A Bibliometric Analysis of Ice Core Science

Pramithas Upreti

Faculty Advisor: Dr. Bess Koffman, Department of Geology

The purpose of this project is to evaluate how research themes and geographic focus areas within ice core science have evolved over the past five decades.

Photo of a freshly drilled ice core by David Ferris

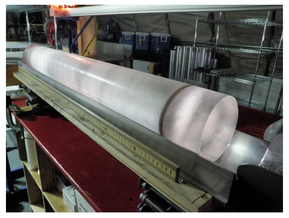
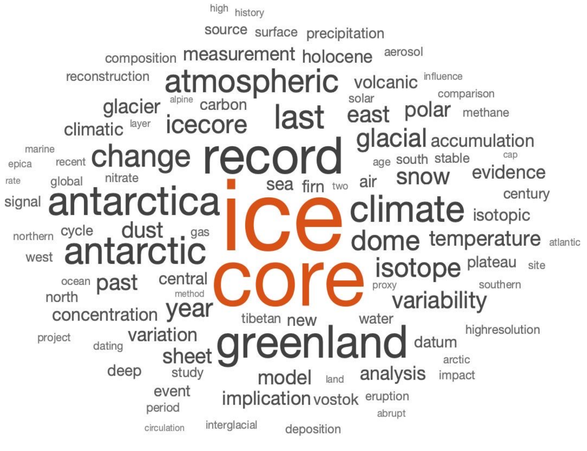


Photo of a volcanic ash layer in an Antarctic ice core by Heidi Roop

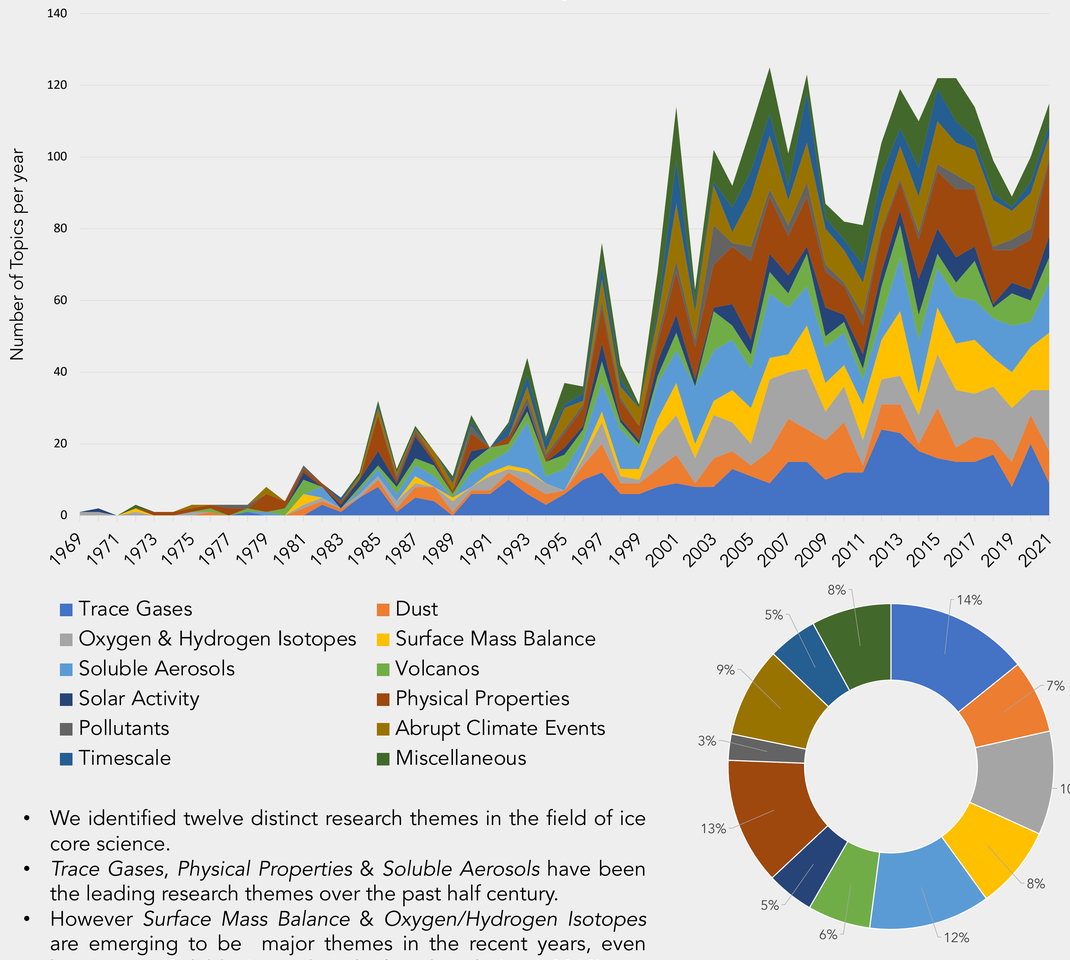
Methods

- We implemented artificial intelligence through a rigorous computational framework in Python and MATLAB.
- For the analysis of temporal themes and hierarchical topic reduction, we used BERTopic's advanced clustering model.
- Among the 2772 abstracts, about 2400 mentioned geographical locations. Each of them was then analyzed and placed into one of three sub-groups: Antarctica, Greenland or Others.



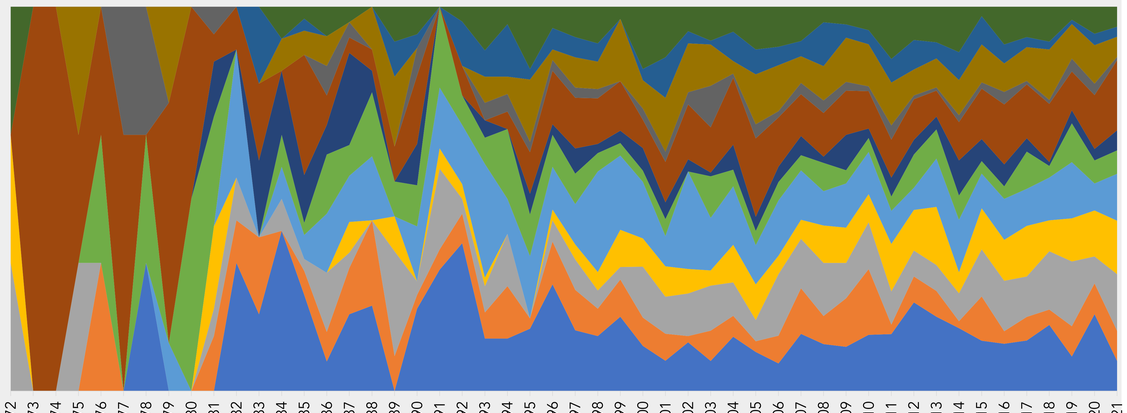
A WordCloud showcasing the frequency count for ice core science literature

Dynamic Topic Modeling



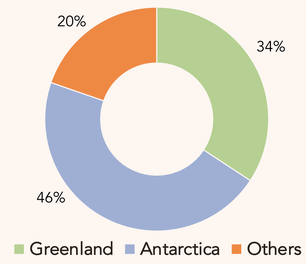
- We identified twelve distinct research themes in the field of ice core science.
- *Trace Gases*, *Physical Properties* & *Soluble Aerosols* have been the leading research themes over the past half century.
- However *Surface Mass Balance* & *Oxygen/Hydrogen Isotopes* are emerging to be major themes in the recent years, even beating out *Soluble Aerosols* in the last decade (post-2013).

A pie chart representing the proportions of the 12 topic groups for the 2772 abstracts.
(same legend as above)

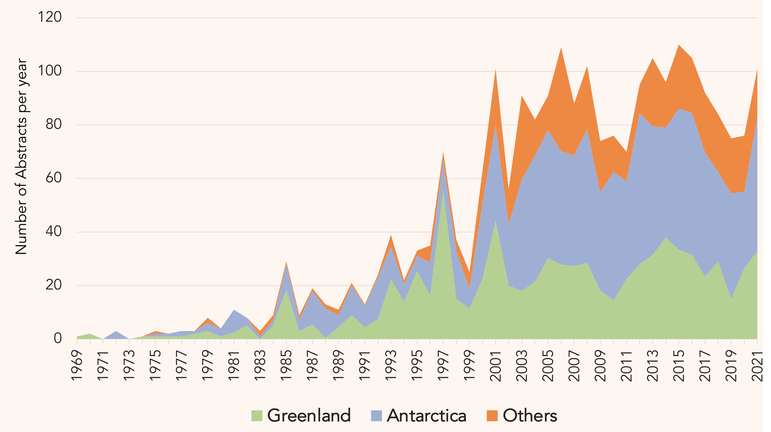


A 100% Stacked Area Chart exhibiting the proportions of each of the12 topic groups over time from 1972-2021
(same legend as above)

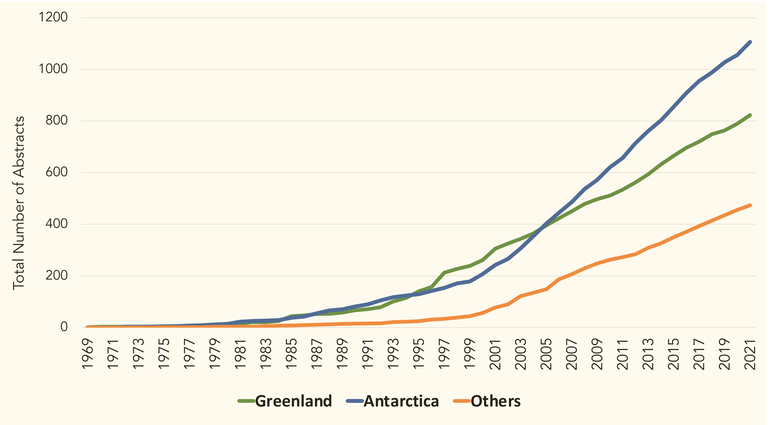
Geographical Distribution



- The geographic exploration portion of the project has unsurprisingly ascertained Antarctica and Greenland as the hubs for ice core science research.
- Antarctica accounts for almost half of ice core literature, Greenland accounts for about a third and the rest of the world accounts for about a fifth of the research landscape.



A 2D area graph showcasing the geographical distribution of ice core science literature over time from 1969-2021



A cumulative line graph showcasing the geographic distribution of ice core science literature over time from 1969-2021. Antarctica overtook Greenland only in 2005.

Acknowledgement

This work was funded by the Data Science Program. Special regards to Dr. Koffman for her support throughout the project. Thanks to Dr. Matt Osman for initial preparation of the dataset.

Future Work

We hope to refine the AI model through a semi-supervised learning approach, which may yield a greater accuracy for topic classification.