**Definitive SS Concentration Camps Database Metadata (7/16/2012)**

**By Benjamin Blackshear and Anne Knowles, Middlebury College, Geography Department**

Genealogy of the Database

This database began as a GIS to record the location of SS main camps and subcamps for purposes of mapping them in the USHMM encyclopedia of camps and ghettos volume I. It was built by Jude Richter, who had taught himself GIS for this purpose. Middlebury College inherited Jude’s database in 2008 when Anne Knowles and Middlebury student Alexander Yule began to standardize and reformat the information, and define the database structure more clearly to maximize its potential for GIS mapping and analysis. Alex’s database was then augmented with information about labor and firms by Middlebury students Roz Vara and Charlie Hoffman in 2009-2010. Paul Jaskot (DePaul University) and Knowles reviewed the data for select camp systems and found further problems of inconsistency. In a final stage of trying to make the database as consistent as possible, Ben Blackshear and Knowles set out to create a definitive SS camps database by:

1. Comparing the various databases, which included:

* Camps Local MXD file using “Camps” table (Alex Yule)
* Merged\_subcamps (by Roz Vara and Charlie Hoffman) (Blackshear combined their datasets which had previously been two separate files)
* Updated\_subcamps (by Ben Blackshear) = Merged\_subcamps minus all camps without opening date or without closing date or with neither date; Ben assigned opening and/or closing dates where a partial date was given, namely last day of month where month only was given, last day of year where only year was given.

2. Defining the necessary fields for the camps project already underway and strategizing ways to combine the relevant aspects of each database into a final definitive version.

3. Printing out a hard copy of “Merged” and “Camps Local” and comparing their entries line by line so that the final database would be as complete as possible. During this process, a variety of other issues were encountered, such as camps with little or no attribute information in the database, camps that did not have encyclopedia entries, camps that switch main camp administration, and camps that had 1 entry in the database but were actually 2 or more separate spatial entities. We realized at this point that subcamps in the two databases had to have the same names so that the databases could be joined in ArcGIS. We created new versions of “Camps Local” (“Yule\_Definitive”) and “Merged” (“Merged\_Definitive”) to edit in preparation for the join.

We also began adding rules for standardizing data entry. See “Rules for Parsing USHMM Camps Database” and “Metadata notes for settling on definitive Camps database.”

4. A number of camps were clearly improperly located [such as a cluster of camps with the coordinates (0,0)] and another large cluster all sharing (51.7, 9.8667) [Location of the Neuengamme main camp]. Ben fixed these errors using the encyclopedia, Google Maps and the application’s latitude/longitude popup, and getlatlon.com.

5. After the join was executed, a new shapefile was created, “Pre\_definitive,” which was proofed against the Encyclopedia. We added an Encyclopedia page reference for every entry that did not have one, and checked dates wherever the date field was <Null>.

6. Throughout this process, we checked questionable entries against the Encyclopedia. This process led to further rules for creating or deleting subcamp entries (see the aforementioned documents containing these rules). We also filled in the LABOR field for each subcamp where possible.

7. We compared the spatial data from the 2 databases for each subcamp. In general, the more detailed coordinates were considered authoritative if they had a populated “HOW\_FOUND” field. We created a new field “share\_loc” which is populated with a 1 if the subcamp shares both coordinates with another subcamp (this occurs a number of times, mostly due to difficulties locating subcamps in a more detailed manner that the center of a given town, but also because there are separate entries for different instances of the same camp). New authoritative fields were created to house the final coordinates, “LAT” and “LONG.”

Even though coordinates of each subcamp were updated and now complete, the points still appeared in the same old locations on the map, so we exported the attribute table of the shapefile and created a new shapefile by bringing the table into ArcGIS, right clicking it to display the XY data of the table using the LAT and LONG fields, and then exporting the XY data as a shapefile called “Camps\_Definitive.”

8. We deleted or renamed a number of fields that were redundant from the join or artifacts of Alex’s standardizing efforts. See “Proposed fields for definitive database” for our decisions

9. We then spatially proofed the definitive database against the maps in the Encyclopedia by creating replicas of their extents for each camp system and comparing printed copies side by side. Some changes were made and when they were, both the LAT and LONG fields and the actual locations were updated so that replotting the entire shapefile from the table another time was unnecessary.

10. Because this will be the template for building the entire camps database for future volumes of the encyclopedia, we created a copy of “Camps\_Definitive,” naming it “SS\_Camps\_Definitive.”