Knowledge Graph Based Visual Search Application

Pawandeep Kaur Betz*[1], Tobias Hecking†[1], Andreas Schreiber[1], Andreas Gerndt[1,2]

[1]German Aerospace Center, Institute of Software Technologies (SC), Germany, [2]University of Bremen

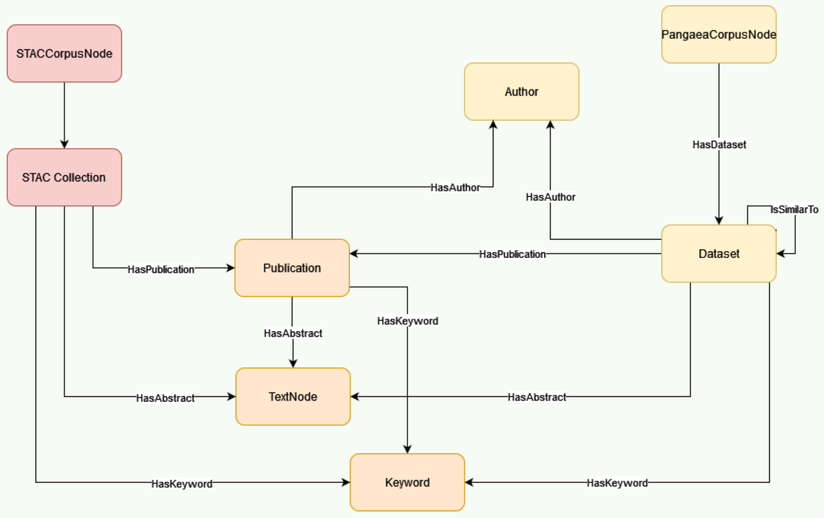
*pawandeep.kaur-betz@dlr.de, †tobias.hecking@dlr.de

MOTIVATION

Traditional text or keyword-based searches often fail to produce needed results unless many different and manual queries are made. While this approach can sometimes lead to success, it often results in long searches with imperfect results. To address these shortcomings, we developed a knowledge graph based visual search application. This application utilizes various chart widgets and a knowledge graph at the backend, connecting two disparate data repositories. Current implementation is done for the Earth System Science (ESS) datasets.

KNOWLEDGE GRAPH

Datasets and Spatio Temporal Asset Catalogues (STAC) are organised in a Knowledge Graph hosted in ArangoDB. Datasets are interlinked by common keyword associations. For Pangaea datasets keywords are provided as metadata. For STAC keywords were associated indirectly from publications that reference the respective catalogues.



VISUALISATION ENABLED SEARCH APPLICATION

Visualisation Enabled Search Application (VESA) is a frontend of this application. It visualizes the metadata harvested by our Knowledge Graph in different dimensions.

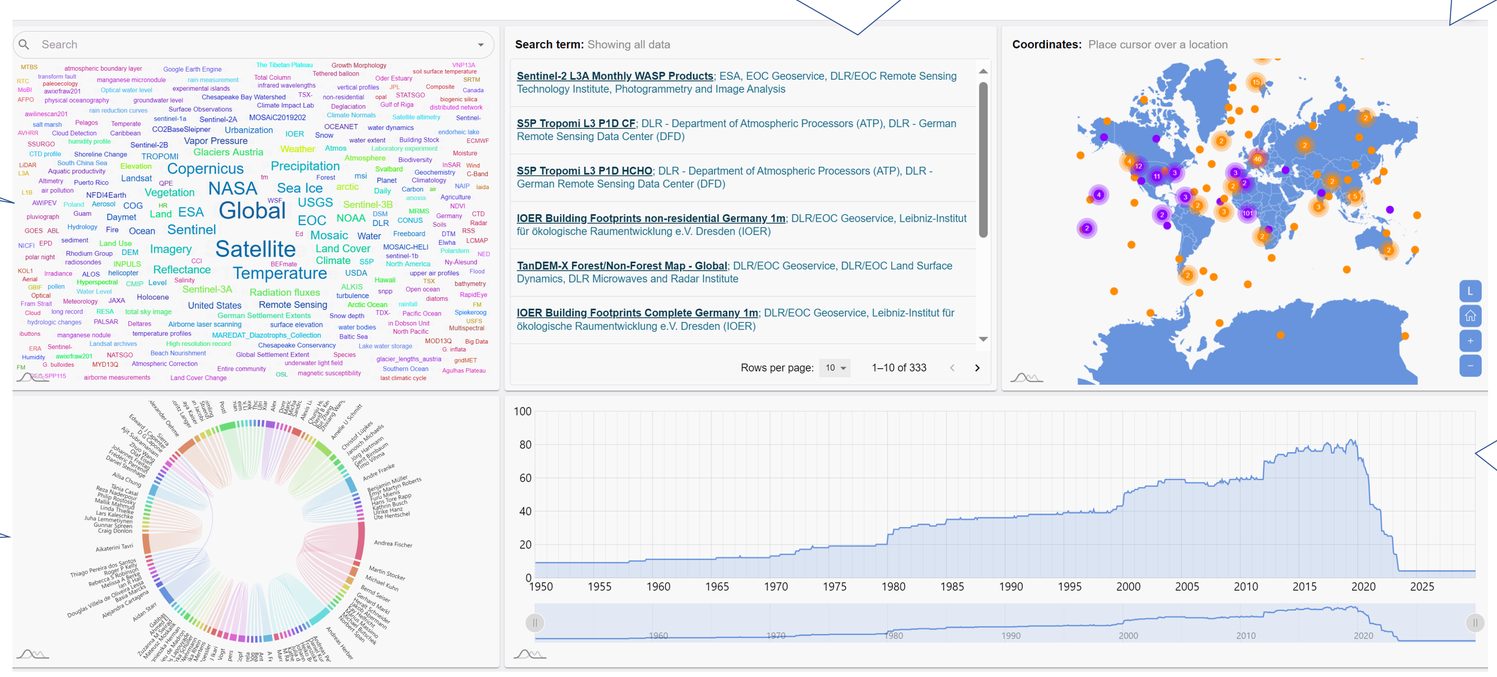
WORDCLOUD
Shows the themantic information about the datasets. Keywords or Tags from the metadata are linked with it. The weight of each keyword shows the frequency of that keyword in the data store

LIST
Shows the resultant datasets from the interactions on the visual filters. These datasets are linked to the actual DOI page where the raw data and metadata is available

MAP
Shows the location where these datasets were observed or collected

CHORD DIAGRAM
Shows the number of datasets that each author has in common with other authors

LINE CHARTS
Shows the year and month when these datasets were observed or collected



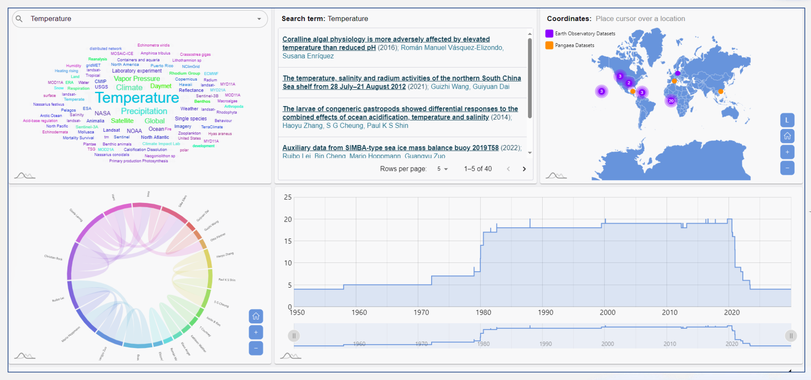
USE CASE

*Meet Dr. Smith who studies Global Warming. He is in need of **historical** **oceanic** data to compare his results.*

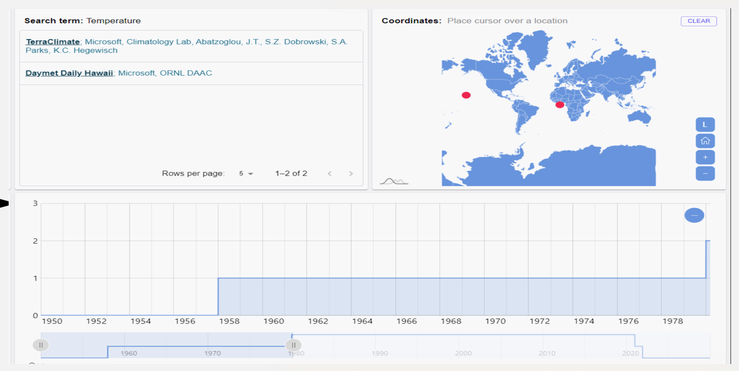
„most of the datasets are close to Africa by Atlantic Ocean“

As he was only interested in oceanic temperature, so he selected one data from Atlantic Ocean and other from Pacific Ocean

He starts by clicking ‚Temperature‘ on the word cloud or typing it over search bar



„rise in temperature dataset since 1980“



As he needed only historical data, so he filters data before 1980

www.vesa.webapps.nfdi4earth.de

This work has been funded by the German Research Founda-tion (DFG) through the project NFDI4Earth (DFG project no.460036893, https://www.nfdi4earth.de) within the German National Research Data Infrastructure (NFDI, https://www.nfdi.de).