

Pairing NASA Capabilities and Western US Water Resource Stakeholders with Data Mining and Machine Learning



Jet Propulsion Laboratory
California Institute of Technology

Western Water Applications Office

Christine M. Lee, Valentino Constantinou, Asitang Mishra (christine.m.lee@jpl.nasa.gov)

NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA

ABSTRACT

Water resource agencies are responsible for the efficient and sustainable management of water resources, which may include rainfall capture or distribution to customers. Partnerships between water resource agencies (stakeholders) and principal investigators (capabilities) serve to improve water resource management in areas such as estimation of current and future snow-water equivalent (SWE) or effective capture of agricultural run-off. In this work, we report preliminary results from an approach that seeks to identify partnership opportunities between capabilities and stakeholders within the Western Water Applications Office (WWAO). By mining texts such as publications and abstracts using natural language processing techniques (NLP), partnerships can be recommended based on shared research interests, resource management goals, and other infomation encoded in natural language - aiding the expert-driven but manual approach for pairing stakeholders and capabilities used currently. This work shows encouraging preliminary results for pairing stakeholders and capabilities using NLP techniques and describes avenues for future work. This poster shows a proof-of-concept of a workflow aimed to better characterize synergies between science data and resource management needs.

MOTIVATION

Western US water issues are highly complex and challenging. This landscape represents a data science problem whereby natural language processing techniques can be used to analyze, prioritize, and optimize strategic partnership efforts. The complexity of the western U.S. water resource landscape presents an opportunity for improved knowledge management.

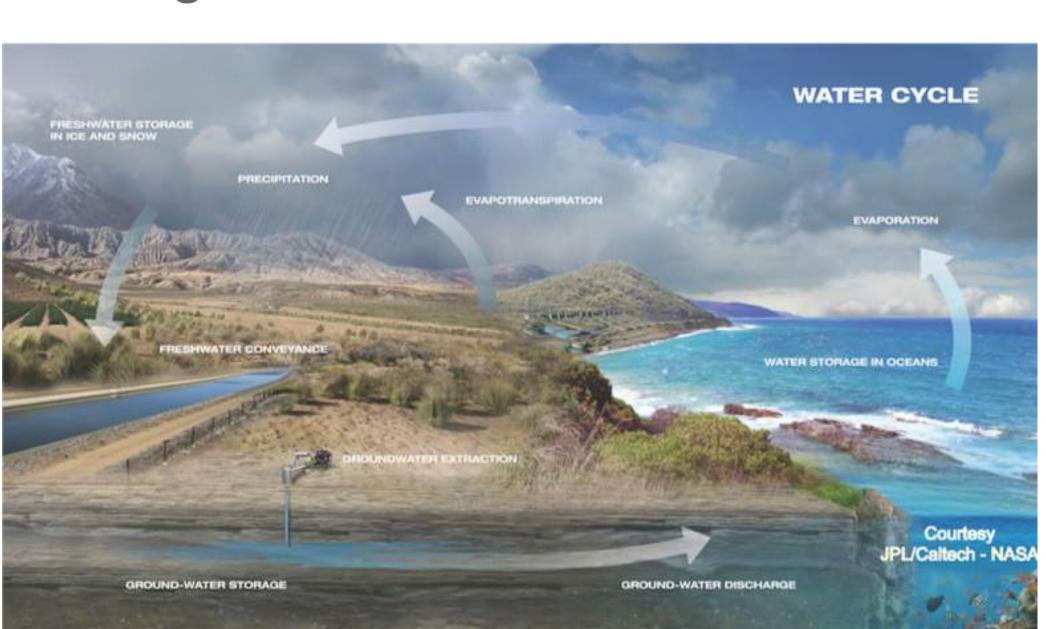


Figure 1: an illustration of the hydrological cycle.

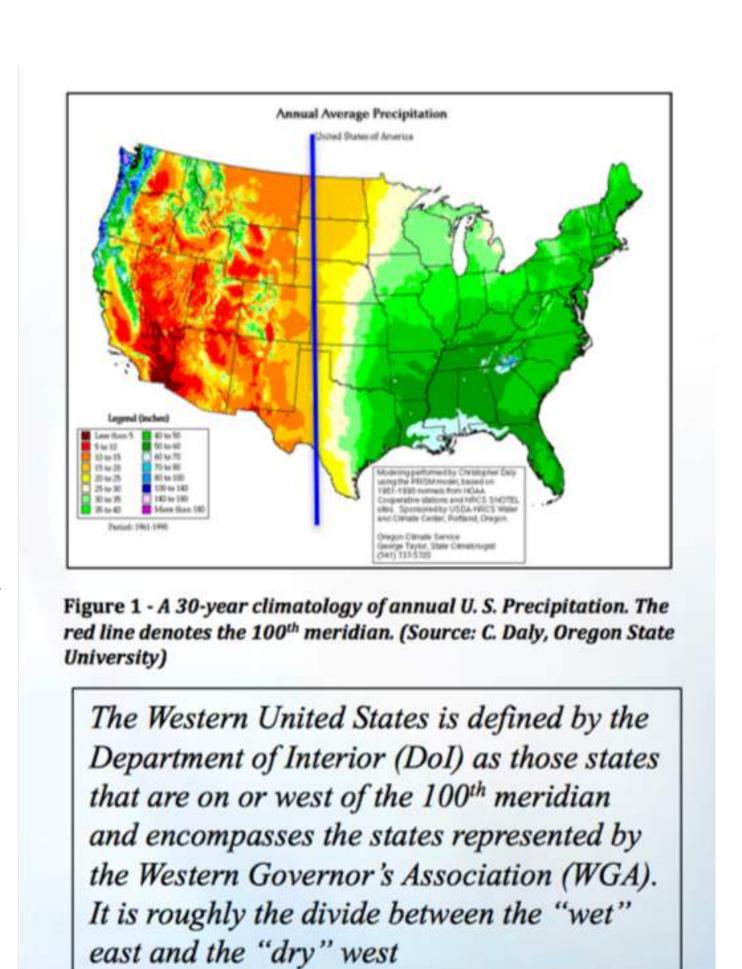
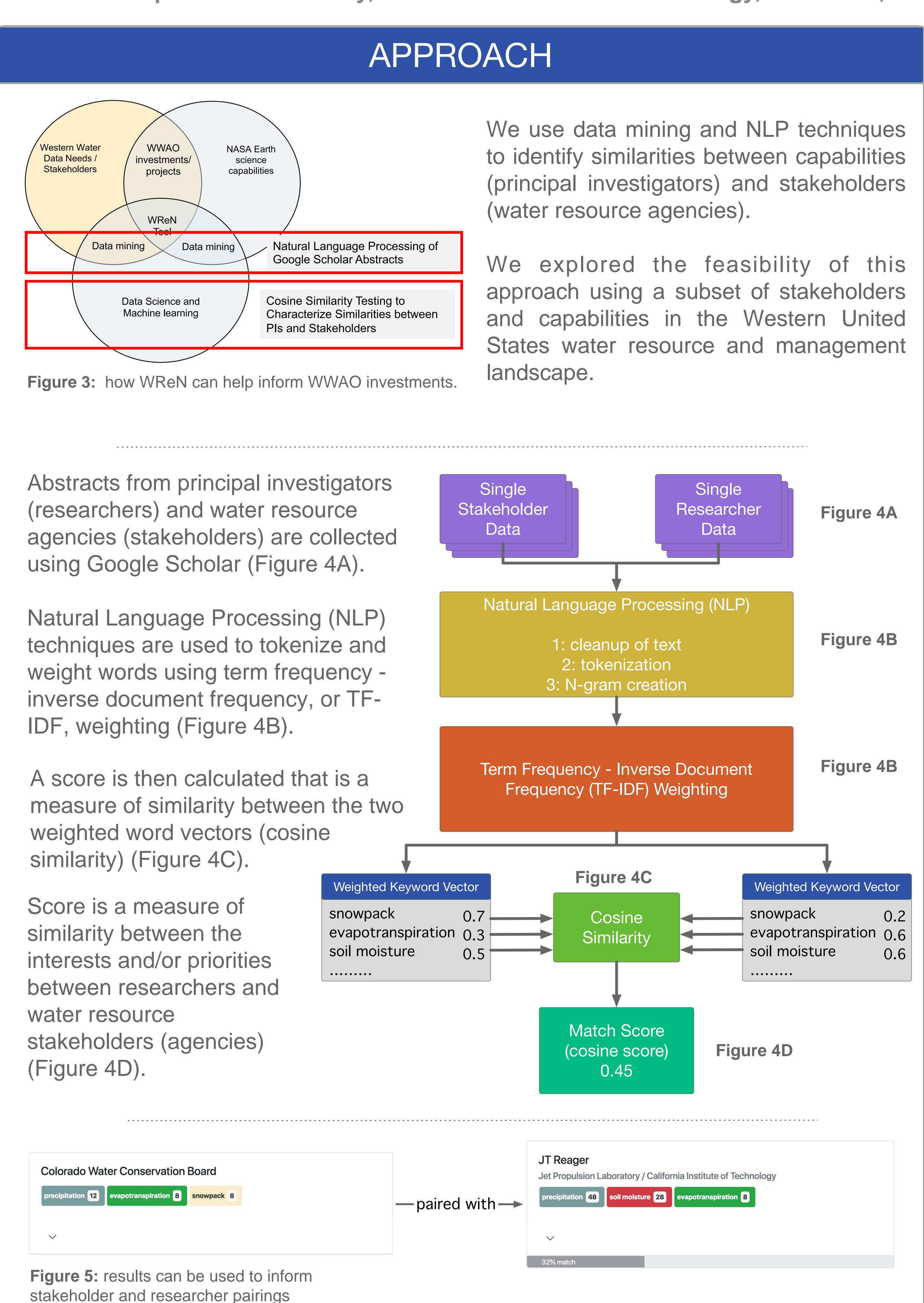


Figure 2: the Western United States as defined by the Department of the Interior (DoI).



APPROACH

Findings: An end-to-end system can be developed for the extraction and organization of information contained in text data. Natural Language Processing (NLP) approaches can be applied for use in knowledge management of water resource interests.

Data collection is a challenge. Google Scholar a one-stop shop for abstracts, but available data limited to most recent 30 listed publications.

Further characterization of water resource stakeholders and researchers is a possibility from the collected textual data. We examined stakeholder water resource variable interests within the Colorado River Basin (preliminary results figure 6 below).

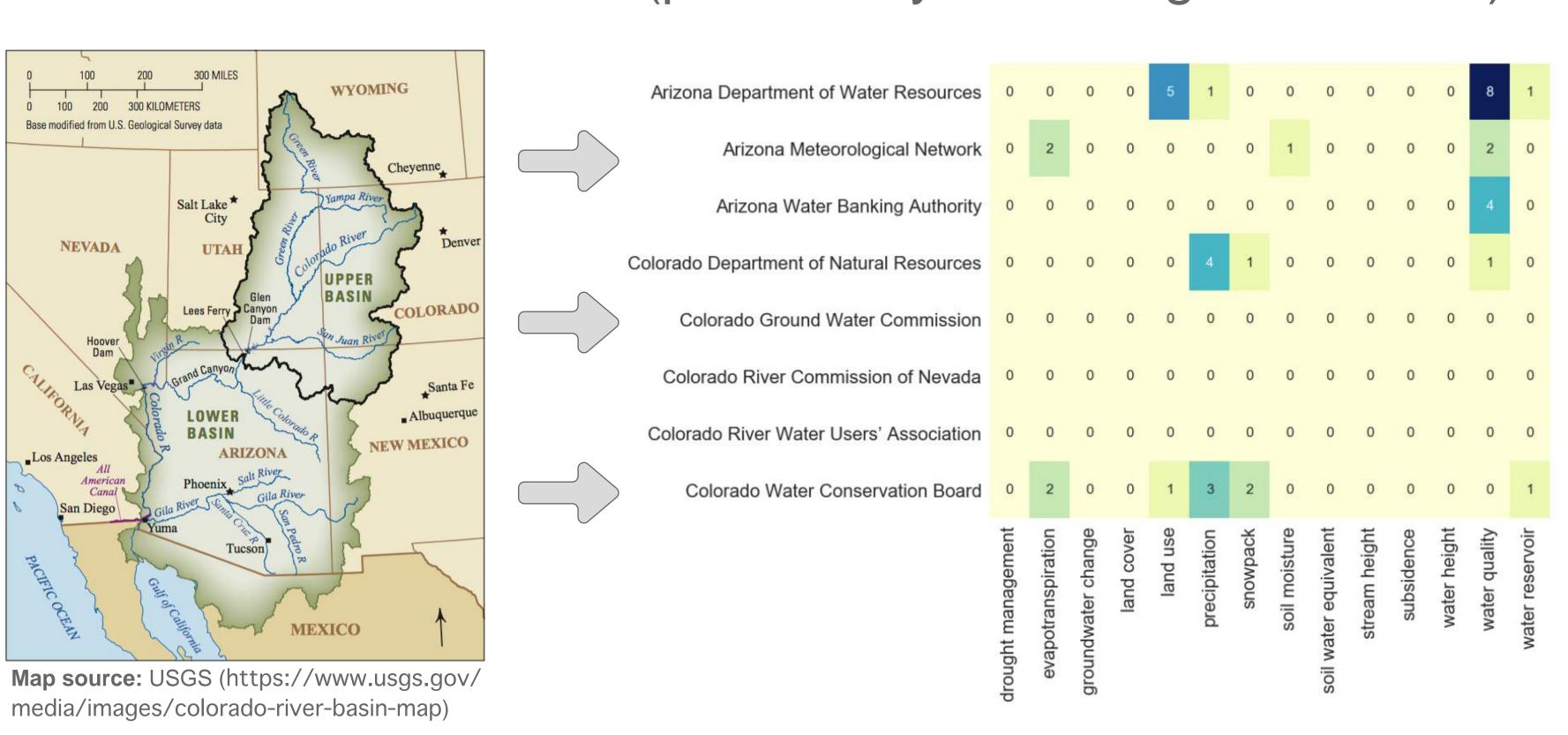


Figure 6: water resource variable interest by Colorado River basin water resource stakeholder. Water quality is mentioned frequently in association with the Arizona Department of Water Resources, Arizona Meteorological Network, and the Arizona Water Banking Authority. Note that some stakeholders show no counts, indicating the need for future work in data mining (gathering texts for analysis).

FUTURE WORK

Demonstrate functionality of approach and tool to supplement Western Water Applications Office (WWAO) planning and project activities through iterative product feedback. Improve generalization and characterization with additional data.

Conduct trends analysis of water resource variables (e.g. drought management, snow-water equivalent) by time period, governance level within a particular river basin.

ACKNOWLEDGEMENTS

We acknowledge the support of the Western Water Applications Office and the Data Science Working Group (DSWG) for funding and guiding this preliminary effort.