

PubMed log dataset, N=2,996,301

Remove rows containing empty user ID, N=1,146

Remove rows containing unusual user ID, N=73

Remove rows containing null query text, N=77,923

Extract distinct search field tags in the user query (N=2,917,159), group similar tags and map them into unique tag

Exclude queries that contain no search field tag (for association mining analysis), N=2,585,183

Exclude queries containing only single search field Tag (for association mining analysis), N=245,838

Include remaining queries in association mining analysis, N=86,138