

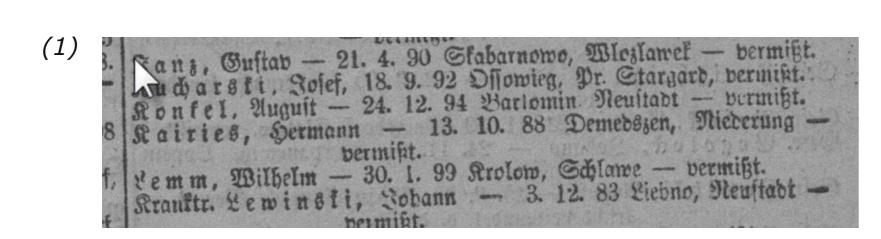
VERLUSTLISTEN.DE — INDEXING THE GERMAN CASUALTY LISTS WITH CROWDSOURCING

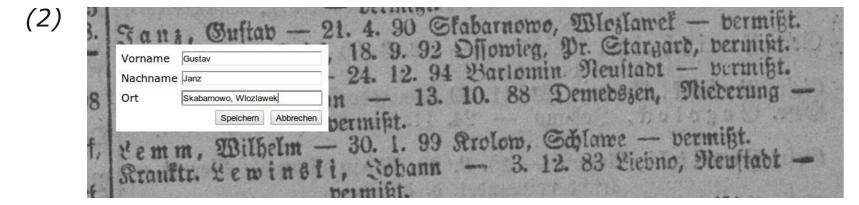
Web-based solution

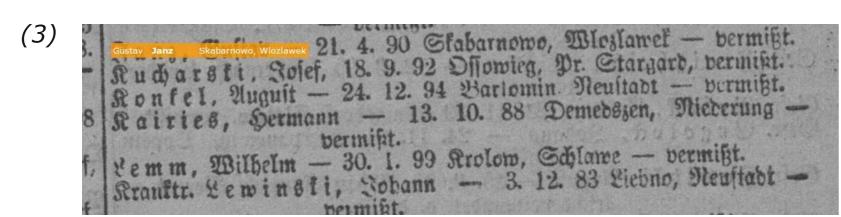
For volunteers working offline a complex management would be necessary to track who is working on which page. Transporting the huge amount of data (93GB in JPEG format) to the users by DVD is not practical. A web-based solution makes it easy to replace faulty scans. Entries are sent to the server immediatly after the contributor has entered them. Assistence can be given quickly by the project supervisors if a volunteer does not follow the editorial guidelines. In a traditional offline data entry such systematic errors are only recognized after hundreds or even thousands of entries have been transcribed wrong.

Work "on-the-scan"

As the entries can be seen directly on the scan, it can easily be verified that each entry of the source has been transcribed. For each entry, its position on the page is recorded automatically. This makes it possible to add further information by editing entire regions (rectangles) of the page. During the later search a user can be led directly to the corresponding position on the page—an enormous help for the large, three-column pages.



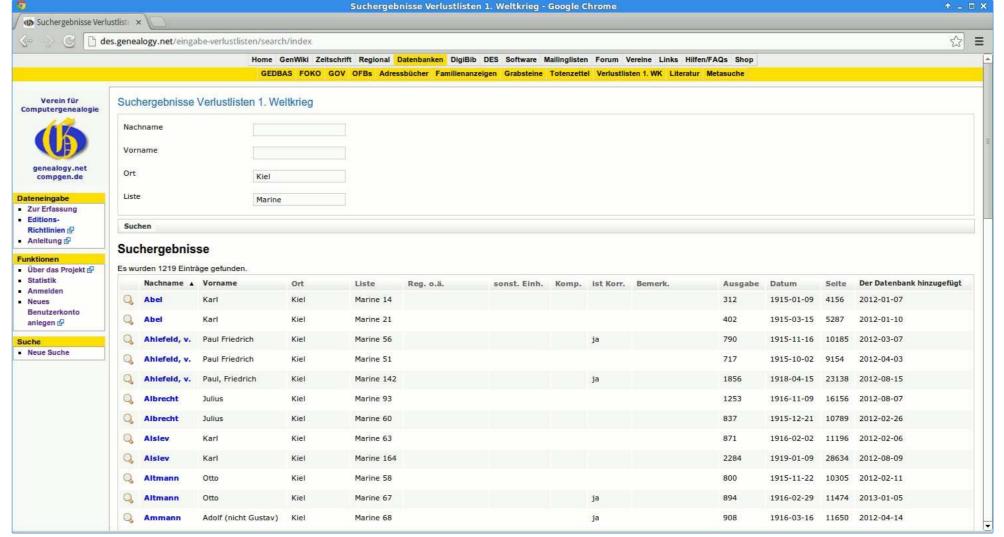




Working on-the-scan: (1) The user clicks on the position of the entry. (2) An input mask (HTML element) opens at that location as an overlay. (3) After the user has entered the data and clicks the save button the data will be sent to the server. The input mask disappears. Instead, the entered data is displayed in a color shaded rectangle at the position of the initial mouse click.

Search during entry

Voluntary transcribers want to see "their" data available online quickly. Visitors that have been attracted via the search might become actively contributing volunteers. In the display of results, a possibility is offered to mark a typo. These typo reports are submitted to the project supervisors for reconsideration as a permanent quality control.



Search results showing marines from Kiel.

Acquisition of "area data"

In parallel with the transcription by the volunteers, the so-called acquisition of "area data" is performed. In this process, information is assigned to larger rectangular regions of a scanned page. Since in the casualty list entries of a regiment and other units are printed en bloc, one can conveniently draw a frame around these entries. With this one working step one can assign the name of the unity to all entries within this block.



Entering "area data": some already defined areas and the input mask for the data of one new rectangle.