San Jose State University
SJSU ScholarWorks

Master's Projects | Master's Theses and Graduate Research

2008

Semantic Categorization Of Online Video

Supraja Swarna
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_projects

Part of the Computer Sciences Commons

Recommended Citation

Swarna, Supraja, "Semantic Categorization Of Online Video" (2008). *Master's Projects*. 114.
DOI: https://doi.org/10.31979/etd.pvwj-j4v3
https://scholarworks.sjsu.edu/etd_projects/114

This Master's Project is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Projects by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

Semantic Categorization of Online Video

by

Supraja Swarna

December 1

© 2008

Supraja Swarna

ALL RIGHTS RESERVED

ABSTRACT

Semantic Categorization of Online Video

by Supraja Swarna

As internet users are increasing day by day, the users of video-sharing site are also increasing. Video-sharing is becoming more and more popular in e-learing, but the current famous websites like youtube are not structured when it come to serving the purpose of providing educational videos for preschool and high school students. There is a need to fill building more educationally focused video site, where the content is more structured, easy to use, support both direct search and browsing, and follow a particular curriculum for preschool and high school students. This report discuss the issues like categorization and search interface of these sites and propose alternatives to existing ones out there. In this project, I have built an educational website for preschool, high school, and college level students concentrating on improved categorization and search interface of the site. This report provides detail description of my system and the results of comparison between my site and youtube.

Table of Contents

1. Introduction....6

2. Categorization....7

3. Faceted Search....9

3. 1 My Site Search....10

3.1.1 Keyword Search....10

3.1.2 Guided search: 12

4. Comparison of my site with Youtube.com....13

4.1 Categorization Comparison....13

4.2 Search Interface Comparison....15

5. Implementation....21

5.1 Drupal Installation:....22

5.2 Embedded Media Field Module:....26

5.2.1 Embedded Media Field Installation:....26

5.2.2 How to make the video size bigger:....29

5.3 Taxonomy Module....31

5.3.1 Installation and Usage:....31

5.4 Taxonomy Manager Module....34

5.4.1 Installation: 34

5.4.2 Usage:....34

5.5 Jrating Module:....36

5.5.1 Installation:....36

5.5.2 Usage:....37

5.6 Tagadelic and Community Tags Module:.....38

5.6.1 Installation:.....38

5.6.2 How to create Tag Cloud:.....40

5.7 Taxonomy Menu Module:.....40

5.7.1 Installation:.....40

5.7.2 Enable Number Of Items in each Category:.....41

5.8 Links Package Module.....42

5.8.1 Installation:.....42

5.8.2 How to add the links:.....44

5.9 Forward Module.....45

5.9.1 Installation.....45

5.9.2 How to forward the Specific node.....45

5.10 Faceted Search Module:.....46

5.10.1 Installation:.....46

5.10.2 Faceted Search: Key word search auto complete functionality: 48

6. Future work.....49

7. References.....49

1. Introduction

Video-sharing is becoming increasingly popular on the World Wide Web. It allows users to easily understand many things and get a good grasp on the subject in little time. There are so many video-sharing websites that are informative and educational such as sutree.com. The main problem with those websites is that they are poorly categorized and use the standard search interface, in which users enter keyword search terms, and results are shown in table format ordered by some measure of relevance. These sites can be effective for searching for very specific items, but do not support browsing and exploratory tasks well [7, 8, 9].

This report discusses how to build a website with meaningful categorization with semantics and how can we modify the standard search interface so that the search is more useful and informative to the users. This report also will give an overview of the website that I have built for this project to provide the concrete examples to support the ideas I have explained in this report.

The first concept of this project is “categorization” of the site. Categorizing the content on the site is an important aspect of the site design. If the site is not organized carefully, then user can feel lost at some point of browsing. One can build meaningful categorization by studying the domain carefully and picking the relevant and right terms for a specific site. Section 2 of this report gives in depth description of categorization and the use of semantics of categories in my site. The second concept of this project is “search interface”. The standard search interface is more useful when the users search for a specific item, but not when the users want to learn the content by browsing and

exploring. The faceted search interface integrates searching and browsing continuously. It supports exploration and learning and also avoids possible dead ends while searching. Section 3 gives in depth overview of faceted search and its advantages and disadvantages over standard search.

In this project I have built an online community for sharing educational videos from grade k-12 and College in specific subjects. This website implements the both the concepts: semantic categorization and faceted search. This project try to fill a need for more educationally focused, safe place for children, teachers, and home learners.

The rest of this report is organized as follows: Section 2 gives a detailed description of the categorization of the site. Section 3 provides a detailed description of faceted search. Section 4 provides comparison of my site with youtube. Section 5 provides the implementation details of the site. Section 6 concludes the report, and discusses future directions.

2. Categorization

Categorizing the content on the site is an important aspect of the site design. If the site is not organized carefully, then user can feel lost at some point of browsing. One can build meaningful categorization by studying the domain carefully and picking the relevant and right terms for a specific site. In this project, I will be referring categorization as "Taxonomy". For my project I studied different websites such as "youtube.com (most famous video sharing site), sutree.com, and teachertube.com (educational video site) to understand the taxonomies used by different sites in education. Finally I have decided to go with traditional based categorization as combining subject and class level.

In my project, the taxonomy of the video site was built according to the "California Department of Education (http://www.cde.ca.gov/be/st/ss/)" standards. Currently the site has about 500 videos in total and planning to add more in future. The site has three main categories: GradeK-7, Grade8-12, and college and professional. Each category has sub categories. Each sub category allows the users to enter synonyms for each term. This feature eliminates the taxonomy duplication. Example: In taxonomy of Gradek-7, a term "Mathematics" may have a synonym "Math". User should be able to type "Math" in a search filed and submit and have it matched up with "Mathematics" in the database and get the "Mathematics" results.

The Gradek-7 category will have the videos for children who are studying from preschool to grade7. This category has 4 sub categories: English, Mathematics, Science, and History-Social sciences. The Grade8-12 will have the videos for children who are studying from grade 8 to grade 12. This category has five sub-categories: Biology, Mathematics, Physics, Chemistry, and SAT. The Biology sub-category has one more level of sub-categories. It has Cell Biology, Ecology, Evolution, Genetics, Nuclear Process, and Physiology. The mathematics also has one more level of sub-categories. It has 7 sub-categories: Algebra, Calculus, Geometry, Linear Algebra, Mathematical Analysis, Probability and Statistics, and Trigonometry. The Chemistry also has one more level of sub-categories. It has 10 sub-categories: Acids and Bases, Atomic and Molecular Structure, Chemical bonds, Chemical Equilibrium, Chemical Thermodynamics, Conservation of matter and Stoichiometry, Gases and Their Properties, Nuclear Processes, Organic Chemistry and Biochemistry, Reaction Rates, and Solutions. The Sat category will have all videos related to SAT test and doesn't have has any sub-categories.

The College and Professional category will have videos for college students and professionals in industry. It has 8 sub-categories: Business, Engineering, humanities and Arts, Law, Mathematics, Medicine, Science, and Social Science.

3. Faceted Search

Faceted search is also called as faceted navigation or browsing. It is a technique for accessing a collection of information represented using a faceted classification, allowing users to explore by filtering available information [Wikipedia: http://en.wikipedia.org/wiki/Faceted_browser].

Instead of waiting for the user to create structured queries from scratch, a faceted search interface allows the user to progressively narrow down the choices by choosing from a list of suggested query refinements [2].

The important element of the faceted search is that there is no enforcement of hierarchy of options. Users can apply and remove facet constraints in any order. There will be no surprises while searching. The user is only given facets and constraints that make sense in the context of items they are searching. In advance, the result count for a particular constraint will be shown to the user, but the users never reach an empty result set. Facets are categories used to characterize information in a given site. In an information collection that supports faceted search, multiple labels are assigned to each item, as opposed to a strictly hierarchical system in which items are placed into single categories or folders **[1].** In the faceted navigation interface, when a term is selected by a user, all items that have been assigned to that term are retrieved, so selecting a term within a facet hierarchy is equivalent to querying on a disjunction over all the terms under the selected

one. When terms from different parts of the hierarchy are selected, the system in effect builds a conjunct of disjuncts over the selected terms and their subcategories.

3. 1 My Site Search

My site supports both direct search and browsing. In my project, I have used faceted search concepts and build facets based on categories instead of meta data of the site as most of the faceted search sites do. My search adopts the concepts from faceted search and also the category search. This search gives results that are themselves entry points into faceted navigation. If a user performs a category search and then selects a result of that search, then the resulting navigation state is identical to that achieved if the user selects the same facet value through navigation.

My site provides two ways for searching the content: one is using keyword search, where the user enters search terms and the other is using guided search, which is more like a navigation and discovery tool. The user can start at any search and use both searches at the same time to narrow down his results. Basically, Keyword search dims at the users who knew exactly what they wanted and the guided search aims at users who are not sure of what they wanted.

3.1.1 Keyword Search

Keyword search provides standard search functionality but also allows the user to search within results. It also interacts with guided search to give the user hint to what to search for next. It also provides different options with search the title, search any where “with all of the words, with the exact phrase, with at least one of the words, without any words” as we can see in figure 1.

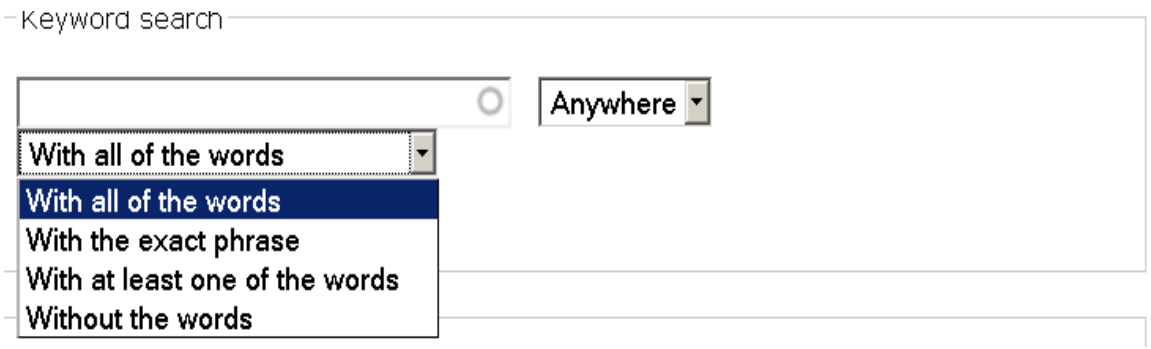


Figure 1: Screenshot of my site keyword search with different filtering options

The keyword search also has "auto complete" functionality as shown in figure 2. Auto complete makes the user not only type less but also it spares him the frustrating experience of over-specifying a query and then being presented with no relevant results.

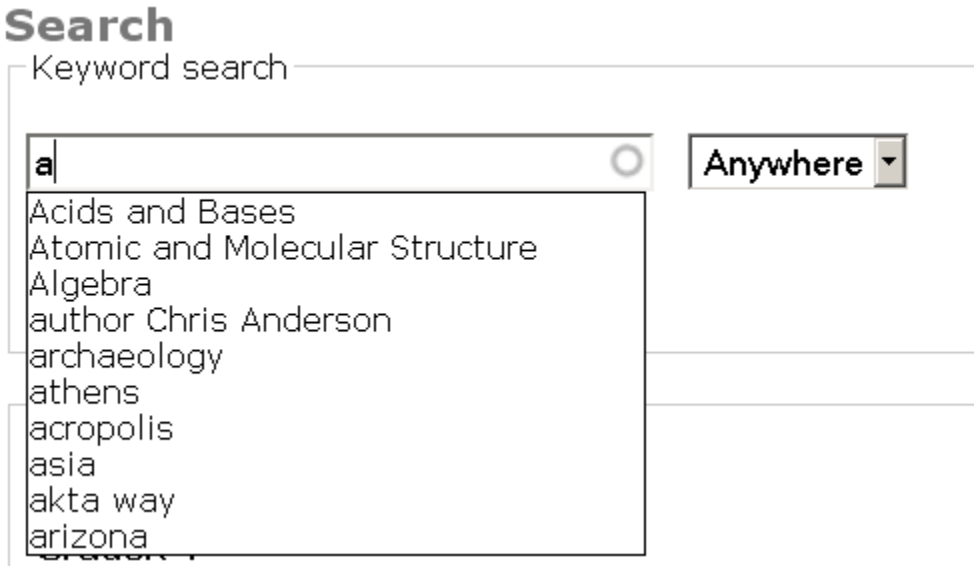


Figure 2: Screenshot of my site keyword search with auto complete functionality

For example, the user is searching for information about the Taylor's theorem in calculus. However, either he has forgotten the exact name or he is not sure about spelling variants (Tylor, Toylor..). He starts typing the letter T and the auto complete brings up all the matching terms that starts with T. User can see list of matching terms and can easily recognize the word he wants to search for and starts his search.

3.1.2 Guided search:

Guided search continuously integrates with keyword search. It gives the user feeling of "browsing the shelves". It helps users to identify what kinds of things are available on the site. It reduces mental work. The user only has to recognize the things, doesn't have to recall them.

Example:

Suppose the user is searching for basic mathematical videos. The user can browse and search at the same time using Guided search. User starts his search by clicking the term "Mathematics" under Gradek-7 in Guided search block as in figure 3. The system will display the results for "Gradek-7 Mathematics". At the same time, the system will display the matching tags under the term Mathematics, so that user can easily recognize some terms that interests him and start refining his search results.

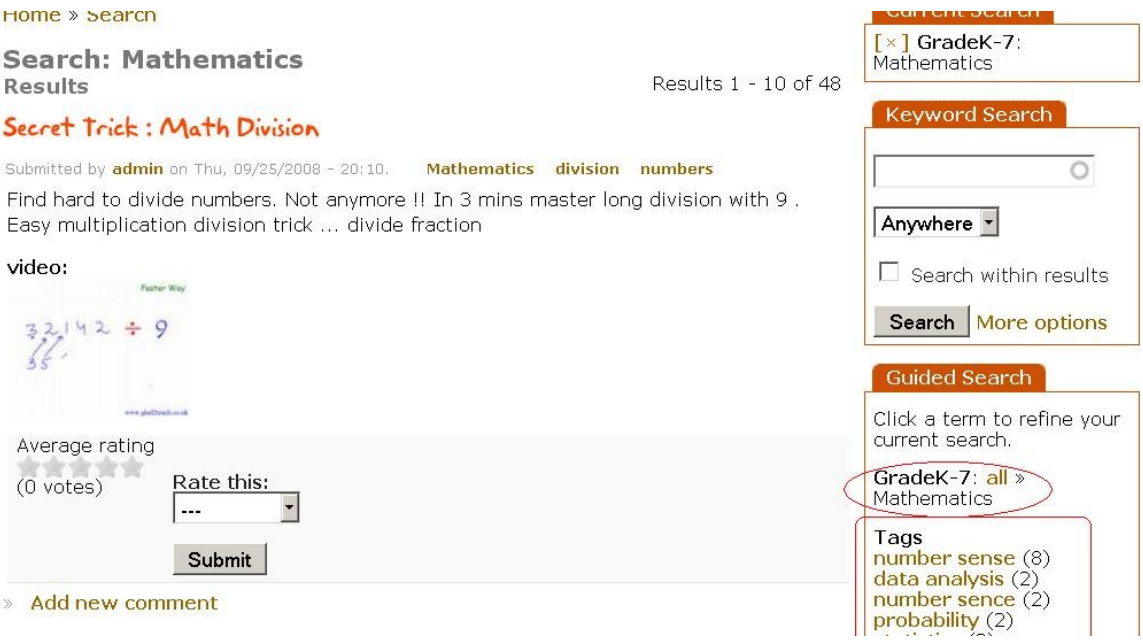


Figure 3: Screenshot of guided search with search term "Gradek-7 Mathematics"

There is also a "current search" block for the user to see what he has searched so far and also to allow the user to remove particular search term form the whole search at anytime as shown in figure 4.

Example: If the user has searched the videos for "high school-> mathematics-> Algebra->Alternate interior angles" so far. Then the current search block will show the search path so far the user has chosen.

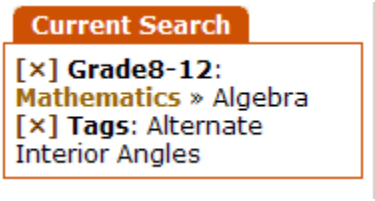


Figure 4: Screenshot of current search with search term "Gradek-7→Mathematics→>Algebra→ Alternate Interior Angles".

Suppose, now the user want to look at all videos with "Alternate interior angles" tag then he can simply remove the "grade8-12 Mathematics" from the search by clicking "x" mark left to each term in current search block. The system will displays the results for all videos from all categories with matching tag "alternate interior angles".

4. Comparison of my site with Youtube.com

4.1 Categorization Comparison

As we can see at http://www.youtube.com/browse?s=mp&c=27&l=, the site doesn't support multi level categorization. It has only one high level categorization called "education" and there are no sub categories that to help the users to navigate the site further in that area. However, it allows the users to create folksonomies by entering tags. However, in this way, there is a high chance of creating duplicate terms that makes the site overwhelming.

As we can see in figure 5, the site supports both pre defined taxonomy terms with subcategories and also taxonomies created by user entered free tagging. This approach gives the users flexibility to choose whether to use already defined taxonomy terms or to create their own terms to include his node item. As shown in figure 6, the site also allows the user to enter synonyms for a particular taxonomy term to eliminate duplication of taxonomies.

Figure 5: Screenshot of my site with Predefined taxonomies

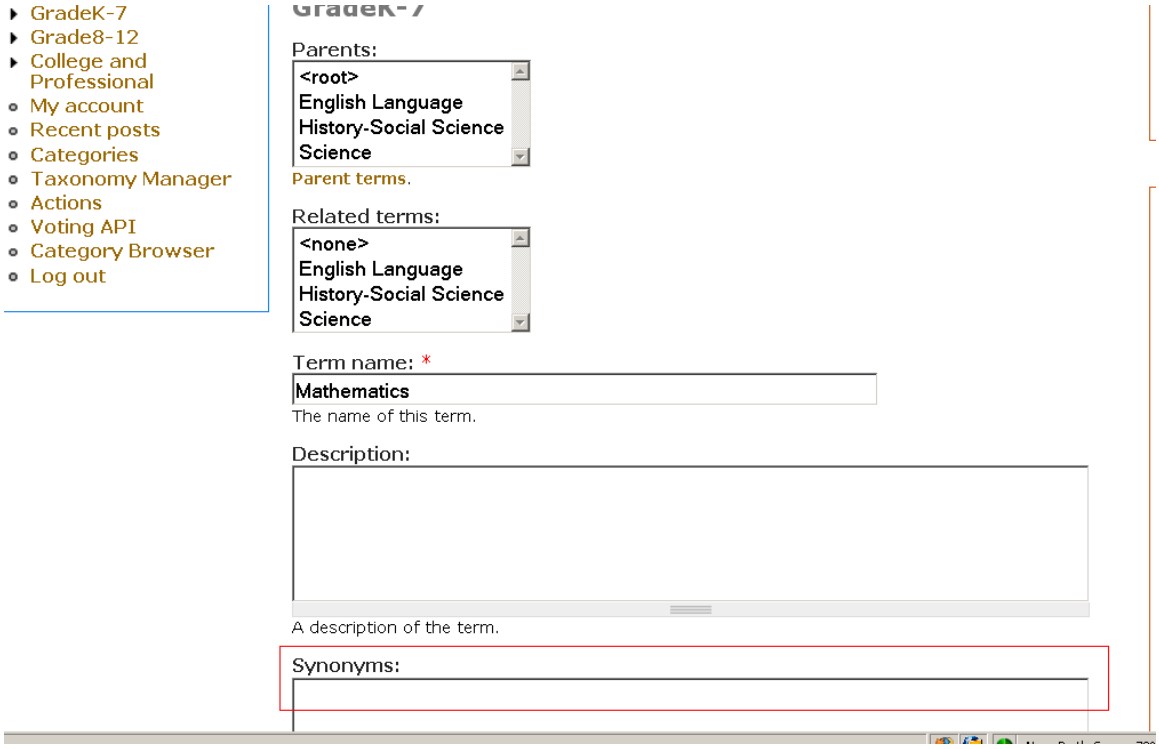


Figure 6: Screenshot of category allowing users to synonyms for taxonomy term

4.2 Search Interface Comparison

Example: Suppose a 8th grade student is studying for algebra exam. The following two tables show the detailed steps of how my system and youtube system reacts to his actions.

You Tube:

Step	Tester Action	System Response
1	User enters "Algebra" as a search term	System displays thousands of videos that match the term "Algebra". (Figure 1)
2	Now user thought it is going to take him so much time to go through each video and learn each concept. So he decided to study the material on "complex numbers" at high school level.	System displays the videos that match the term "High school Algebra complex numbers". Still it has hundreds of videos.

	Now he has to restart his search by retyping the whole phrase as High school Algebra complex numbers.	
3	Since the results are in hundreds, the user decides to look at the videos only on quadratic formulas in complex numbers. Now user wants to refine the above results. Again User has to restart the search by entering the whole phrase High school Algebra Complex numbers quadratic formulas.	System displays the results that match the all the terms in the phase “high school algebra complex numbers quadratic formulas”. Now the results are about 30 videos.

Table 1: Youtube Search test case

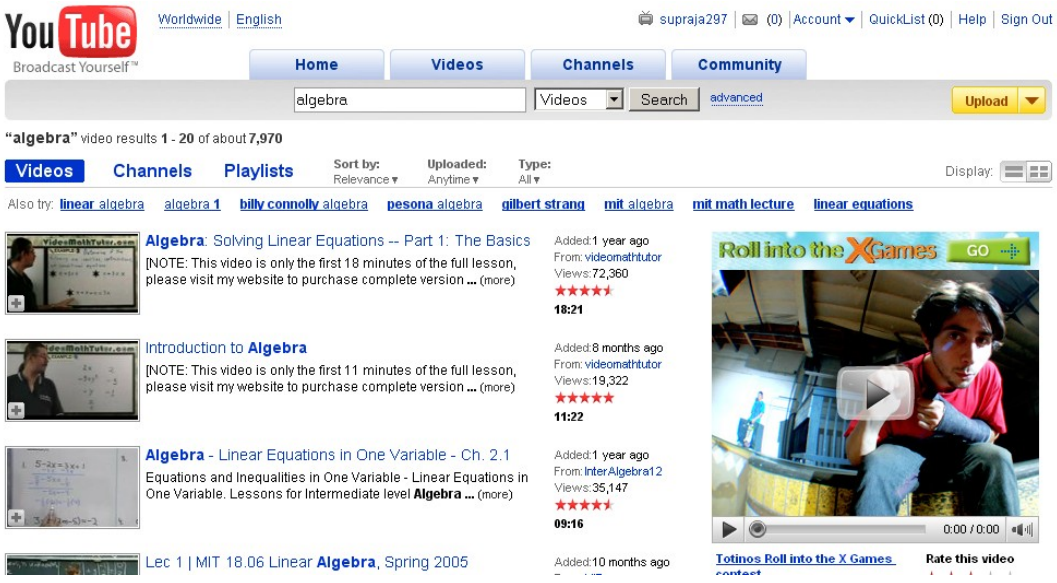


Figure7: Youtube Search for term Algebra

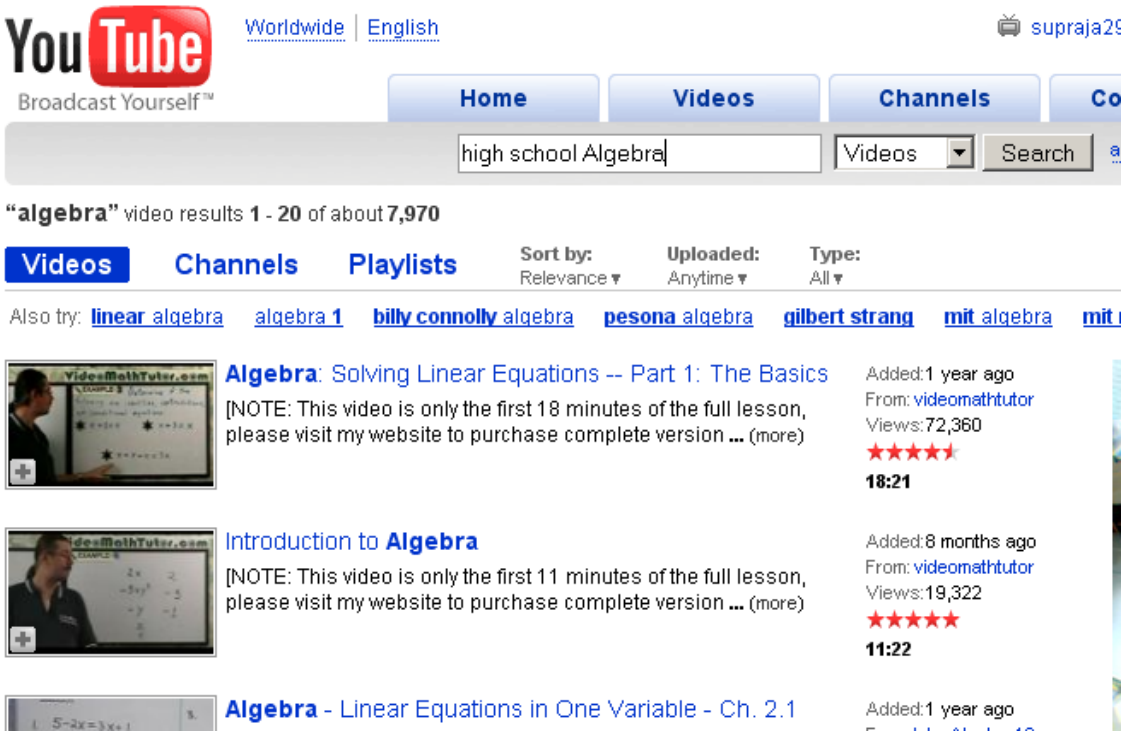


Figure 8: Youtube Search for High School Algebra

My System:

Step	Tester Action	System Response
1	User enters "Algebra" as a search term in "Key word Search" block.	System displays all the videos that match the term "Algebra". At the same time the "Guided Search" block just below the "Keyword Search" displays only the categories that match the term "Algebra" with the count. (Figure 3)
2	Now user thought it is going to take him so much time to go	System displays the videos that match the term "High school Algebra".

	through each video and learn each concept. So he decided to study the material on "complex numbers" at high school level. User can simply select the "Algebra" sub category under "Grade8-12→ Mathematics" in guided search block.	At the same time "Guided Search" block will display the tags with node count that are only related to the term "High school Algebra" to give the user chance to refine further. (Figure 4)
3	Since the results are in hundreds, the user decides to look at the videos only on quadratic formulas in complex numbers. Now user wants to refine the above results. Now user can select "search within results" checkbox in keyword search block and type in "quadratic formulas" in search box.	System displays the results that match the all the terms in the phase "High school Algebra Complex Numbers quadratic functions". (Figure 5)

Table 2: Mysite Search testcase

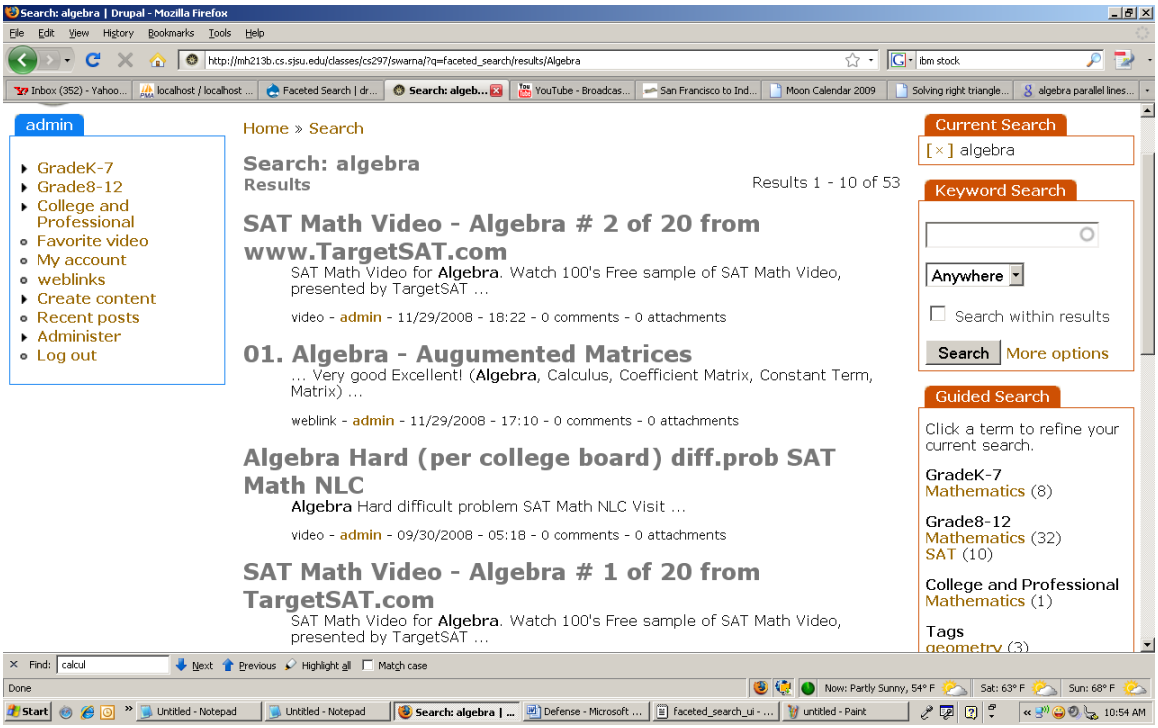


Figure 9: Mysystem Search for term Algebra

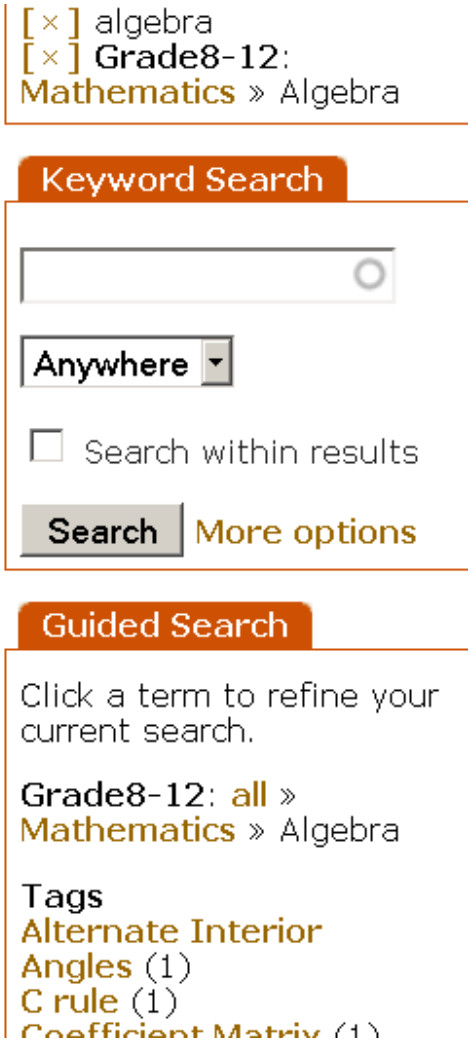


Figure 10: Mysystem Search for term High school Algebra

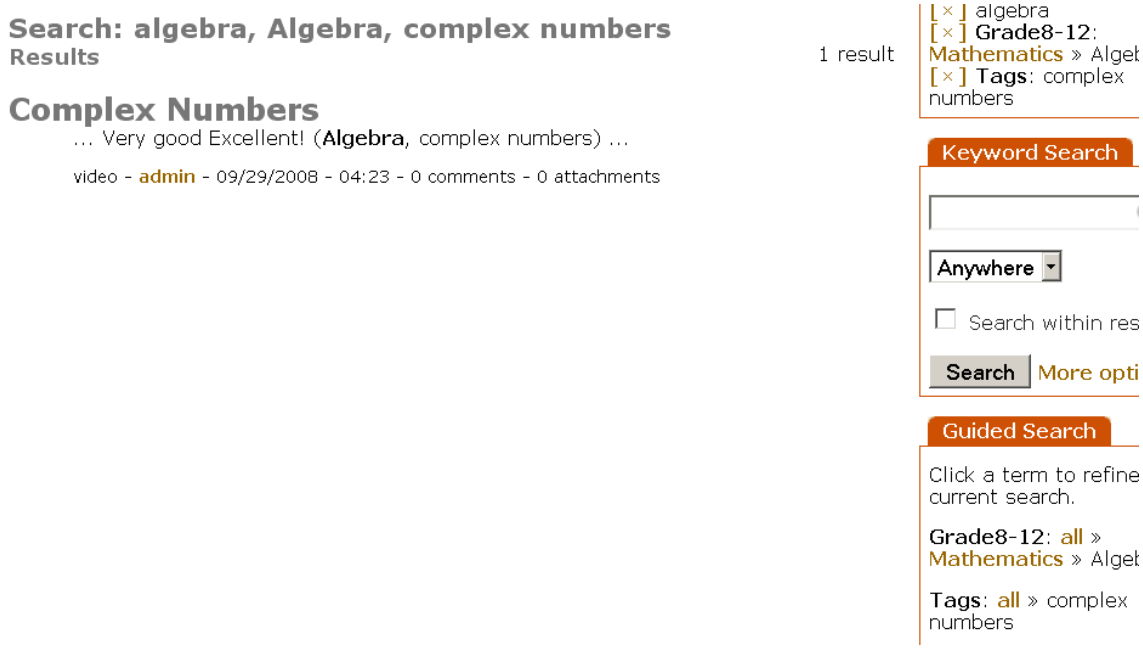


Figure 11: Mysystem Search for term High school algebra complex numbers

Conclusion:

As we can see above, using "You tube search" if the user wants to filter his results he has to restart his search, there is no way of filtering his results in the middle of the search. In this way, user gets frustrated if he has to restart search every time.

Using "My System" User can filter his results multiple times in the middle of the search without restarting new search. This way user feels more interactive with the system. In each step system is giving him some hint what to search for next. By looking at count with each term, user can decide whether he wants to refine his search or not before hand.

5. Implementation

To build the site, the Drupal Content Management System Software has been used.

This section covers module configuration and installation, and code customization. The following modules have been used from Drupal to build the site.

5.1 Drupal Installation:

1. Download Drupal latest version from http://drupal.org/download
2. Extract the files into a directory called drupalx-x.
3. Move the contents of that directory into a directory within your web server's document root(…webroot/htdocs/drupal57)
4. Go to http://localhost/phpMyAdmin
5. Create new database called drupal57
6. Go to http://localhost/drupal57 and enter the database details for the database which you have created in step 5.

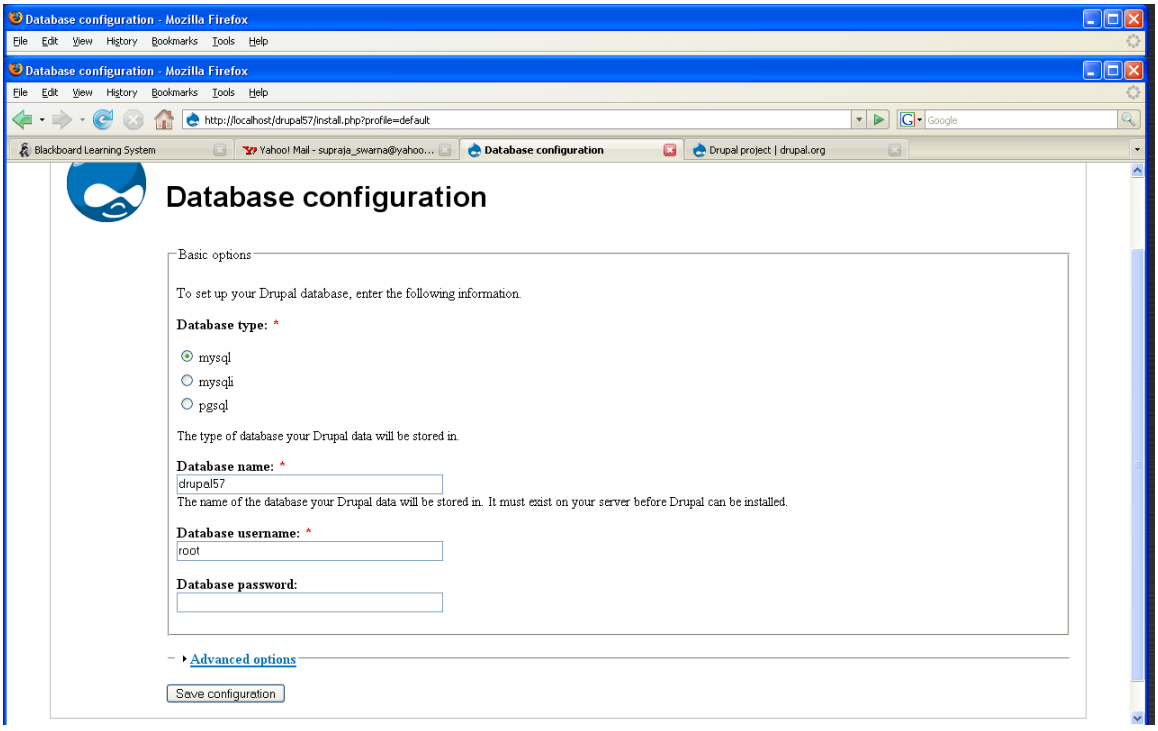


Figure 12: Screenshot of database configuration for drupal installation

7. Once you enter all the details, Installation complete page displays.

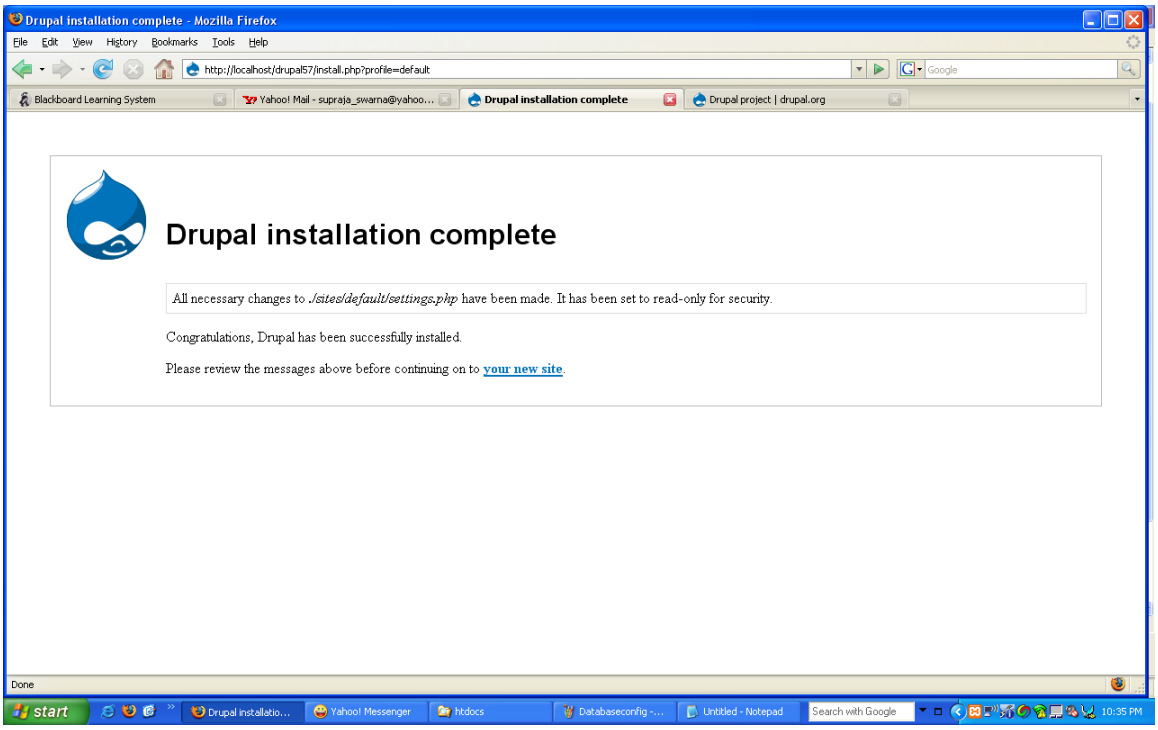


Figure 13 : Screenshot of drupal installation complete page

8. Click on “your new site” link on Drupal installation complete page.
9. Drupal doesn't automatically create admin account, so you have to create admin account by yourself.
10. Click on “create the first account” link as shown in figure

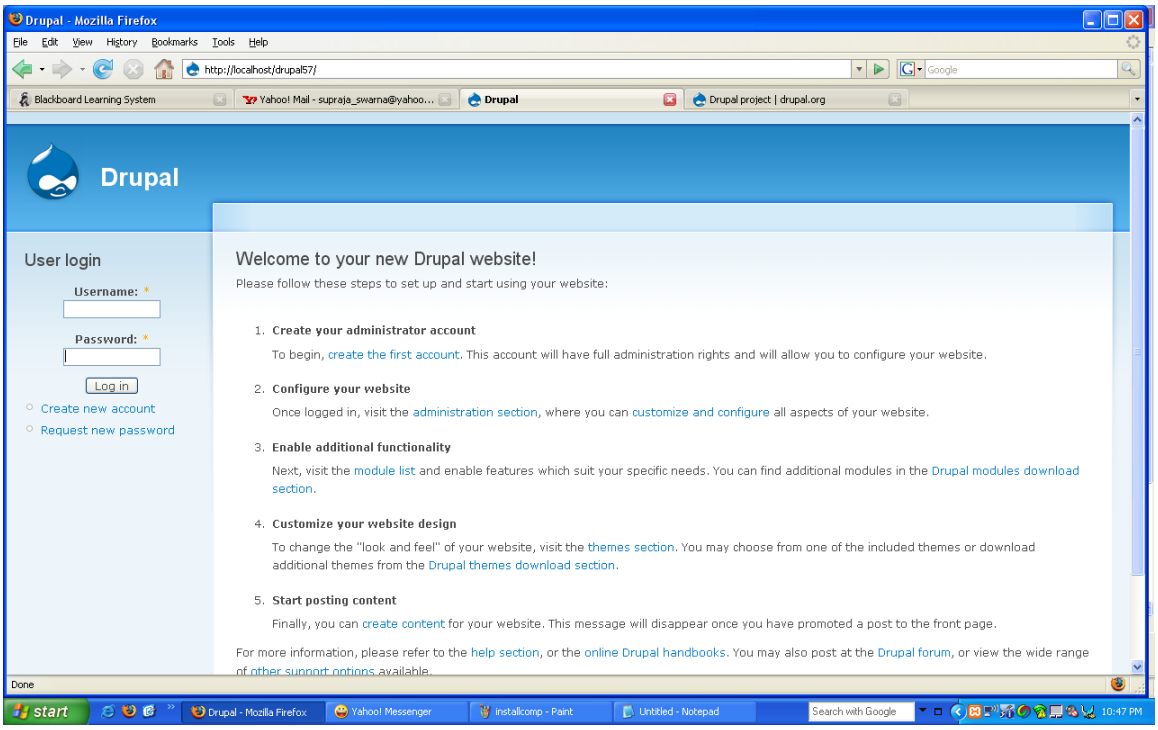


Figure 14: Screenshot for creating new account

11. Enter username and email address and click create new account button as in figure .

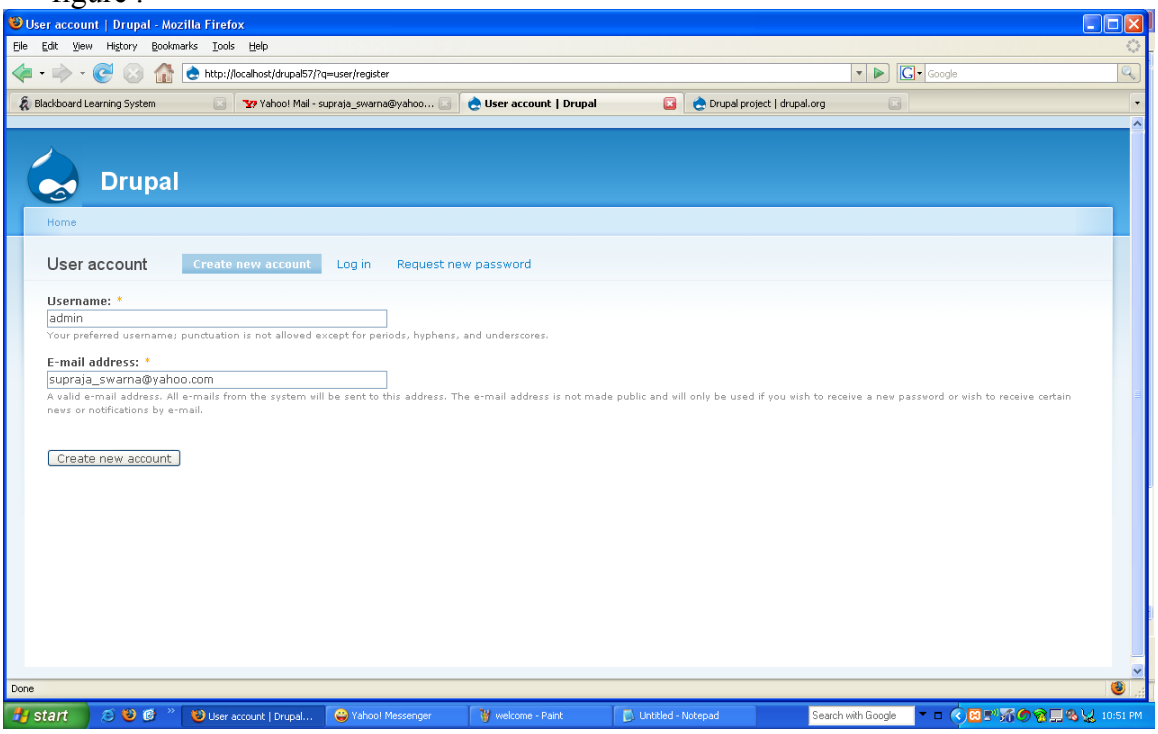


Figure 15: Screenshot for creating new account form.

12. After creating an account, the system will give a password.
13. You can change the password by clicking on Edit button.

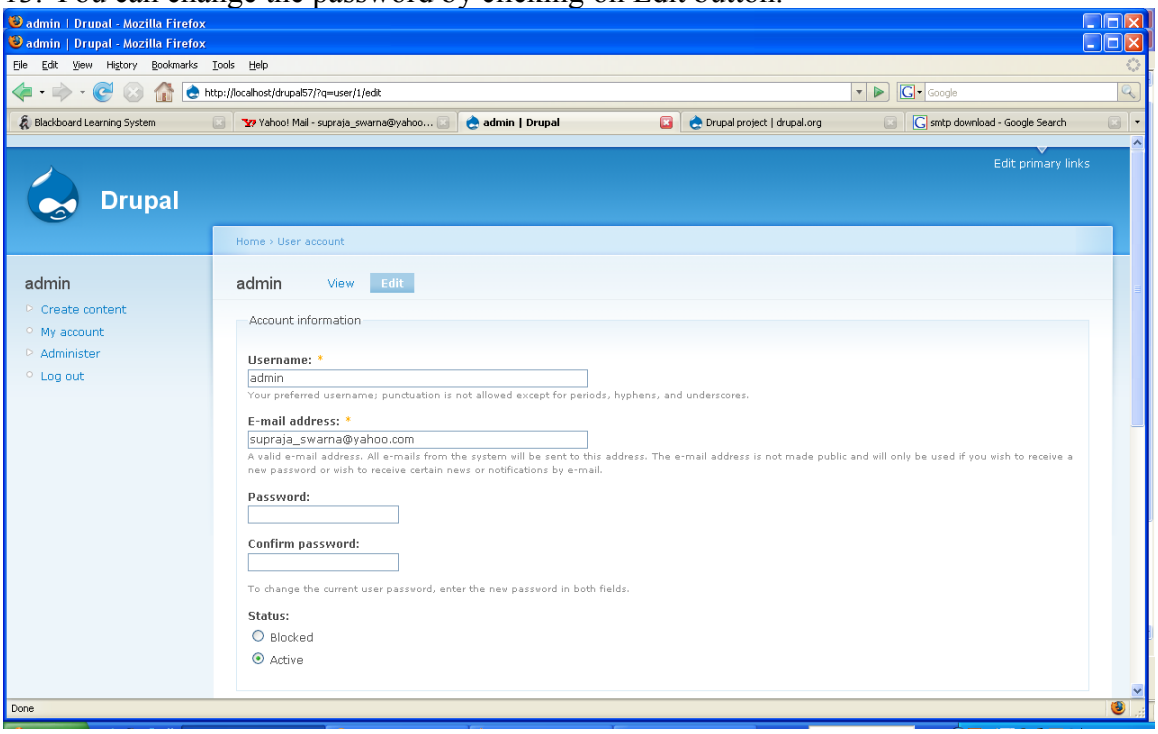


Figure 16: Screenshot for editing the user account form

5.2 Embedded Media Field Module:

To upload videos, you need Content Construction kit module (CCK) and Embedded Media Field module.

This module supports You Tube, Google, Revver, My Space, MetaCafe, Jump Cut, Bright Cove, Seven Load, iFilm, Blip.TV, Daily Motion, and Live Video video formats.

5.2.1 Embedded Media Field Installation:

1. Download module CCK (http://drupal.org/project/cck)
2. Download module Embedded Media Filed(http://drupal.org/project/emfield)
3. Unzip both of them and copy into modules sub-folder.
4. Go to Administrator\Site Building\Modules and enable both (CCK and Emfield)

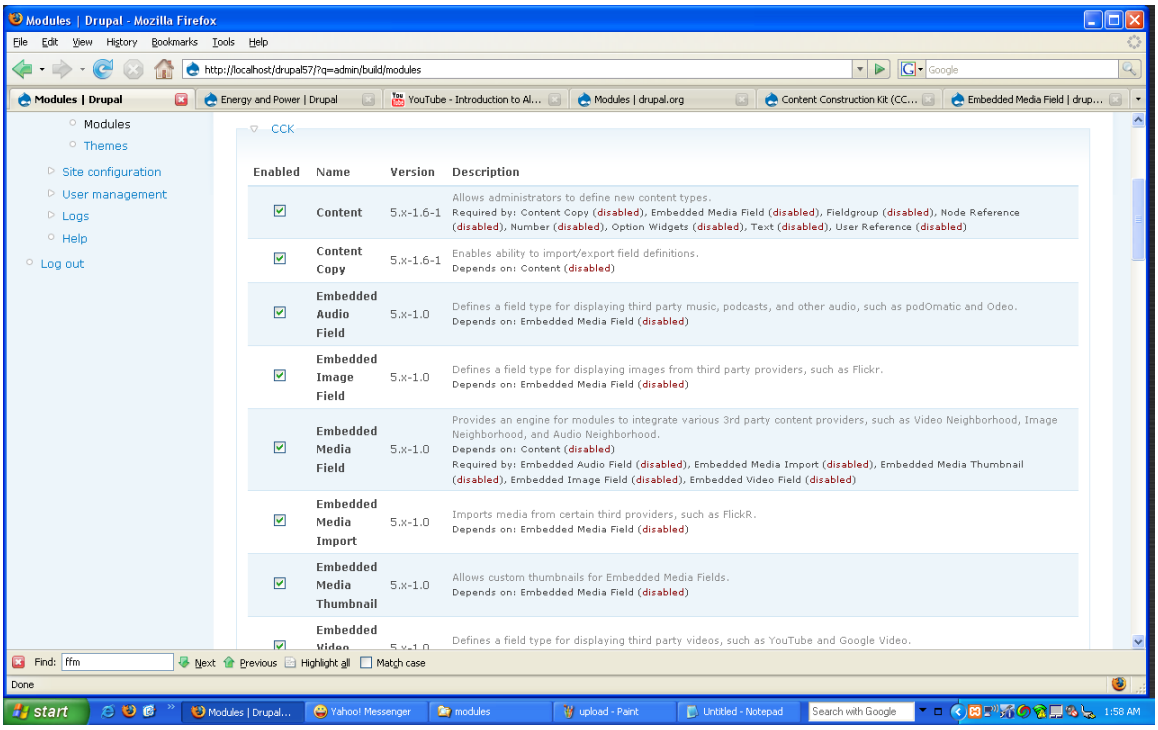


Figure 17 : Screenshot of enabling module page

5. Go to administer/content types/video/edit and click on Add field" button on top.

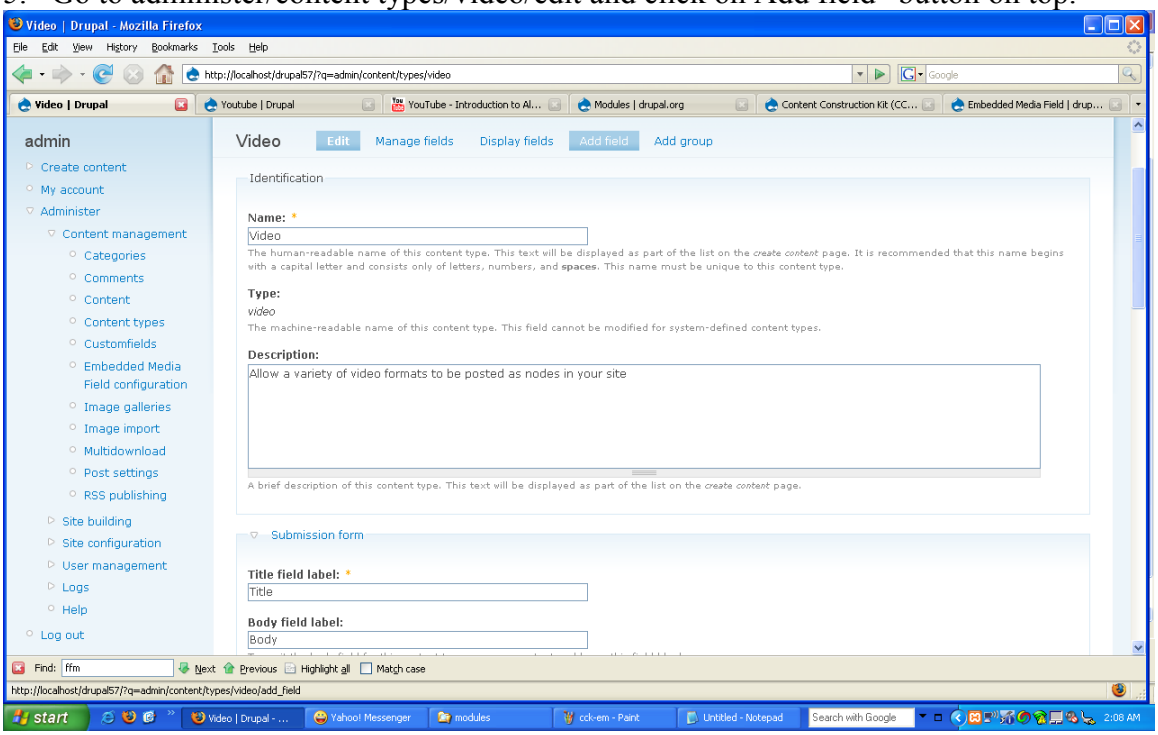


Figure 18: Screenshot of adding a video field using "Embedded media filed" module.

6. Enter Video as Title and choose Embedded Video 3rd party at the bottom of the page.

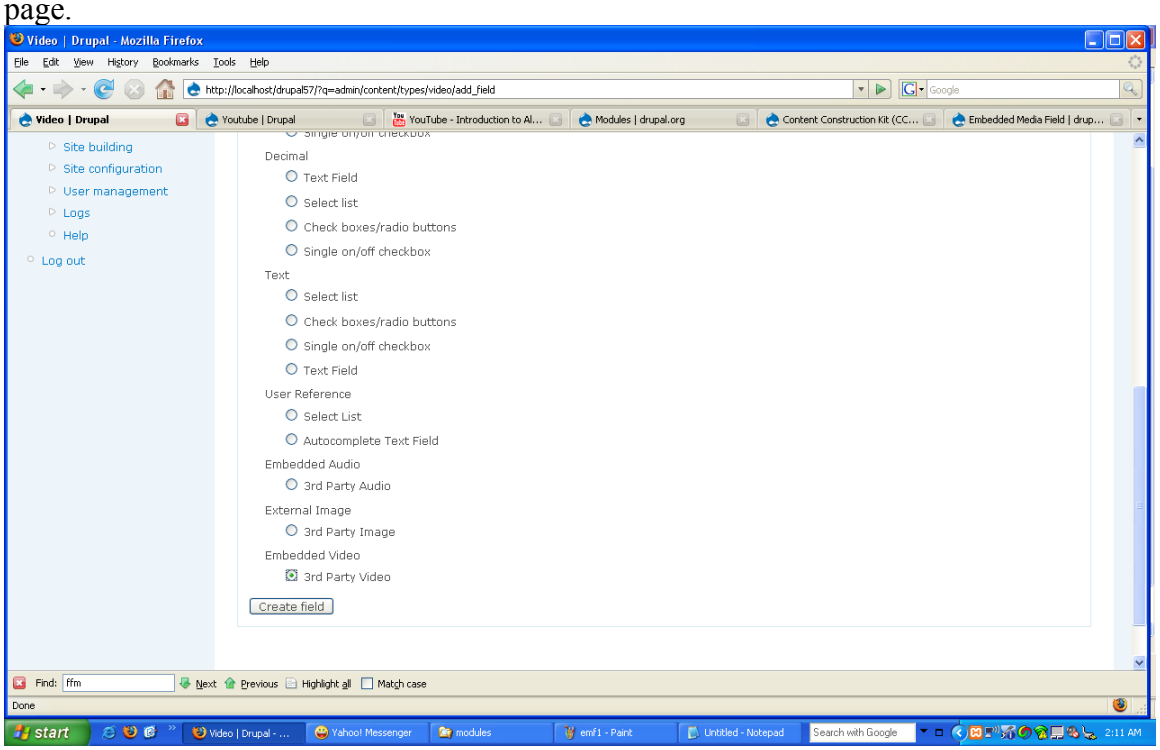


Figure 19: Screenshot of settings of video field to embed third party videos

7. Submit "create field".
8. Now choose from which providers you want to upload videos to your site.

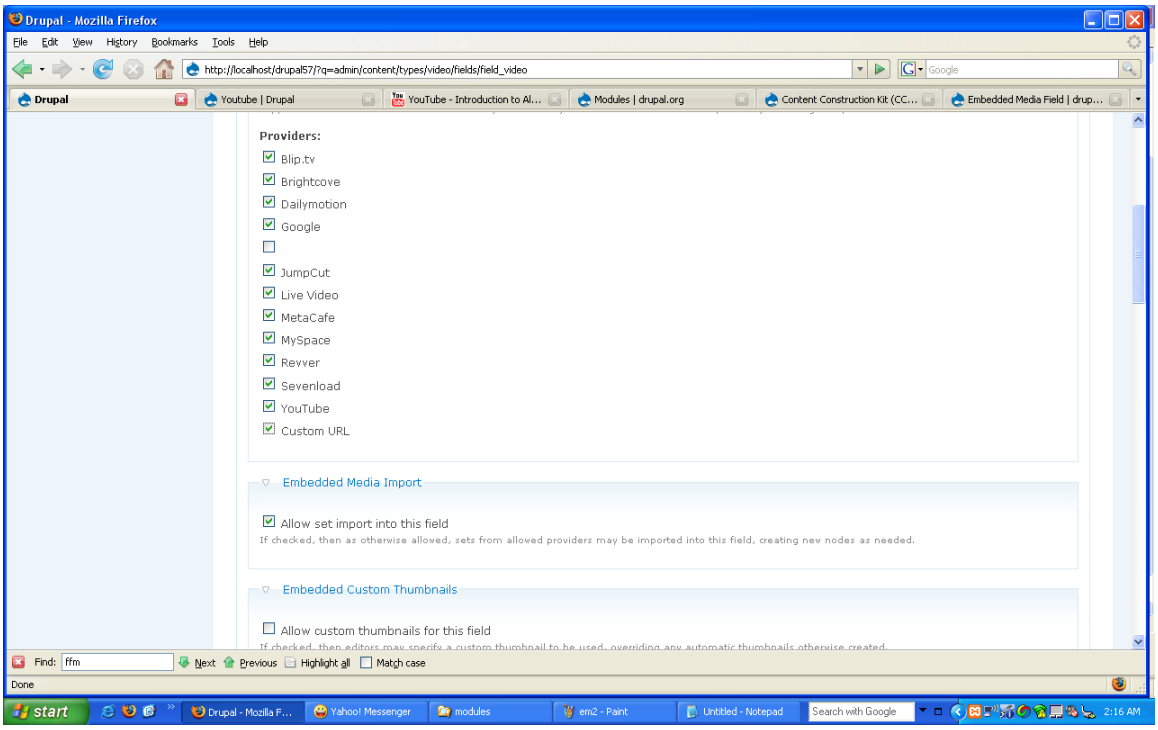


Figure 20: Screenshot of available third party video providers

9. Go through rest of the configuration page and enable the other functions such as "auto play" etc you wanted for your site.

5.2.2 How to make the video size bigger:

1. Go to Administrator/content types and click on video.

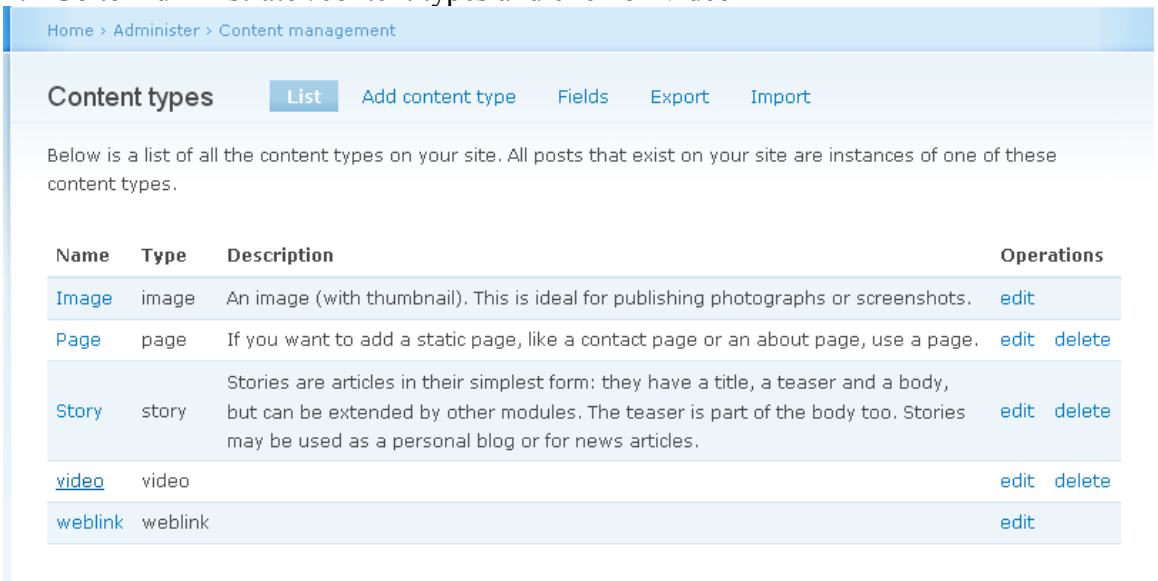


Figure 21: Screenshot of Content types once the user create video field

2. Click on video and then click on “Manage fields”.

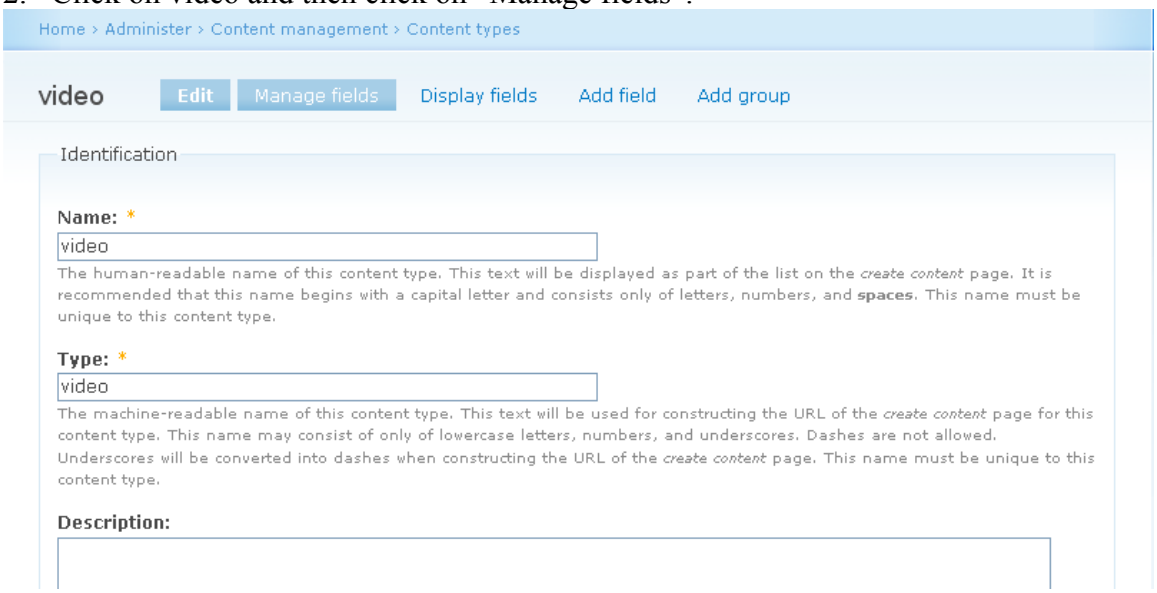


Figure 22: Screenshot of video field form

3. Now clink on “Configure” for video field under manage field.

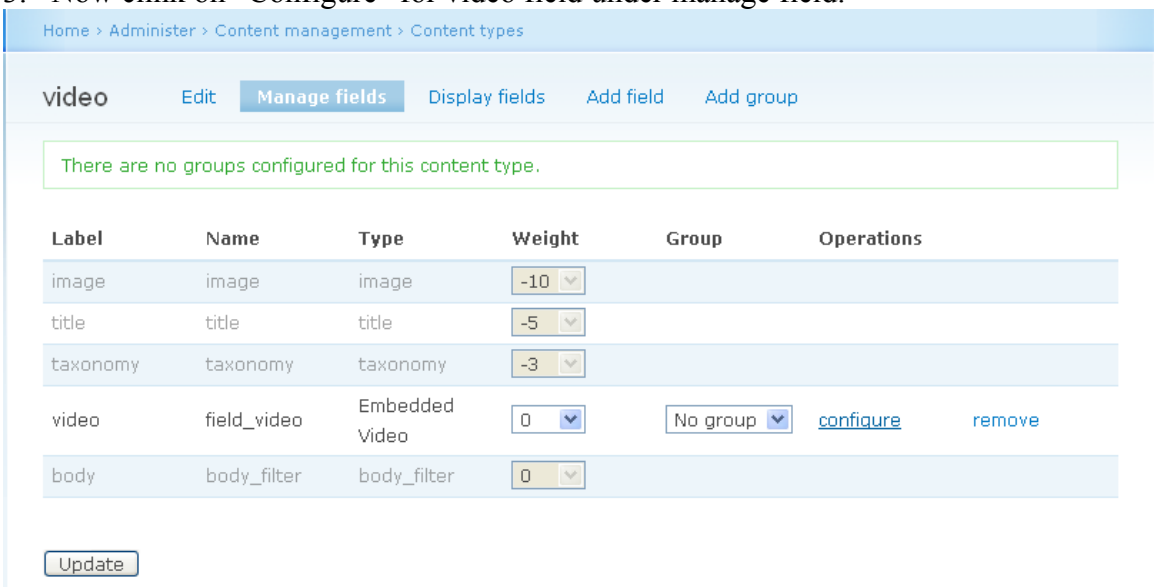


Figure 23: Screenshot of manage field form

4. In the “video display” section give the dimensions you wanted to play the video. The default video size is 425x350.
5. As we can see in the following screen shot, we can change the default video display settings to the new dimensions 800x800.

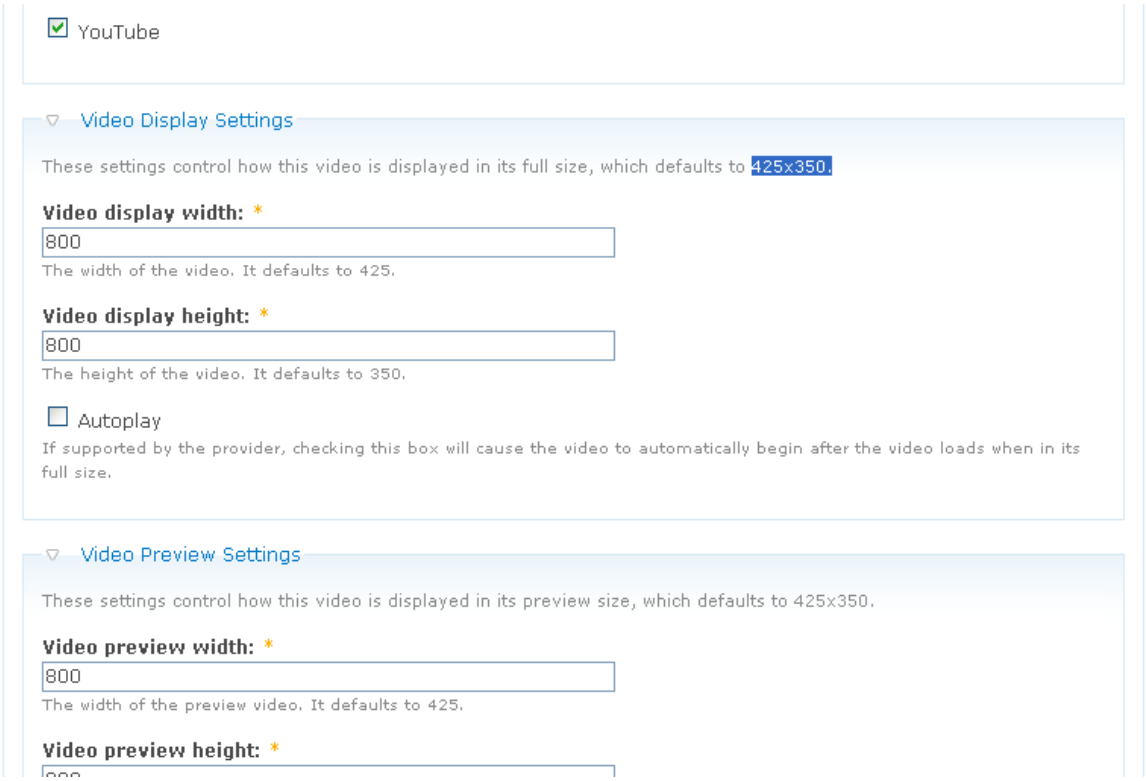


Figure 24: Screenshot of video field settings form

5.3 Taxonomy Module

The Taxonomy module helps classify content and allows vocabularies to be setup as either user defined tags or administrator defined terms. It is a default module in drupal.

5.3.1 Installation and Usage:

1. Go to Administer/site building/modules and enable taxonomy module.
2. Go to Administer/Content management/Categories and click on "Add Vocabulary".

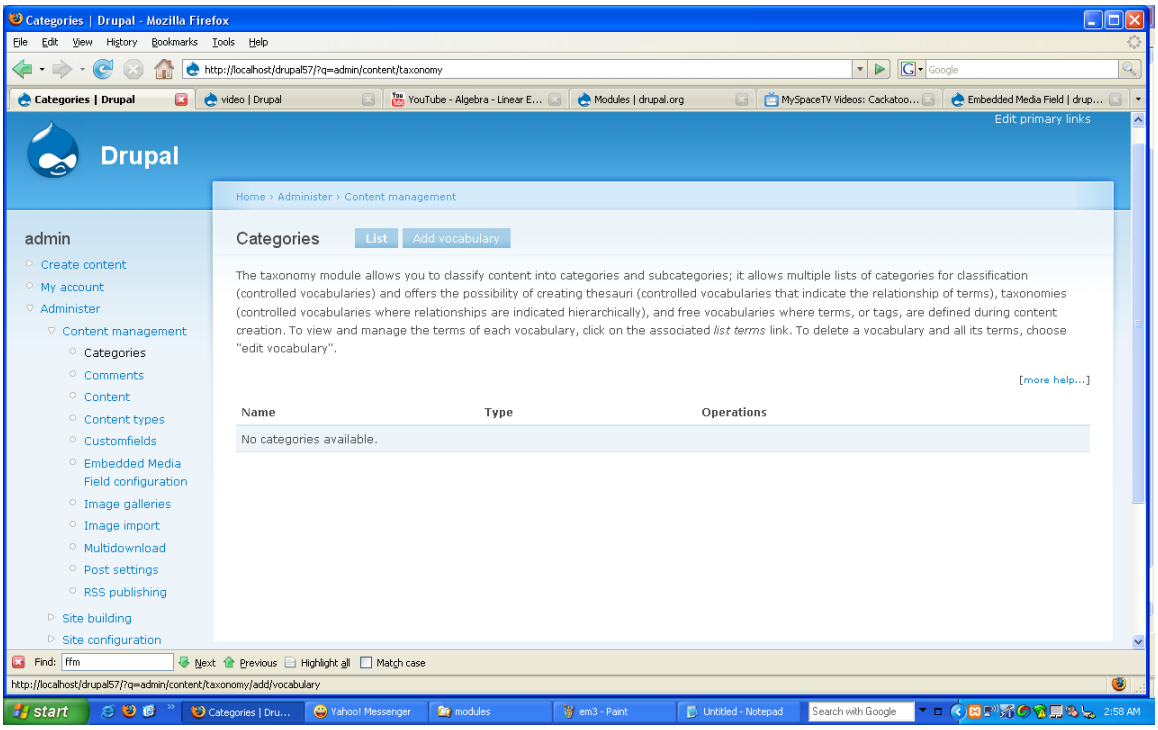
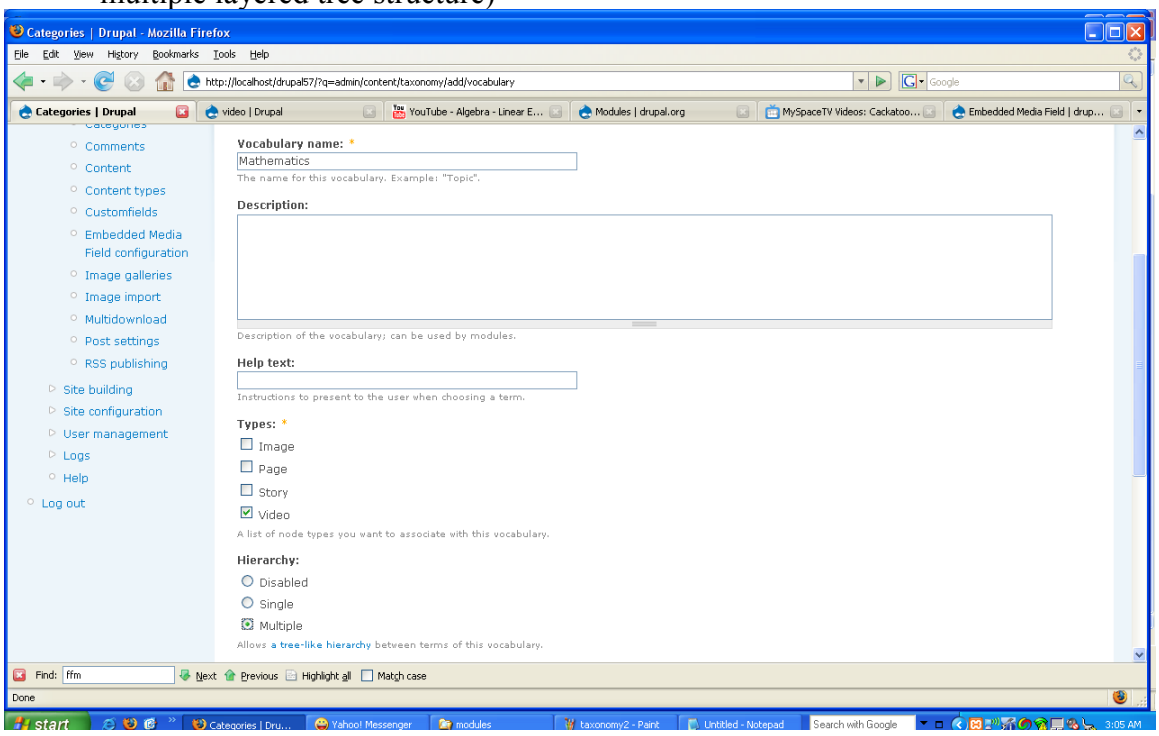


Figure 25: Screenshot of Categories page

3. Now go to Add vocabulary page and enter vocabulary name, description, select type as video, select hierarchy(single for single layered node and multiple for multiple layered tree structure)



Screenshot of "Add vocabulary page".

4. Repeat Step 3 to add more vocabulary.
5. To add second layered terms click on "add terms" for which term you want to add more layers.

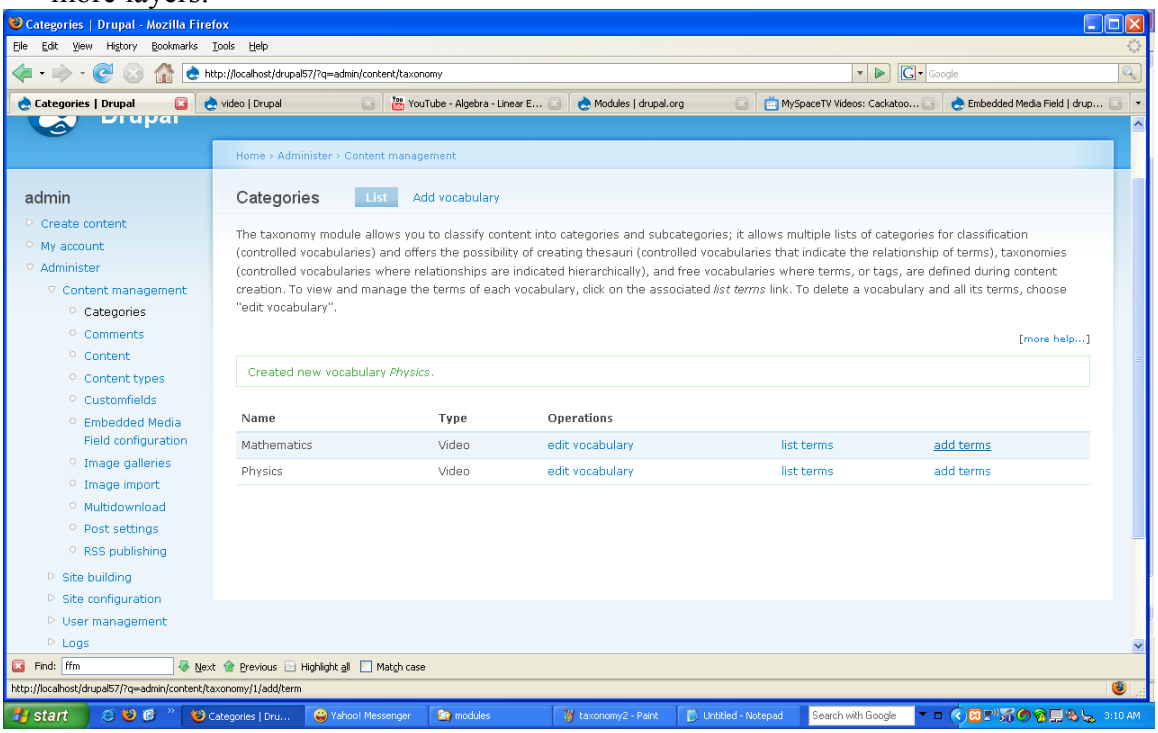


Figure 26: Screenshot of categories listing page

6. Enter term name, description, and the weight (in which order you want the term to be displayed) for that term.

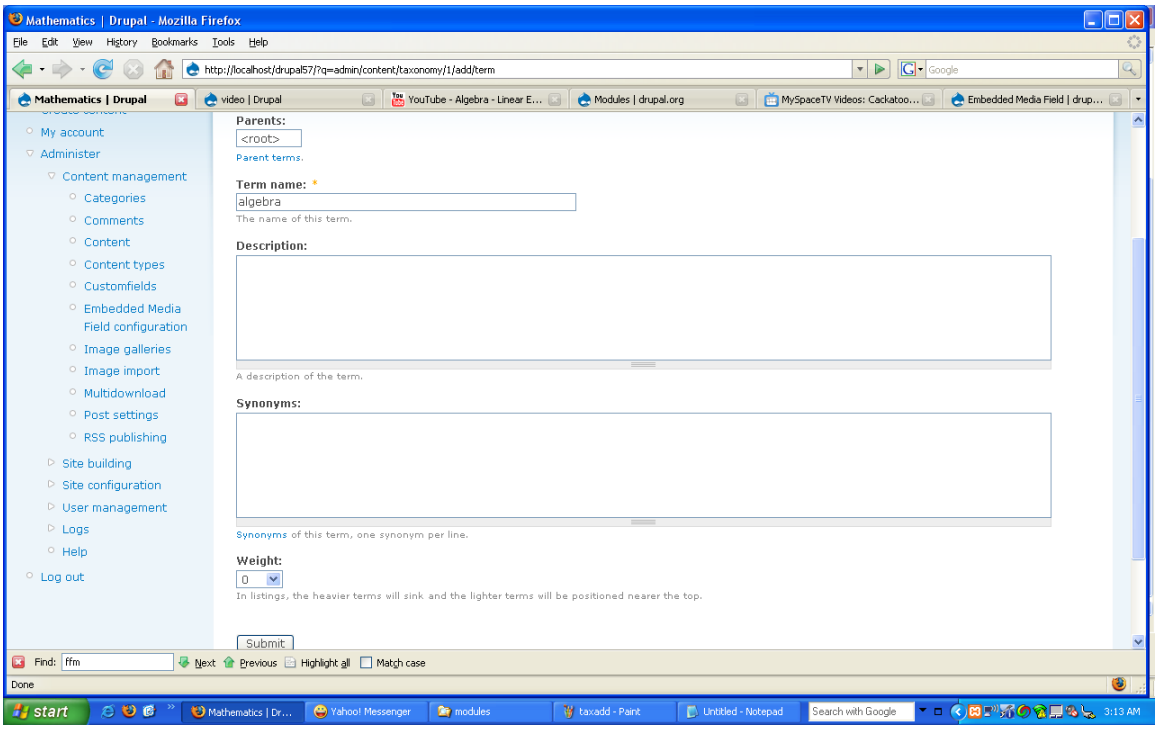


Figure 27: Screenshot of edit form for taxonomy term

5.4 Taxonomy Manager Module

This module provides a powerful interface for managing a taxonomy vocabulary. A vocabulary gets displayed in a dynamic tree view, where parent terms can be expanded to list their nested child terms or can be collapsed.

5.4.1 Installation:

1. Download module Taxonomy Manager (http://drupal.org/project/taxonomy-manager)
2. Unzip it and copy into modules sub-folder.
3. Go to Administrator\Site Building\Modules and enable Taxonomy manager module.

5.4.2 Usage:

1. Go to Administer/Content management/Taxonomy Manager
2. Choose the vocabulary you want to display.

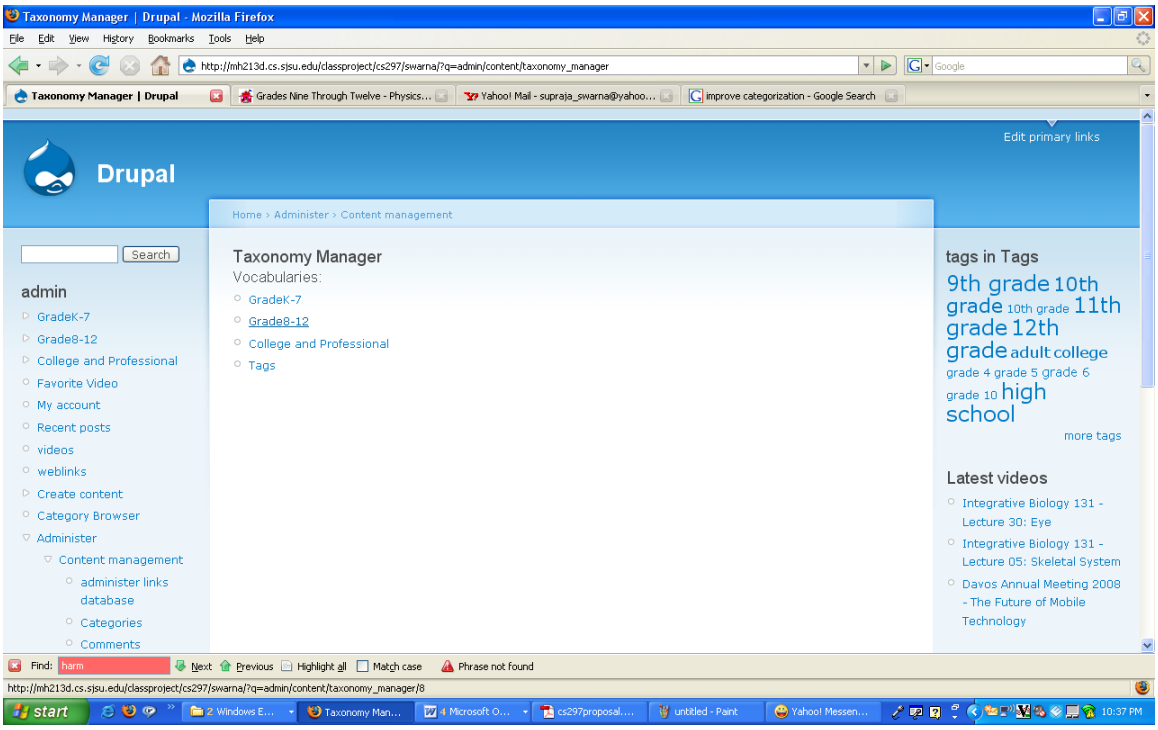


Figure 28: Screenshot of taxonomy manager page

3. Now you can add more terms, delete terms, move terms, or merge terms.
4. We can also add sub categories in each term by choosing that term and clicking on add term.

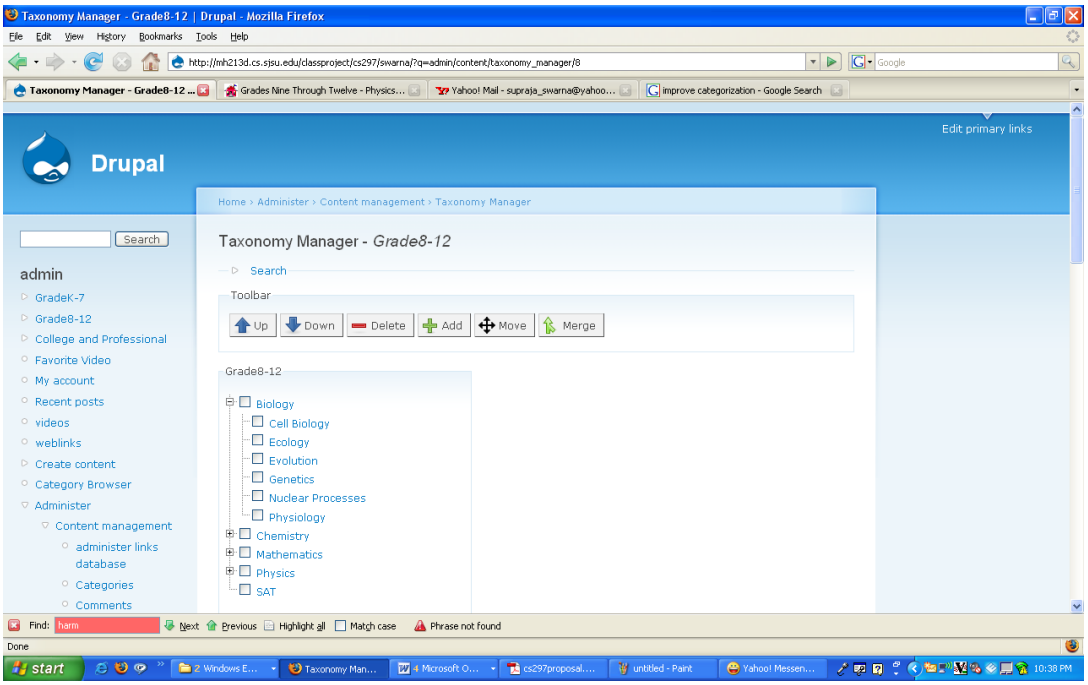


Figure 29: Screenshot for adding multiple taxonomy terms to a single category

5.5 Jrating Module:

This module allows the user to rate specific node or comment.

5.5.1 Installation:

1. Download module votingAPI (http://drupal.org/project/votingapi)

2. Download module Rating (http://drupal.org/project/jrating)

3. Download module actions (http://drupal.org/project/actions)

4. UnZip them and copy into module sub-folder

5. Go to Administrator\Site Building\Modules and enable three of them (jRating requires voitingAPI to be there and enabled and actions required by voting API).

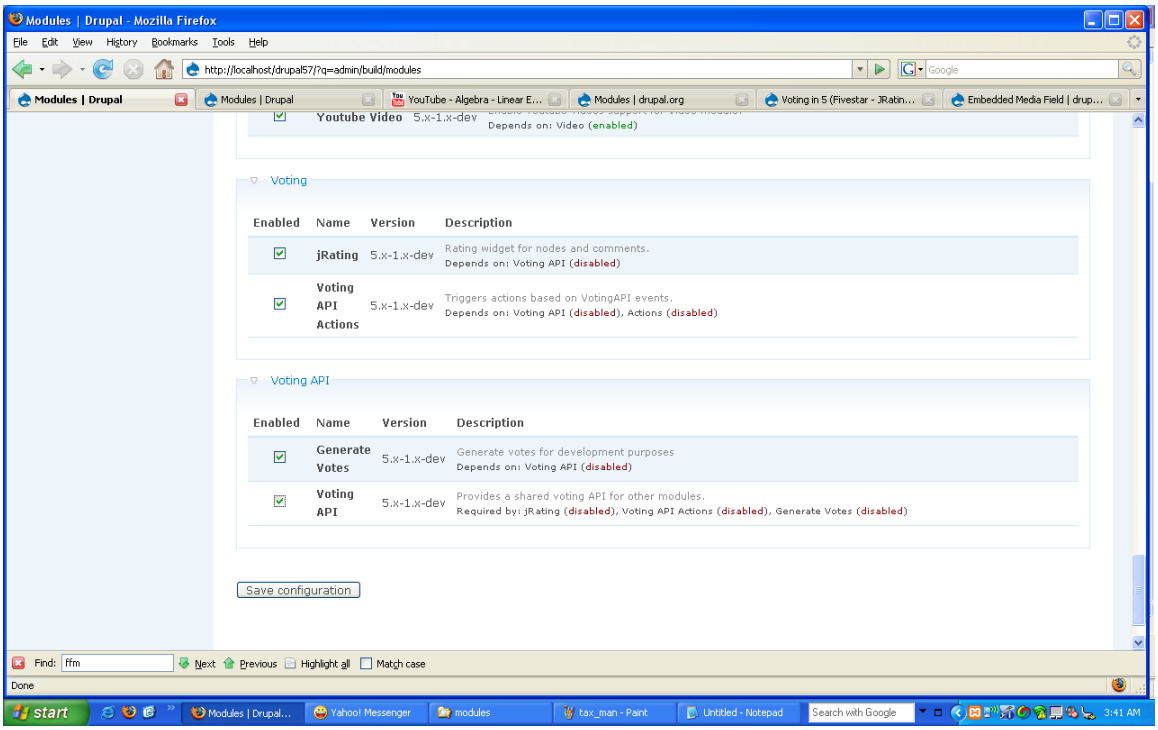


Figure 30: Screenshot of module page to enable Jrating

6. Go to Administrator\Site Configuration\jRating and configure it.

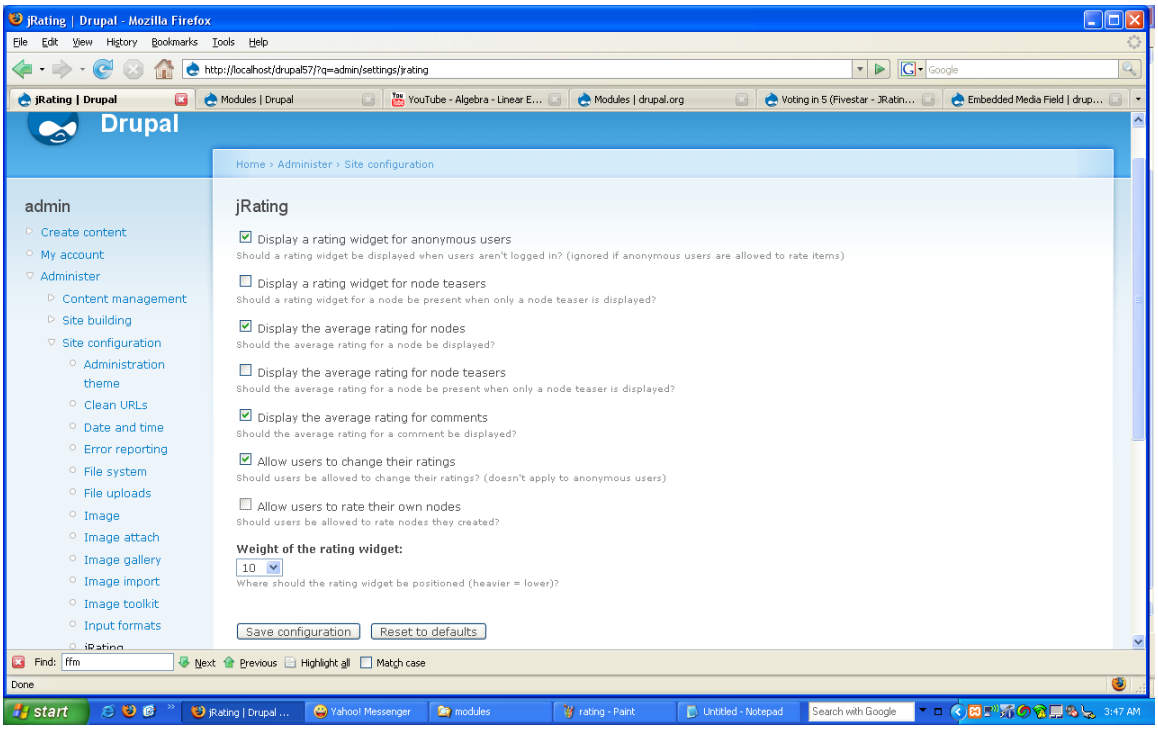


Figure 31: Screenshot of Jrating settings page

7. Go to Administrator\Site Configuration\Voting API to configure tally.

   a. Select Tally results whenever vote is cast.

5.5.2 Usage:

1. Go to Administer\Content Management\Content Types
2. Select type you want, let's say "Video"
3. Enable it as such:

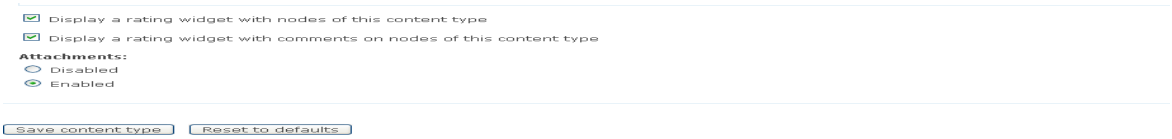


Figure 32: Screenshot of jrating setting s to enable ratings for video

5.6 Tagadelic and Community Tags Module:

Tagadelic module generates a page with weighted tags. The cool thing is that by merely altering font sizes, these lists suddenly gain a dimension.

Community module allows members of the community to tag other users' content, uses the module to show a "tag cloud" for a node based on popularity of tags.

5.6.1 Installation:

1. Download module Tagadelic (http://drupal.org/project/tagadelic)
2. Download module Community Tags (http://drupal.org/project/community-tags)
3. Unzip them and copy to module sub-folder.
4. Go to Administrator\Site Building\Modules and enable tagadelic and Community Tags module.
5. Go to administer/site configuration/community tags and follow instructions as specified.

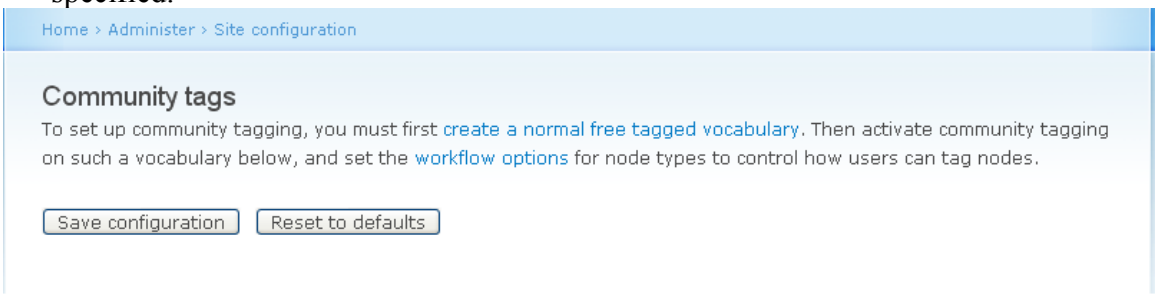


Figure 33: Screenshot of community tags settings page

6. Click on “create a normal free tagged vocabulary”.
7. It takes to administer/Content management/categories.
8. Add a new vocabulary called Tags and choose “free tagging” as below.

Types: *

- [ ] Image
- [ ] Page
- [ ] Story
- [x] Video

A list of node types you want to associate with this vocabulary.

Hierarchy:

- (•) Disabled
- ( ) Single
- ( ) Multiple

Allows a tree-like hierarchy between terms of this vocabulary.

- [ ] Related terms

Allows related terms in this vocabulary.

- [x] Free tagging

Content is categorized by typing terms instead of choosing from a list.

- [ ] Multiple select

Allows nodes to have more than one term from this vocabulary (always true for free tagging).

- [ ] Required

If enabled, every node **must** have at least one term in this vocabulary.

Weight:

0

Figure 34: Screenshot of Tags vocabulary with enabling free tagging

9. Go to Administer/site building/blocks and enable “tags for the current post” block.(This block is because of Tagadelic Module)

Figure 35: Screenshot of showing Tags for current post

5.6.2 How to create Tag Cloud:

1. Go to Administer/site building/blocks and enable “tags in Tags” block, this will create tag cloud.

Figure 36: Screenshot of Tagcloud

2. Enable “Community Tagging form” this will display tags for current post and allow user to tag that post.

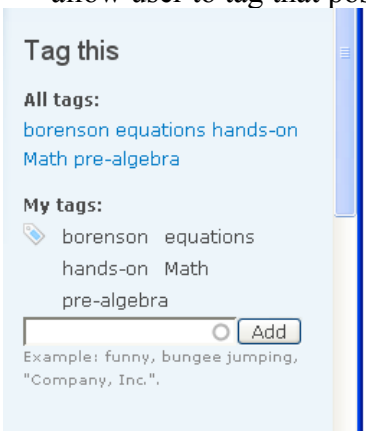


Figure 37: Screenshot of community tagging block

5.7 Taxonomy Menu Module:

Transform any of taxonomy vocabularies into menus easily.

5.7.1 Installation:

1. Download module Taxonomy (http://drupal.org/project/taxonomy_menu)
2. Unzip it and copy to module subfolder.
3. Go to Administer/site building/modules and enable taxonomy menu module.
4. Go to Administer/site configuration/taxonomy menu and select the categories which you want them to show in Navigation Menu.

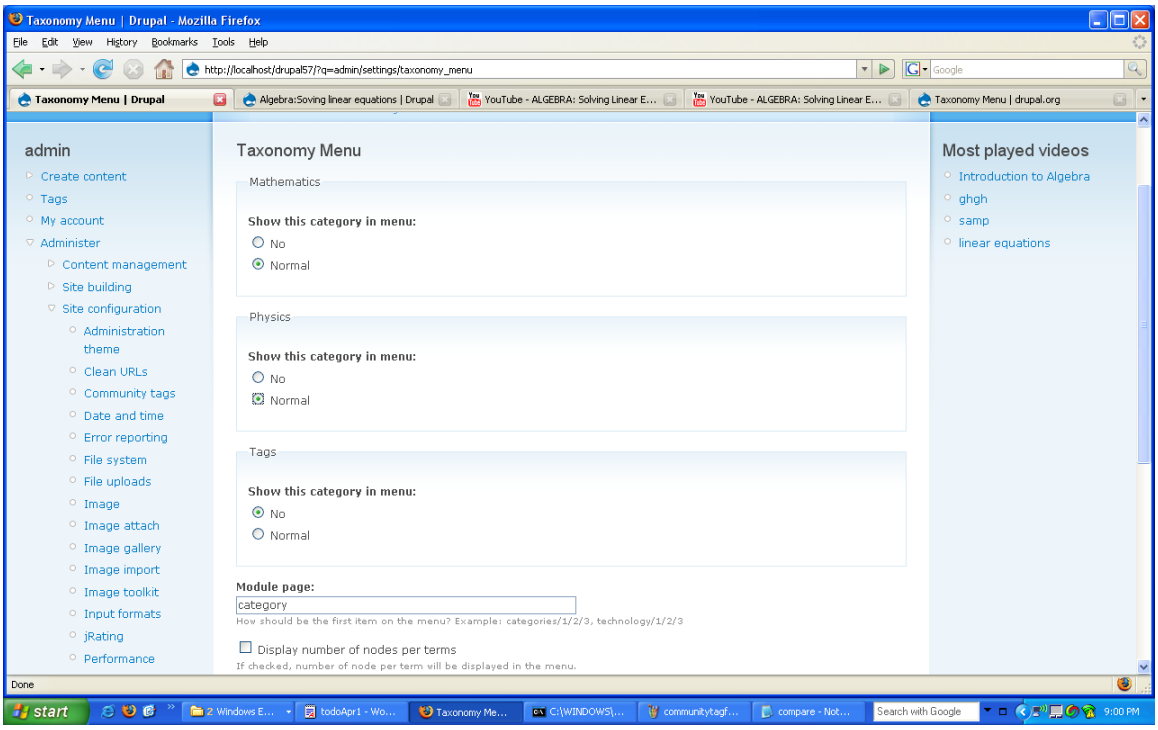


Figure 38: Screenshot of taxonomy menu settings page

5.7.2 Enable Number Of Items in each Category:

Go to Administer/Site Configuration/taxonomy menu and enable display number of items check box.

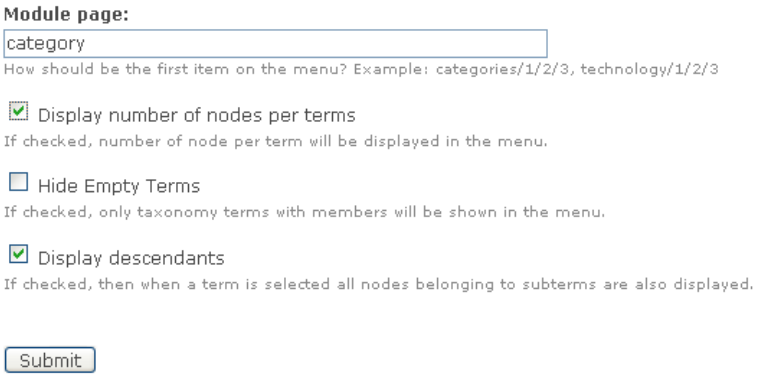


Figure 39: Screenshot of Taxonomy menu setting page to enable number of items per node

5.8 Links Package Module

The Links Package is a set of modules that facilitate use of web links (URLs) in various content types.

5.8.1 Installation:

1. Download module Links Package (http://drupal.org/project/links)
2. Unzip it and copy to module subfolder.
3. Go to Administer/site building/modules and enable Links and weblink modules under Links.

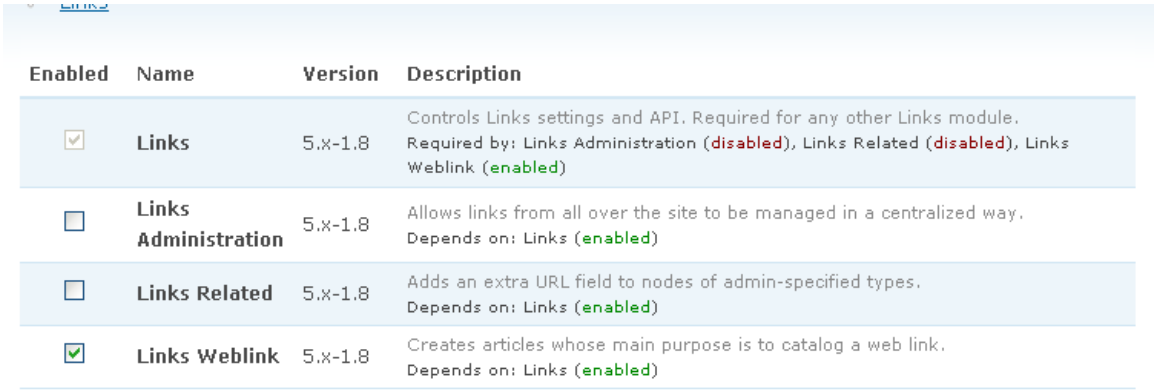


Figure 40: Screenshot of Module gage enabling links module

4. Go to Administer/site configuration/Links package and choose how the links node should display on the system.

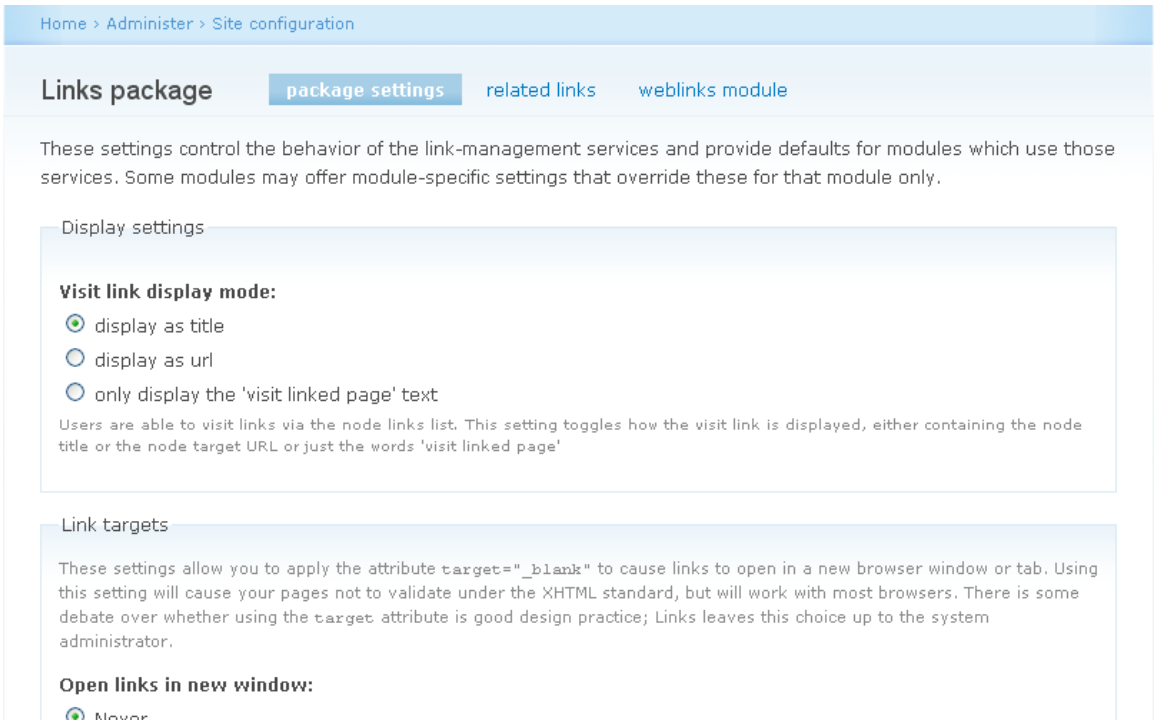


Figure 41: Screenshot of links package module settings page

5. To open the links in new window select “external links only” in links targets section of site configuration of links package.

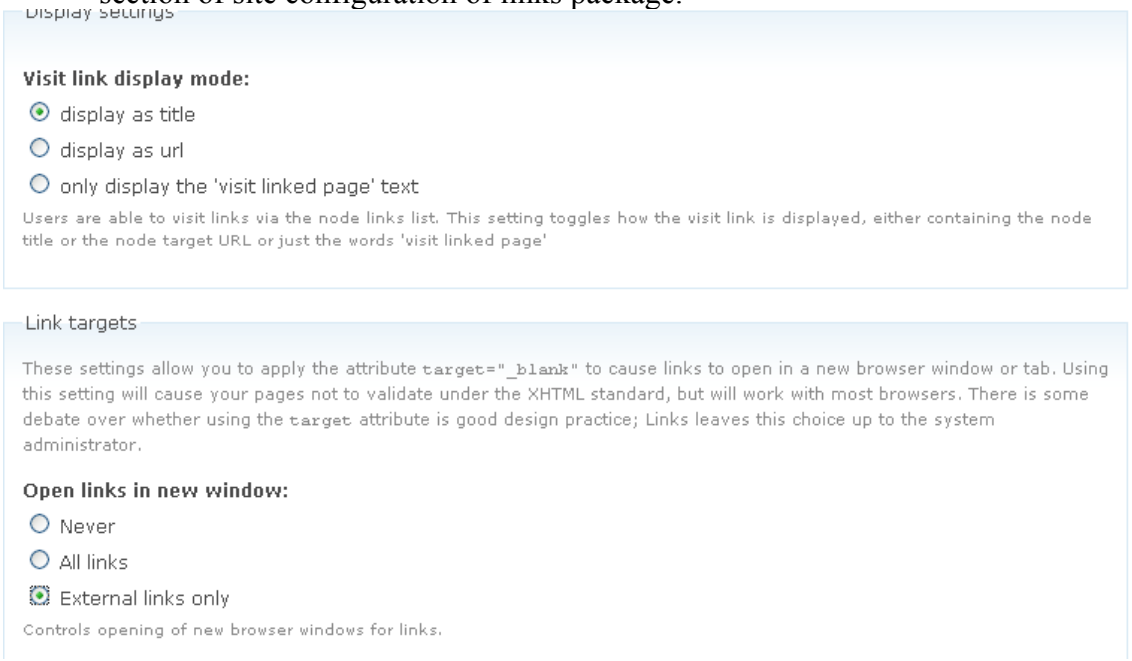


Figure 42: Screenshot of links package settings page to open links in new window.

5.8.2 How to add the links:

1. Once we install links package and enable links and web links modules, “weblink” data type will be automatically generated under create content.
2. Go to create content and click on welink to add the new link.

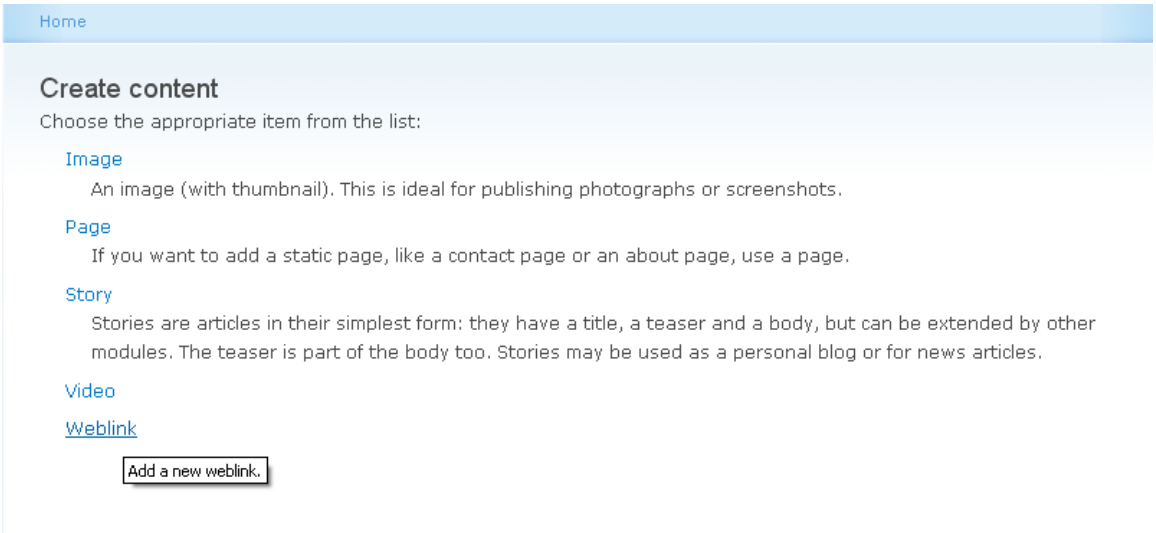


Figure 43: Screenshot of create content page

3. Now “the submit link” form will display.
4. Enter the title of the link you wanted to add in “Title” text box and enter Url of the link in “Link URL” text box.

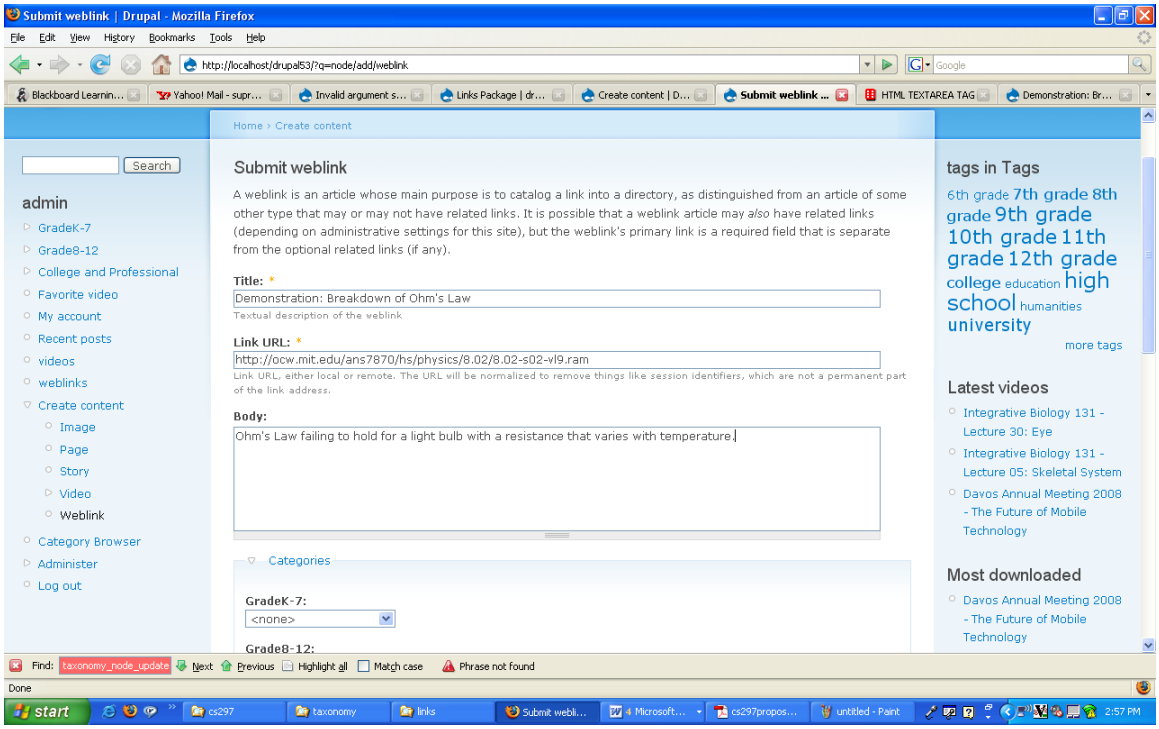


Figure 44: Screenshot of how to create weblink

5. Fill out the rest of form and click submits.
6. Now the link is successfully added into the system.

5.9 Forward Module

Forward module allows the users to email the specific link or page to other user.
This module needs SMTP server to send an email, so it only works if the SMTP server is working properly.

5.9.1 Installation

1. Download module Forward (http://drupal.org/project/forward)
2. Unzip it and copy to module subfolder.
3. Go to Administer/site building/modules and enable forward module.
4. Go to Administer/site configuration/forward and select the options you wanted for forwarding the node.
5. Fill the form and save the configuration.

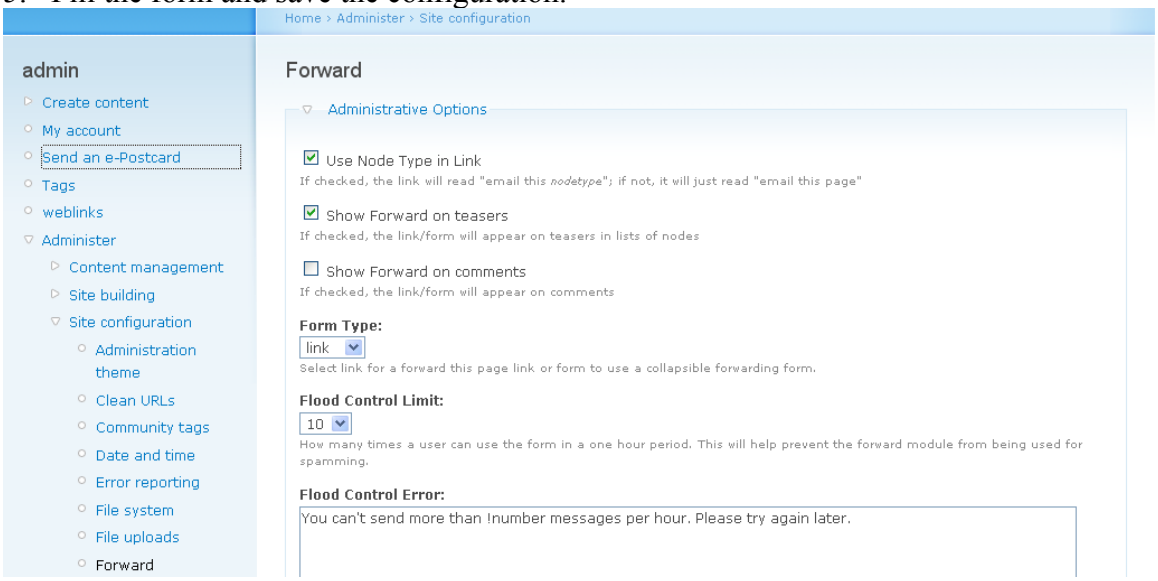


Figure 45: Screenshot of Forward module settings page

5.9.2 How to forward the Specific node

1. Once you install “Forward module”, every node will have a link at the bottom or top as “Email this video” depending on your forward display settings.

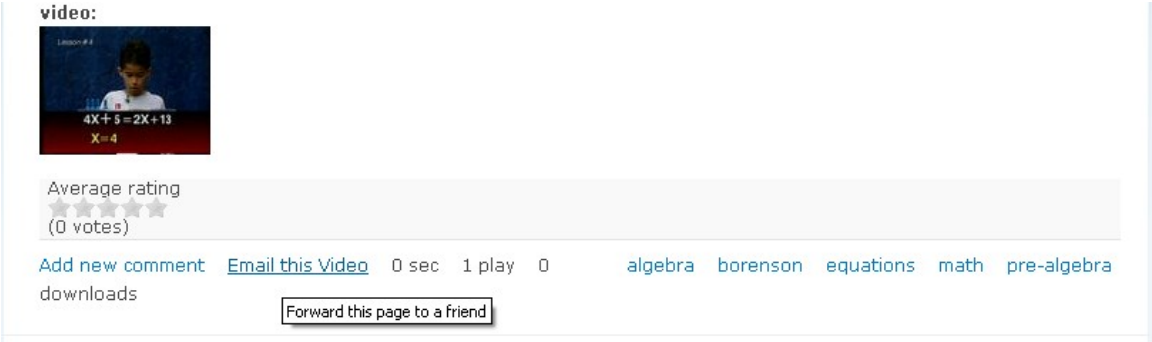


Figure 46: Screenshot of showing forward link on specific node

2. Click on "Email this video", it will open "forward this page" form.

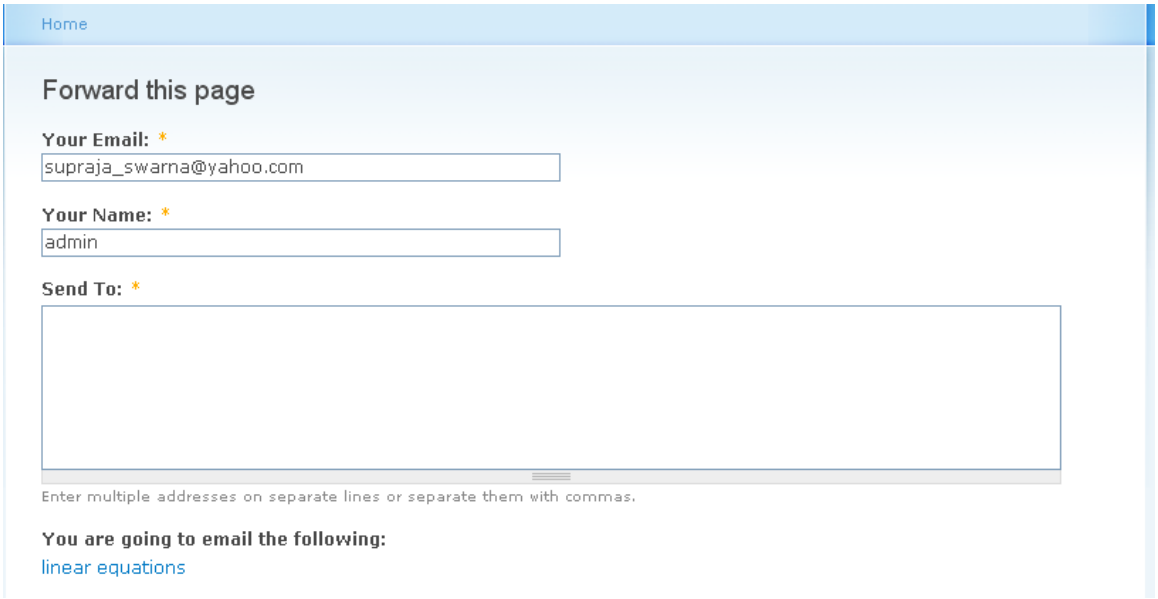


Figure 47: Screenshot of forward page submit form

3. Enter sender email address in "your Email" field, name of the sender in "Name" field, and recipient email address in "send to" field.
4. Fill the rest of the form and submit.
5. If the both entered email addresses are valid and the SMTP server is working properly, then the link will be emailed to the recipient successfully.

5.10 Faceted Search Module:

The Faceted Search module provides a search API and a search interface for allowing users to browse content in such a way that they can rapidly get acquainted with the scope and nature of the content, and never feel lost in the data **[drupal].** It works as a discovery tool as well as search interface.

5.10.1 Installation:

1. Download module Faceted Search (http://drupal.org/project/faceted_search)
2. Unzip it and copy to module subfolder.
3. Go to Administer/site building/modules and enable the following modules: Faceted search, Faceted Search UI, Search (Drupal core search module), At least one of the following modules (which are provided with Faceted search): Author Facet,Content Type Facet, Date Authored Facet, Taxonomy Facets, Field Keyword Filter.
4. Go to the Administer/site configuration/ Faceted Search page, and click on the "Add Environment" tab.

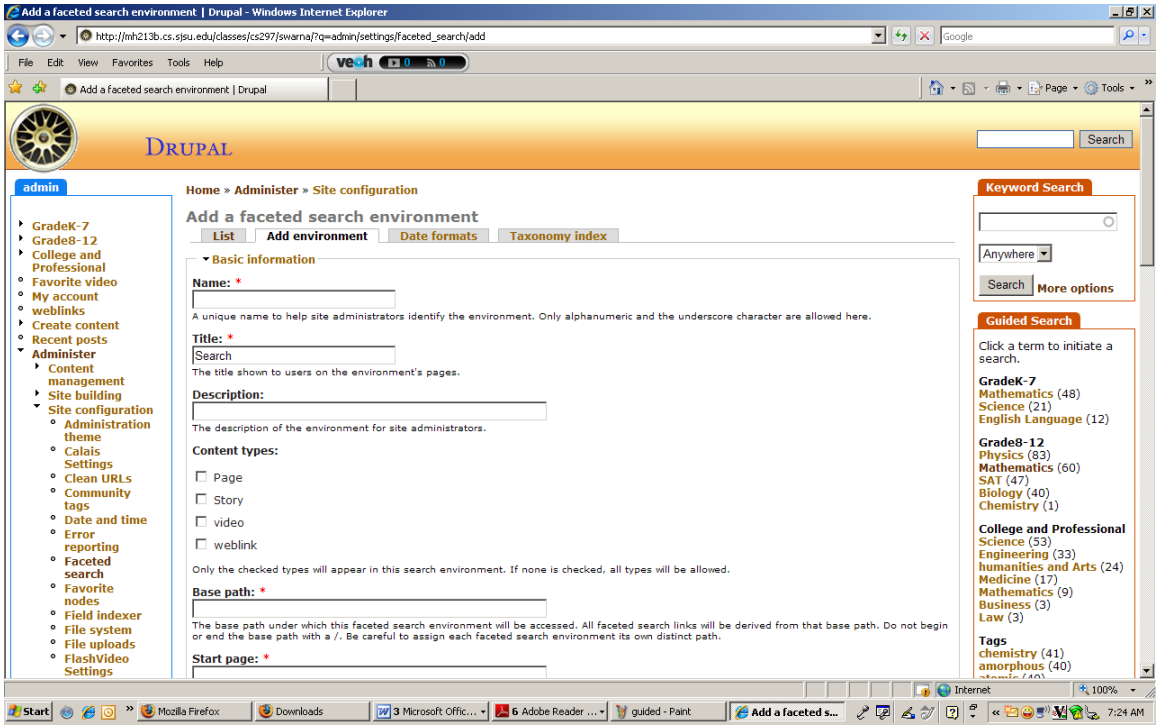


Figure 48: Screenshot of faceted search module settings page

5. Now fill the "Add a faceted search environment" form to define the environment. Now click save to save the new environment. This will takes you back to the Administer/ Site configuration / Faceted Search page.
6. Go to Administer/ site building /blocks page, and enable the following blocks: environment name/ Current search, environment name / Keyword search, and environment name/ Guided search
7. Go to Adminster/ site building/ blocks, and click on configure page for keyword search or guided search block and set visibility of the block. Select "Show on every page except the listed pages", then enter the following paths in the Pages field: search environment base path, base_path/*.
8. Go to Administer / user management / Access control page, and give the permission to the roles you want to give access to "Use faceted Search".

5.10.2 Faceted Search: Key word search auto complete functionality:

The auto complete functionality doesn't come with default package of faceted search module. The user has to insert following code in faceted_search_ui.php file.

Code:

```
/**This has to be add under function faceted_search_ui_menu($may_cache)**/
$items[] = array(
  'path' => $base_path .'/categories',
  'callback' => 'faceted_search_ui_stage_categories',
  'callback arguments' => array($env_id),
  'access' => user_access('use faceted search'),
  'type' => MENU_CALLBACK,
 );


/** autocomplete functionality for search box
* Retrieve a pipe delimited string of autocomplete suggestions for existing users
*/
function tax_autocomplete($search) {
  $matches = array();
  $result = db_query_range("SELECT name FROM {term_data} WHERE
LOWER(name) LIKE LOWER('%s%%')", $search, 0, 10);
  while ($user = db_fetch_object($result)) {
    $matches[$user->name] = check_plain($user->name);
  }
  print drupal_to_js($matches);
  exit();
}

/** This has to be add under function faceted_search_ui_form($search)*/
// Add the keywords to search.
  $form['keywords'] = array(
    '#type' => 'textfield',
    '#title' => '',
    '#default_value' => '',
    '#size' => $current_stage == 'select' ? 35 : 20,
    '#autocomplete_path' => 'faceted_search/autocomplete',
    '#maxlength' => 255,)
```

6. Future work

The search interface could be improved in a number of ways. For example, the keyword search and guided search block could be placed just below the banner of the site so that the user doesn't have to scroll down to see all the categories. One more additional nice-to-have feature would be a checkbox next to each term so that user can always see all the terms in entire time even user chooses a specific term. Currently, once the user clicks on a specific category in guided search block, the system hides all the other catogories and shows only the selected one. So far this project concentrated on proving certain new features (the interface for entering synonyms for categories, allow the user to search those synonyms, faceted search, keyword search with auto complete), and showing that my site can support them very efficiently.

7. References

[1] Hearst, M. 2008. UIs for Faceted Navigation Recent Advances and Remaining Open Problems. *Microsoft Research.*

[2] Koren, J., Zhang, Y., and Liu, X. 2008. Personalized interactive faceted search. In *Proceeding of the 17th international Conference on World Wide Web* (Beijing, China, April 21 - 25, 2008). WWW '08. ACM, New York, NY, 477-486.

[3] Daneil, T. 2006. Dynamic Category Sets: an approach for faceted search. *Paper presented at the ACM SIGIR 2006 Workshop on Faceted Search*.

[4] Kules, B. and Shneiderman, B. 2008. Users can change their web search tactics: Design guidelines for categorized overviews. *Inf. Process. Manage.* 44, 2 (Mar. 2008), 463-484.

[5] Yee, K.-P., Swearingen, K., Li, K., & Hearst, M. (2003). Faceted metadata for image search and browsing. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems, Ft. Lauderdale, FL* (pp.401-408). New York: ACM Press.

[6] Hearst, M., Elliot, A., English, J., Sinha, R., Swearingen, K., & Yee, P. (2002). Finding the flow in web site search. *Communications of the ACM, 45*(9), 42-49.

[7] Garber, S.R. and Grunes, M.B. The art of search: A study of art directors. In *Proc. of CHI-92*, Monterey, CA, 1992.

[8] Jose, J.M., Furner, J., and Harper, D.J. Spatial querying for image retrieval: a user-oriented evaluation. In *Proceedings of ACM SIGIR '98*, pages 232–240, 1998.

[9] Markkula, M and Sormunen, E. End-user searching challenges indexing practices in the digital newspaper photo archive. *Information Retrieval*, 1:259–285, 2000.

[10] http://drupal.org/