapted and stored in specific Directory hierarchy.
 - Update content in the local store based on the date of request.
- Todo
 - Crawler need to be integrated.
 - User preference need to be taken
 - Search should be tried with more search engines.
 - Search need to be enhanced with semantic dictionaries such as Wordnet
 - Local store need to be replaced with a storage service after considering the performance penalties.

Contextual Caching- result of crawling- list of urls'

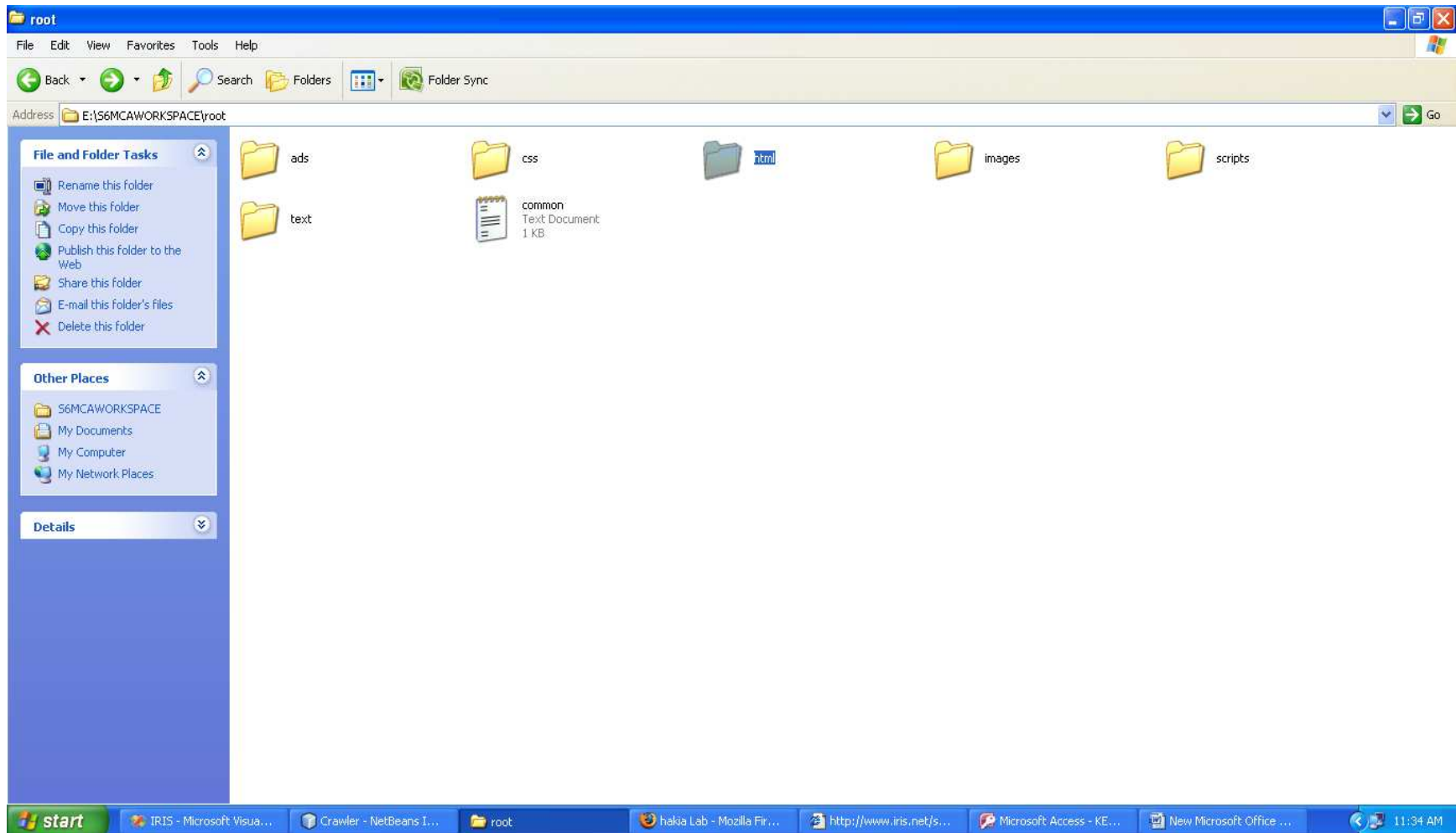


```
FC - NetBeans IDE 6.1
File Edit View Navigate Source Refactor Build Run Profile Versioning Tools Window Help
<default config>
Output
FC (run) x FC (run) #2 x
init:
deps-jar:
compile:
run:
Crawler Initialized...
Crawler is ready...
http://en.wikipedia.org/wiki/Gandhi
downloadPage : java.io.IOException: Server returned HTTP response code: 403 for URL: http://en.wikipedia.org/w/index.php?title=Gandhi&redirect=no
http://en.wikipedia.org/wiki/Gandhi_(disambiguation)
http://en.wikipedia.org/wiki/Assassination_of_Mohandas_Karamchand_Gandhi
http://en.wikipedia.org/wiki/Harilal_Gandhi
http://en.wikipedia.org/wiki/Manilal_Gandhi
http://en.wikipedia.org/wiki/Gujarati_language
http://en.wikipedia.org/wiki/Satyagraha
http://en.wikipedia.org/wiki/Civil_disobedience
http://en.wikipedia.org/wiki/International_Day_of_Non-Violence
http://en.wikipedia.org/wiki/Swaraj
http://en.wikipedia.org/wiki/Sabarmati_Ashram
```

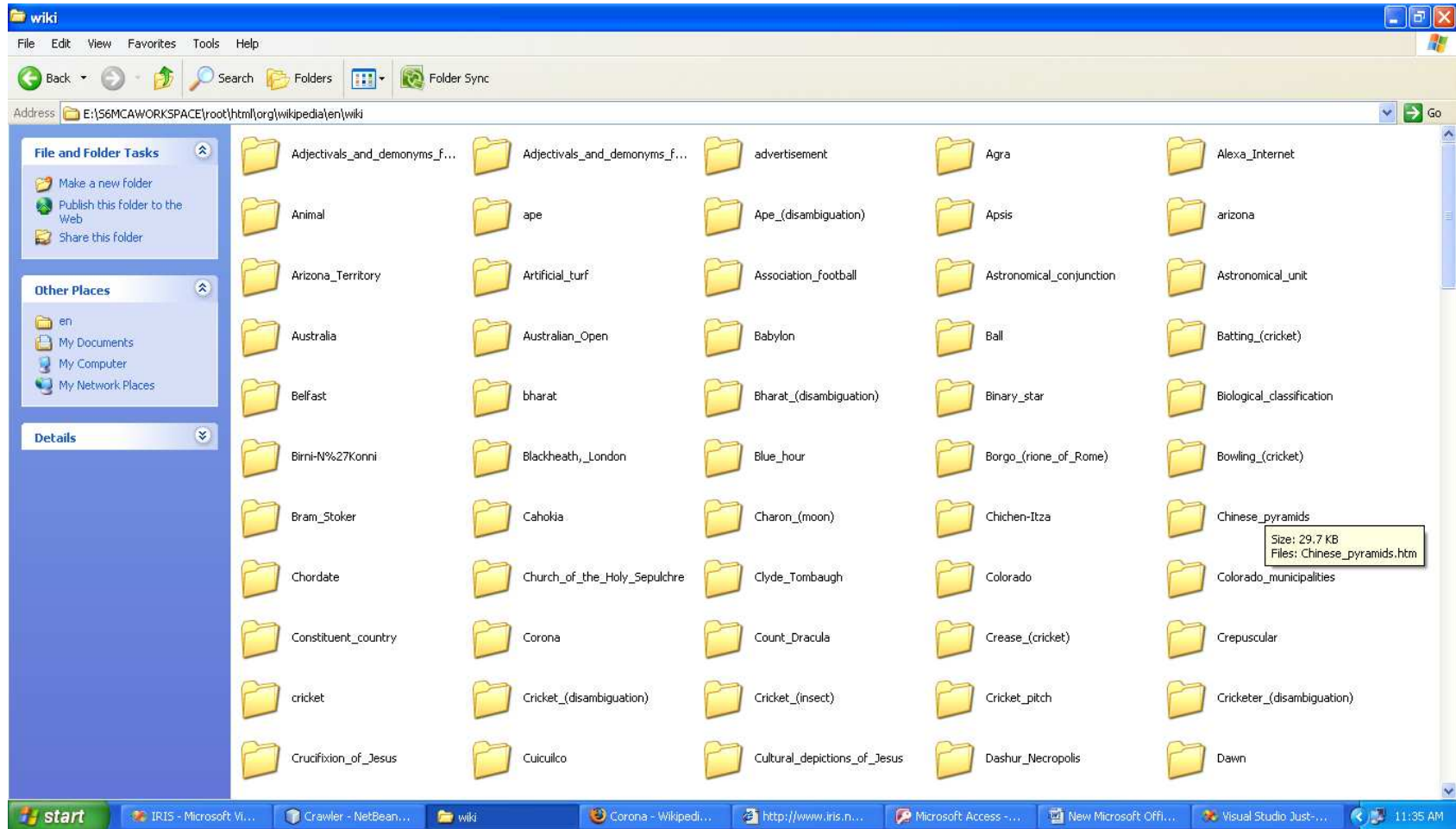
Contextual Caching- Directory structure of local storage



Contextual Caching- Directory structure of local storage



Contextual Caching- Directory structure of local storage



Modules to be added

- Automatic E book Creation and payment
- Fair sharing across users

References and Related work

- <http://cater.cs.nyu.edu/wiki/index.php/RuralCafe>
 - Rural Café: - By CATER (Cost effective Appropriate technologies for emerging regions) , New York University
- www.loband.org - For Simplification and compression of web pages to reduce bandwidth requirements.
- Lucent.net – Open source web crawler
- Wordnet – large lexical database of English interlinked by means of conceptual semantic and lexical relation.
- http://www.googleguide.com/google_works.htm
- <http://www.en.wikipedia.org>



Questions?

Thank You