EM Places: Search and Details review

Work done so far on EM Places GUIs focused on the detailed view of a single place record (https://github.com/culturesofknowledge/emplaces/blob/master/design/display.md) including all the data that are potentially available for an entry. While there is room for improvement of this interface, more efforts should now be dedicated to search functionalities, and specifically to the optimization of the display of search results.

Discussion with scholars and with developers helped us identify some suggestions, requirements, and constraints that can be distinguish broadly by belonging to contents or to interaction and functionalities.

Content

Search criteria

Search in EM Places, at least for v. 1, will be completely **place-centric**, i.e all the results shown will be places, either current or historical. The basic search criteria then will be related to the **core metadata** that characterize a place (https://github.com/culturesofknowledge/emplaces/blob/master/design/concepts.md). However, what makes EM Places unique are calendars and hierarchies, so specific ways of querying and displaying these two classes of metadata should be designed for future implementation.

Calendar search can take advantage of the calendar conversion APIs (https://github.com/culturesofknowledge/emdates/tree/master/dev) created for EM Dates and of the guidelines for calendar attribution mentioned in https://github.com/culturesofknowledge/emplaces/blob/master/design/display.md. Also related to this is deciding to what dates apply in search.

Hierarchy search should include the possibility of querying through one of the 4 types of hierarchies (of which 2 are currently implemented in the data model: political and ecclesiastical). One of the most relevant challenges is to design ways of querying and displaying hierarchies on which scholars gave contradictory/conflicting opinions. More in general, even before search, the display and navigation in hierarchies requires specific work.

In addition to this, scholars expressed their wish to have:

- Th possibility to have a flag in search about **uncertain information** if marked as so in the data model..

Search Results

Currently the results of a search are presented as a simple list having the following **problems**:

- Results are ordered by the time of their ingestion into a particular collection in Timbuctoo.
- The label of the results is only the place name. No additional information (e.g. country, province, etc.) to help disambiguate between same toponyms corresponding to different places is provided.
- Other additional details that would help further refinement of search (such as displaying the place short description in the results) are not available.

Consequently, the main and basic improvements to the display and evaluation of search results will consist in proposing **additional information** to be displayed for each result.

In addition to this, scholars expressed their wish to have:

- A blurb showing **general statistics** about a search (such as the number of results retrieved in a particular query on the total number of the items in the collection; a map of their spatial distribution; etc.).

Display

Wrt the current display interface, scholars suggest to include:

- **Feature type of a place** (isn't the current GeoNames hierarchy enough to start with?)
- **Contextual information**: other things that happened in the same time at the same place (could it be presented in a similar way as “related places”?).

Interaction and functionalities

As described in https://github.com/culturesofknowledge/emplaces/blob/master/design/search.md two main methods exist in similar projects concerning search. The first is about displaying all the results including “related” places in a long list; the second is about displaying only the matching strings (i.e. place names) and keep any detail for the dedicated place page. A combination of the two methods would probably represent the best solution for our case because it will allow users to pick the right results thanks to additional information displayed in the result list, while at the same time being able to subset the potentially long result list using filters.

Basic improvements include:

- **Sorting options** (e.g. alphabetical, by date of last edit, etc.).
- **Pagination** and/or other incremental ways of navigating through the list of results.
- **Filters** and/or **facets** to allow sub-querying of the results
- **Export** the result list.

Timbuctoo’s current search capabilities (front-end side) are very limited. Widgets that can be used in the GUI are:

- **Free text** search (text box). This is completely based on Elasticsearch (ES). Therefore it supports the writing of queries based on ES grammar for e.g. regex or wildcards. This search supports also the ISO 8601 date format. No autocompletion is available (but see: ES “Completion Suggester”). Some scholars we’ve talked with think that Google Maps-like suggestions would be very useful.
- **Collection based** search (a button for every collection type). This takes the elements in the graph marked by a certain rdf:type and converts them into buttons for filtering. Apparently not configurable to restrict the list of elements to display.
- **Facets** (checkboxes). Facets can be be configured based on any textual attribute contained in the graph of one or more collections.
- **Number ranges** (slider). In case there of fields with datatypes integers or dates, a slider can be used to select ranges.
- Currently there is **no map** display. The possibility to draw buffers/polygons around a place for the same data/timespan to do a search has been mentioned as a nice-to-have feature by different scholars.