

<< Re-Segment Sections In Local Monographs >>

</opt/lampp/htdocs/interface/>

check_sections.php	Find wrongly segmented sections
check_sections_details.php	Re-segment a section, should be accessed from check_sections.php . If to be accessed directly, book id and overlapping threshold (i.e. count) should be given in the url, e.g. check_sections_details.php?book_id=12345&count=5
edit_section_db.php	Included by check_sections.php and check_sections_details.php
insert_new_columns_into_books/	Contains programs parsing book information from other resources and the results

<< Search For Keywords In Local Monographs >>

</opt/lampp/htdocs/search/>

search.php	Search for keywords
search_locust_temple.php	Search for keywords and mark the results with tags if they match the locust-temple-related syntax
search_function.php	Included by search.php and search_locust_temple.php and providing common functions used by these two programs
search.js	Javascript file for search.php
search.css	Css file for search.php
search_results/	All the results from search*.php will be exported to html format and stored in this folder
csv_files/	The results of search*.php in csv format. Will be used by map/get_coordinates_for_listed_books.php

Note:

1. Before users are linked from the search results to the visualization page, **map.php**, they are firstly directed to **map/get_coordinates_for_listed_books.php**, in which the program gets from database the coordinates of the books listed in the search results stored in **search/csv_files/**, writes them in a file under **map/datasets**, and redirect the users to **map.php**. So that users do not have to re-run the search program if the coordinates in the database is changed, instead clicking on the link in the html version of search results will do.
2. Indexing is not done yet, because the default indexing method by MySQL only applies to space-separated languages. The searching still takes a long time, possible solution:
 - a. Additional column storing reversed content
<http://stackoverflow.com/questions/7848445/how-to-use-prefix-wildcards-like-abc-with-match-against>
 - b. Plugin Mysqlcft
The plugin is already installed, only have to execute the command (takes very long):
ALTER IGNORE TABLE Gazetteer.contents ADD FULLTEXT INDEX content_mysqlcft (content) WITH PARSER mysqlcft;
<http://zyan.cc/post/356/>
<https://code.google.com/p/mysqlcft/>

<< Visualize Data From Local Monographs >>

</opt/lampp/htdocs/map/>

map.php	Loads a predefined csv, overlay, and map file and visualize using sebastian's library
map.php?mode=1	map.php with different toolbar layout
map.js	Javascript file for map.php
map.css	Css file for map.css
about.php	Description of the map.php program, included by map.php
get_coordinates_for_listed_books.php	Get csv of the search result then generate another csv to feed into map.php. Read the specified file with a book list stored under search/csv_files/ , get the coordinates for the books, write to the specified file, store it under datasets/ , and redirect to map.php File name must be given: file=filename.csv; Name must be given: name=name , i.e. the url should be get_coordinates_for_listed_books.php?file=filename&name=name
WindowWidget.*	Javascript class enabling windows within the html document
datasets/	Input of map.php
images/	The image files for the program map.php
coordinates/	The programs getting coordinates of each local monograph from CHGIS

Change the list of background maps:

Find the variable "backgroundMapArray" in **map.js** and alter the list in accordance with the format.

Change the list of overlays:

Find the variable "overlayArray" in **map.js**, and insert a new overlay by giving "name" and "layer name". The name is to be displayed to the users. The layer name refers to the "name" as seen in GeoServer in layer preview.

Change the list of dataset (csv format):

Find the variable "datasetArray" in **map.js**, and insert a new dataset by giving "name" and "path". The name is to be displayed to the users. The path refers to the path to the csv file.

Change the colors representing the datasets:

Find the variable "colorArray" in **map.js**, and insert new colors. r1, g1, b1 refer to the color shown when the dataset is active/chosen. r0, g0, b0 refer to the color shown when the dataset is inactive/not chosen.

Locate the "overlay" panel in the map:

Give the parameter mode=1 in the url for **map.php**, i.e. **map.php?mode=1**

<< Get the Coordinates for Each Local Monograph >>

</opt/lampp/htdocs/map/coordinates/>

get_coordinates_from_chgis.php	The the coordinates of each local monograph listed in local_monographs_list.txt from CHGIS and write the results to csv files under csv_files/ . Use the parameter list=176 to change the input to local_monographs_list_176.txt , i.e. get_coordinates_from_chgis.php?List=176
local_monographs_list.txt	List of 1824 local monographs, input of get_coordinates_from_chgis.php
local_monographs_list_176.txt	List of 176 local monographs, input of get_coordinates_from_chgis.php
local_monographs_coordinates.html	List of 1824 local monographs and their coordinate information, output of get_coordinates_from_chgis.php
local_monographs_176_coordinates.html	List of 176 local monographs and their coordinate information, output of get_coordinates_from_chgis.php
1820_1911/	The gis data of china in 1820 and 1911, input of get_coordinates_from_chgis.php
csv_files/	The raw data of the output of get_coordinates_from_chgis.php . Files with book id as the file name contains the candidate coordinates of a book. list.csv contains the list of books and their candidate coordinates

local_monographs_coordinates.php	Insert the coordinates of each book into database
local_monographs_coordinates_filtered.txt	Input of local_monographs_coordinates.php
books_coordinates.csv	Output of local_monographs_coordinates.php , in the format required as input for Sebastian's library (Need to manually copy to map/datasets folder if you want to update the book data set as well)

Create txt input files from excel:

Copy all the needed cells from excel and paste it in text wrangler, the columns should be automatically separated with tabs and rows with line feeds. Save the file in txt. (as

local_monographs_coordinates_filtered.txt)

<< Visualize the Possible Coordinates of Each Local Monograph >>

</opt/lampp/htdocs/map/coordinates/>

map.php	Draw the coordinates of the 1824 books on the map
map.php?list=176	Draw the coordinates of the 176 books on the map

map.js	The javascript file for map.php
map.css	The css file for map.php
map_input_files/	The copy of the files under csv_files/ as the input for map.php

<< Get the Coordinates of Provincial Capitals for Each Local Monograph >>

[/opt/lampp/htdocs/map/coordinates/](#)

provincial_capital_coordinates.php	Insert the provincial capital coordinates into database and list all the local monographs along with their provincial capital coordinates
provincial_capital_coordinates.txt	Input of provincial_capital_coordinates.php
provincial_capital_coordinates.csv	Output of provincial_capital_coordinates.php , in the format required as input for Sebastian's library

<< Insert Additional Book Information Into Database >>

[/opt/lampp/htdocs/interface/insert_new_columns_into_books/](#)

insert_new_columns_into_books.php	Parse the book information from localmonographs.xml and insert the result into database
insert_new_columns_into_books.css	The css file for insert_new_columns_into_books.php
insert_176_rows_into_books.php	Insert and update the information for 176 books in the database
localmonographs.xml	Book information to be inserted
localmonographs.txt	The txt version of book information to be inserted
local_monographs_176.txt	The list of 176 books and their information to be inserted

Note: the above was to supplement edition info in the books_info table.

<< Get New Books From Sinica >>

[/opt/lampp/htdocs/interface/insert_new_columns_into_books/](#)

get_data_from_sinica.php	Parse the book information from the website of sinica and write to files stored under data_from_sinica/
parse_data_from_sinica.php	Group the duplicated books of source 1 and write the results to data_from_sinica/merged_books.csv
analyze_data_from_sinica.php	Concatenate all the csv from 01-71.csv to all_data.csv Count the # of books of each source (Each action is written in a function) (Choose one of the above at will. Comment another before you run.)
data_from_sinica/	Csv files storing book information, encoding in utf8 and big5 (originally encoded format) column_name.csv contains the mapping between column name and source all_data.csv contains all the data concatenated from 01-71.csv merged_books.csv contains the grouped list of duplicated books list_of_local_monographs_from_sinica.xlsx contains the grouping

of the duplicated books which are assumed to be the same one, the
excel version of merged_books.csv

Note:

These books are not yet stored in the database

<< Sebastian's Library >>

[/opt/lampp/htdocs/geotemco/](#)