

## DISCERN Deliverable 4: User requirement list

The following list of user requirements of the CLARIAH Media Suite is based on the following student assignments:

- Tool criticism templates (the students' own explorations with the Media Suite)
- User study reports (the students' analysis of peer use of the Media Suite, based on their own developed exploratory search tasks)

<b>Topic</b>	<b>User requirement reflection</b>
Source	There is unfortunately no information on how sources are selected, and on what basis sources are included or left out of the database. There is no information on (if any) pre-processing steps undertaken in order to digitize the sources (Ref: AB)
Unclear functionality	<p>Interestingly, although the website enables searches within supported collections, the website does not offer a search function for the contents of the website itself. This makes it more difficult to find information on the general infrastructure of the website, since you're dependent on how the website is structured by others to find information. (Note: I found out that there is a search function within 'help'). (Ref: AB)</p> <p>Server issues when working with big groups (JB)</p>
Dead links	Some 'dead links', for example if you go to Data>Explore>'More information about the dataset' (JB)
International audience	<p>Most resources are in Dutch, but the infrastructure itself is aimed at a more international audience (the navigation and tool are in English) (JB)</p> <p>Exclusively makes use of Dutch archives, although the website's main language is in English. It feels like this last point feigns usability for non-Dutch speaker. (TH)</p>
Facets	There seem to be a lot of facets, with 1 or 2 search results. Seeing as this tool is meant for researchers and academics, this might be useful to scroll through to get an idea of the data, however, it can also become too much

	<p>to process. (e.g. the keyword facet). Big differences in options for different archives. (JB)</p> <p>User can look at one collection in general. Then they can make more specific queries to extract more specific information. There are some layers of search that are provided by the platform that can make a query very specific (e.g., we can search for some words only in the titles if we do not want to search in the entire dataset. Although, more specific search tasks by applying more facets in the search task sometimes are quite blurry. Most of the time no results are generated because the user does not know how to handle facets. Moreover, poor explanation is given about the facets that are offered from the page.</p>
Suite infrastructure	<p>Each archive contains these ways of specifying queries, but they all might have different metadata fields, different types of date field or different facets.</p> <ul style="list-style-type: none"> <li>- The use of these kinds of filters is as a result quite inconsistent between archives.</li> <li>- In addition, I feel like that these filtering options are often ill explained.</li> <li>- Lastly, especially a lot of different ‘facets’ seem incomplete, have inconsistencies, duplicates or sometimes barely readable.</li> </ul> <p>(TH)</p> <p>An element that could support a new user to navigate and manage the platform, at least during the first tasks, would be a search box at the top right side of the home page, instead of the ‘Help’ feature. (MS)</p> <p>Across the archives, and in some cases within the same archives, there is not a lot of consistency in terms of data and metadata. It is very much to the researcher to find this out. (JB)</p>
Search function	<p>The ‘Search’ is a very useful and with a clear goal tool dedicated to searching and exploring. When a user starts the exploration using a simple query, the process is simple and direct. However, for</p>

users who are not familiarized with filtering using the so-called 'facets' or systematic search, the process can be quite complex and overwhelming, especially for the beginners since the plethora of options make it hard to identify and select the most suitable. (MS)

"Sort by date" is only available when having selected a 'date field', maybe useful to provide it next to the sort by relevance button but refer users to the date selection option when they click on it. (JB)

The most important ways to sort the data is by date and by relevance. It is in the media suite unclear what sort by relevance does. If it even does anything at all! (TH)

Sentences like "It works as an aggregator of the metadata, but also enables access to the media content itself. These collections are registered in one instance of a web-based open source management system called CKAN." are somewhat cryptic and might exclude researchers who are less familiar with the lingo or junior researchers. (RT)

Some search results return tags/keywords on the overview of search results, and others do not. In order to foster a serendipitous search, it would be interesting if tags are returned in the search results. It would also be helpful if tags would be clickable to return new search results instead of writing down interesting tags/keywords and returning to the search field, especially since users can save queries. • Looking at this result: <https://mediasuite.clariah.nl/tool/resource-viewer?id=2101608110098121131&cid=daan-catalogue-aggr&st=gentrifica.e>, there are three versions of the evening news bulletin, but users have to figure this out from the little arrow down. The metadata is precise. However, the annotation contains a speech transcript (340 lines), but the transcript does not contain the search term. However, according to the description on the result overview, it should be in there. (RT)

Cold start problem	Maybe an additional and optional layer/page, 'start your research here', could be helpful. Like Spotify or Eventbrite ask for some initial preferences, this might point researchers to possible complementary sources in the other focus areas and thus enrich research. (RT)
Boolean operators	The search function does work well, however, it becomes more powerful when using Boolean operators. It does not make inferences for you like Google would for example, but that is more useful for a tool like the media suite (allows for serendipitous discovery) (JB)
Resource viewer	Why not immediately a link to the page where users can get access instead of a stern message and a link in the metadata. • A user might expect the question mark to return the information from the help section rather than additional information on a tool. Maybe there could be a page with additional tools where screencast and additional tools are explained. • Back to results is also counter-intuitive at first: it sits next to < >, which is for jumping backwards and forwards. (RT)
Interface	<p>Select date field not too intuitive, first need to click on it to see all the features, might be more practical to have it open</p> <p>The 'starting date' did not always change into the actual date when selecting a new starting date. Here, the words 'Start Date' remained on screen instead of the actual start date (01-01-1990, for example)</p> <p>If you are on the 2<sup>nd</sup> page and select a different date, the query runs but it puts you on the 2<sup>nd</sup> page straight away instead of presenting the 1<sup>st</sup> page. (TB)</p>

	<p>The illustrations and icons do not fit the identity of the website and the tools; they are an example of so-called (corporate) flat design. (RT)</p>
<p>About page</p>	<p>Looking at the about section gives a timeline of the events of the development of CLARIAH and what its aims are for the future. However, the about section does not tell users much about why, how and especially the connection between what and how. The pages on the tools, data resources, standards, and workflows give some hints but no details or examples. • The same also goes for the about section of the Mediasuite; a compelling story for people who first come across the Mediasuite is missing, and more explanation about the actors in the project, e.g. why did these actors become partners in the project? Sentences like "It works as an aggregator of the metadata, but also enables access to the media content itself. These collections are registered in one instance of a web-based open source management system called CKAN." are somewhat cryptic and might exclude researchers who are less familiar with the lingo or junior researchers. (RT)</p> <p style="text-align: center;">-</p>