H25V-1275

Challenges and opportunities in coding the governance of water resource systems: problems, procedures, and potential solutions based on the Lake Mendocino reservoir, California.

Samuel Park[1], Tomás Olivier[2], Parsa Pouladi[1], Peyman Yousefi[1], Margaret Ellen Garcia[3], and David J. Yu[1]

[1] Purdue University, USA, [2] Florida Atlantic University, USA, [3] Arizona State University, USA

Summary

- A network representation of water infrastructures and actors (nodes) linked by regulatory rules (edges) can facilitate understanding how a change in infrastructure rules might affect a complex water resource system.
- We developed an automated tool that identifies such a network representation using text analysis of policy documents.
- Our tool can be used to explore potential misfits in interdependent nodes involved in infrastructure management

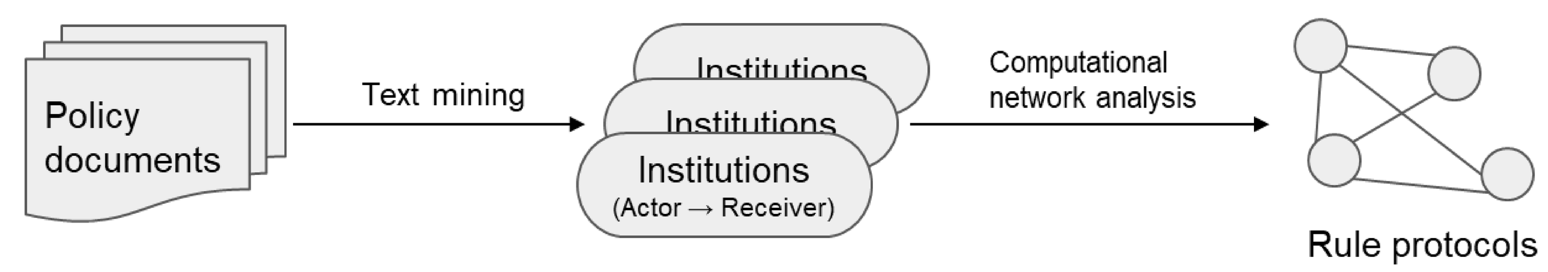


Figure 1. Overview of developing the automated text mining tool

1. Introduction

- Many cities and counties in the US establish formal plans & regulations to shape stakeholders' decisions that concern the use and provision of water resources and relevant public infrastructures.
- Such rules are created to **prescribe what actions are to be done, by whom, and under what conditions**, with the assumption that following these rules will likely lead to more effective outcomes.
- Such prescriptions (also referred to as *institutions*) can be **conceptualized as a network of actors and infrastructures linked by rules**.

Box 1. Prescribed Rules as a Network (an example based on a water resource system)

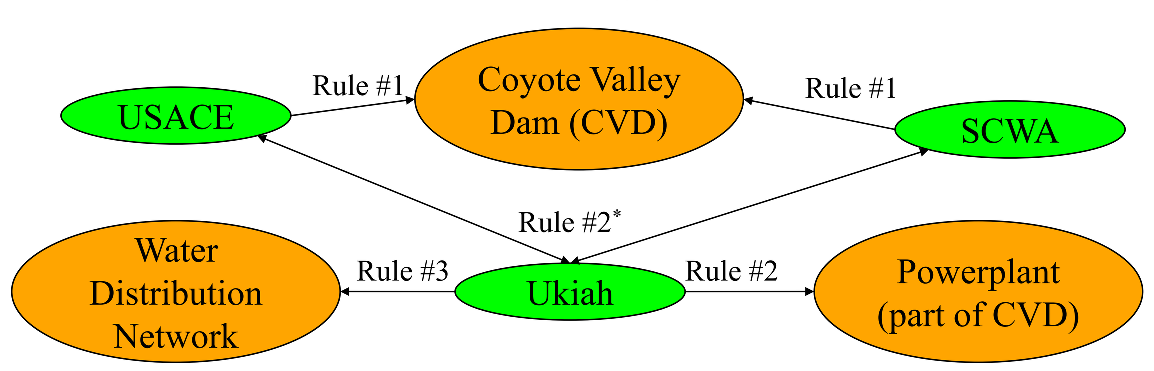
Rule #1: USACE determines the schedule and amount of water released from Lake Mendocino (Coyote Valley Dam) during flood control operations, while Sonoma County Water Agency (SCWA) manages releases from the conservation pool.

Rule #2: The City of Ukiah is responsible for: (1) Physically operating the powerplant to meet the release requirements.

USACE (1986), Coyote Valley Dam and Lake Mendocino Water Control Manual.

Rule #3: [City of Ukiah] Work to avoid preventable water loss from the City's water distribution system.

The City of Ukiah (1995), Ukiah Valley General Plan and Growth Management Program.



* "meet the release requirements…" (Rule #2)

This network-based approach can help address **two research problems** in the study of interdependencies in critical infrastructure (e.g., Rinaldi et al. 2001):

1) Hard to obtain a systematic view into how infrastructures and actors are *logically* interconnected (i.e., linked by human decisions prescribed by rules).
2) Hard to detect logical interdependency *misfits* (i.e., mismatch of a rule with other rules or physical nodes) or propagative effects of a change in infrastructure rules.

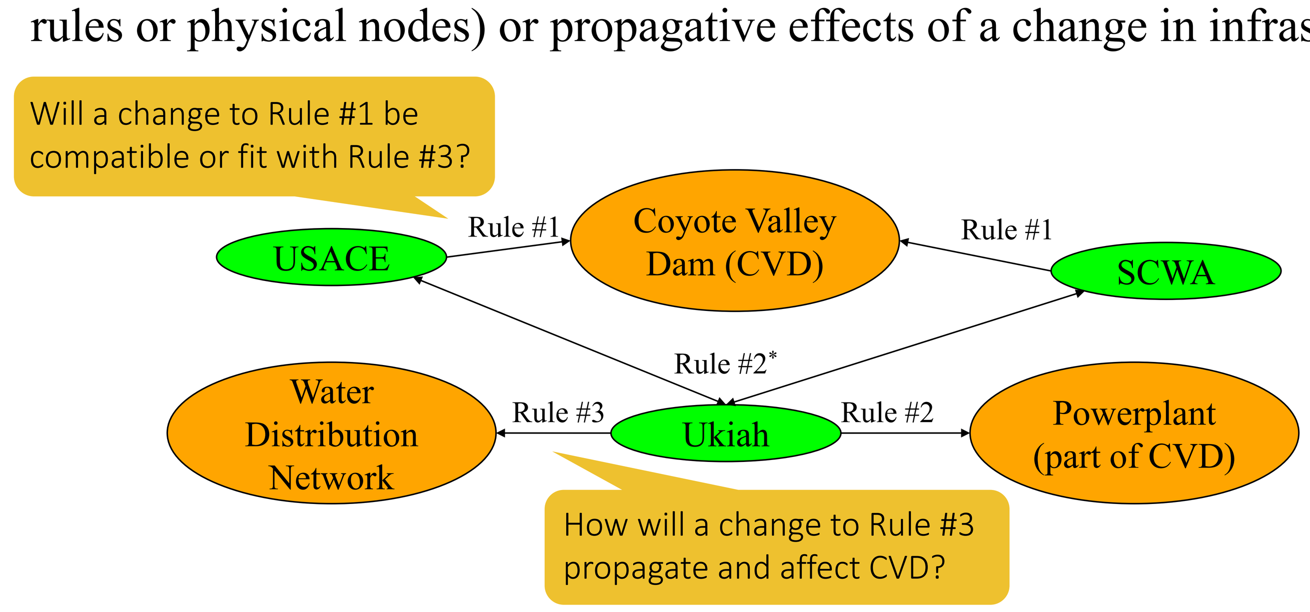


Figure 2. Logical interdependency issues in a complex water resources system

2. Purpose of the study

- We developed an automated approach to create a parsimonious network representation of rules, actors, and infrastructure in three interlinked situations (dam operation, flood response, and water supply).
- We used planning documents for the dam operation, flood response, and water supply situations of Lake Mendocino (Coyote Valley Dam), California as our case data.

3. Methods

Our approach is composed of two parts. The first part was to train an automated text mining tool for policy documents by adapting the approach used by Rice et al. (2021). In order to systematically figure out institutions from policy documents, we used a theoretically informed syntax, called the Institutional Grammar (IG) (Crawford & Ostrom, 1995).

Table 1. Syntax of Institutional Grammar (IG)

IG syntax	Meaning	Explanation
Attribute [A]	Who	Actor to whom an institutional statement applies
oBject [B]	Whom	Receiver of the focal action
Deontic [D]	Must, shall, may	Prescriptive operator
aIm [I]	What	Action of the statement
Condition [C]	When, Where, How	Temporal, spatial or procedural boundaries
Or else [O]	If not	(Dis-) Incentives for (not) doing the focal action

E.g., "Facility owners shall annually submit a permit renewal application or pay a fine"
Attribute Deontic Condition aIm oBject Or else

After preprocessing raw texts from policy documents, we annotated them with syntactic and semantic properties (e.g., Parts-of-speech, Dependency, Named Entity Recognition) using Natural Language Processing (Stanford's CoreNLP). Using training data manually labeled with IG, we trained a neural network classifier (Multilayer Perceptron) to map the annotated word to one of IG elements (ABDICO).

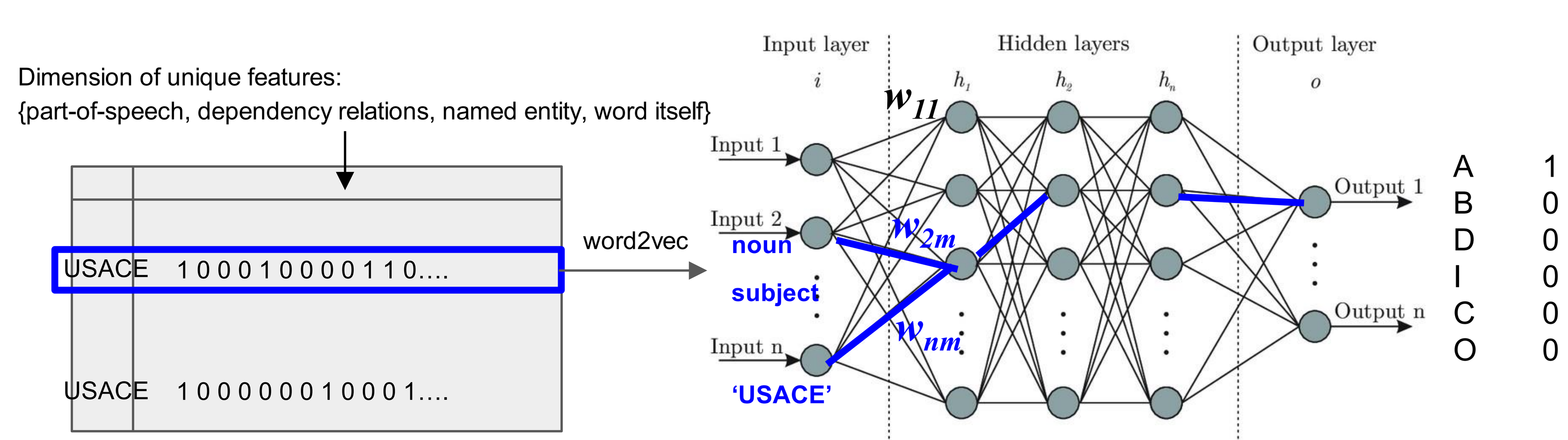


Figure 3. Deep learning model to classify Institutional Grammar elements

The second part of this tool was to visualize and analyze the designed policy network using the given classified institutional statements (including AIC at least). We made a pair of extracted Attribute and oBject elements for identifying actor-receiver relations. Then, we created and analyzed designed policy networks.

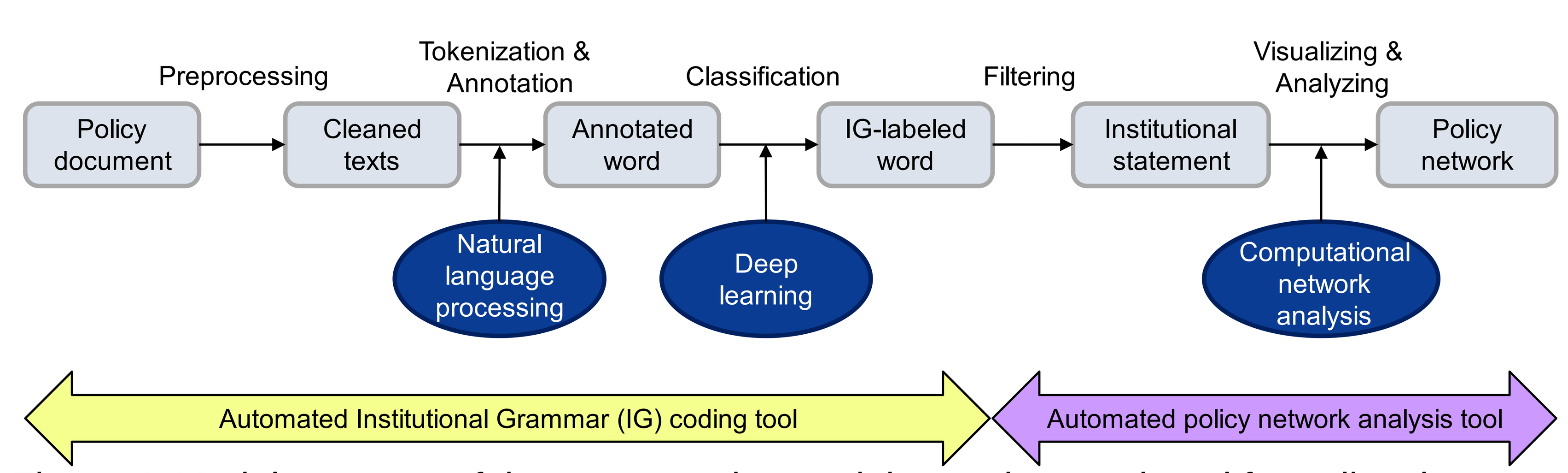


Figure 4. Model structure of the automated text mining and network tool for policy documents

4. Results

In results, the automated machine coding showed reliable accuracy and model performance in comparison with the criteria of acceptable intercoder reliability test (0.7~0.9) (Quinn et al., 2010). The results provide evidence of the potential of supervised machine learning for the automated classification of IG.

Table 2. Classification performance of 3-layer and 5-layer NN model

Performance of 3-Layer Neural Network Classifier

Class	Precision	Recall
Attribute	0.68	0.37
Object	0.89	0.95
Deontic	0.84	0.81
Aim	0.64	0.29
Condition	1.00	1.00
Or/Else	0.73	0.92
Overall accuracy	0.73	

Performance of 5-Layer Neural Network Classifier

Class	Precision	Recall
Attribute	0.81	0.30
Object	0.87	0.95
Deontic	0.85	0.76
Aim	0.65	0.28
Condition	1.00	1.00
Or/Else	0.73	0.93
Overall accuracy	0.73	

We created a formal representation of how identified stakeholders and critical infrastructure systems are linked by prescribed rules and strategies. Figure 5 is an example of rule protocols in terms of indegree centrality.

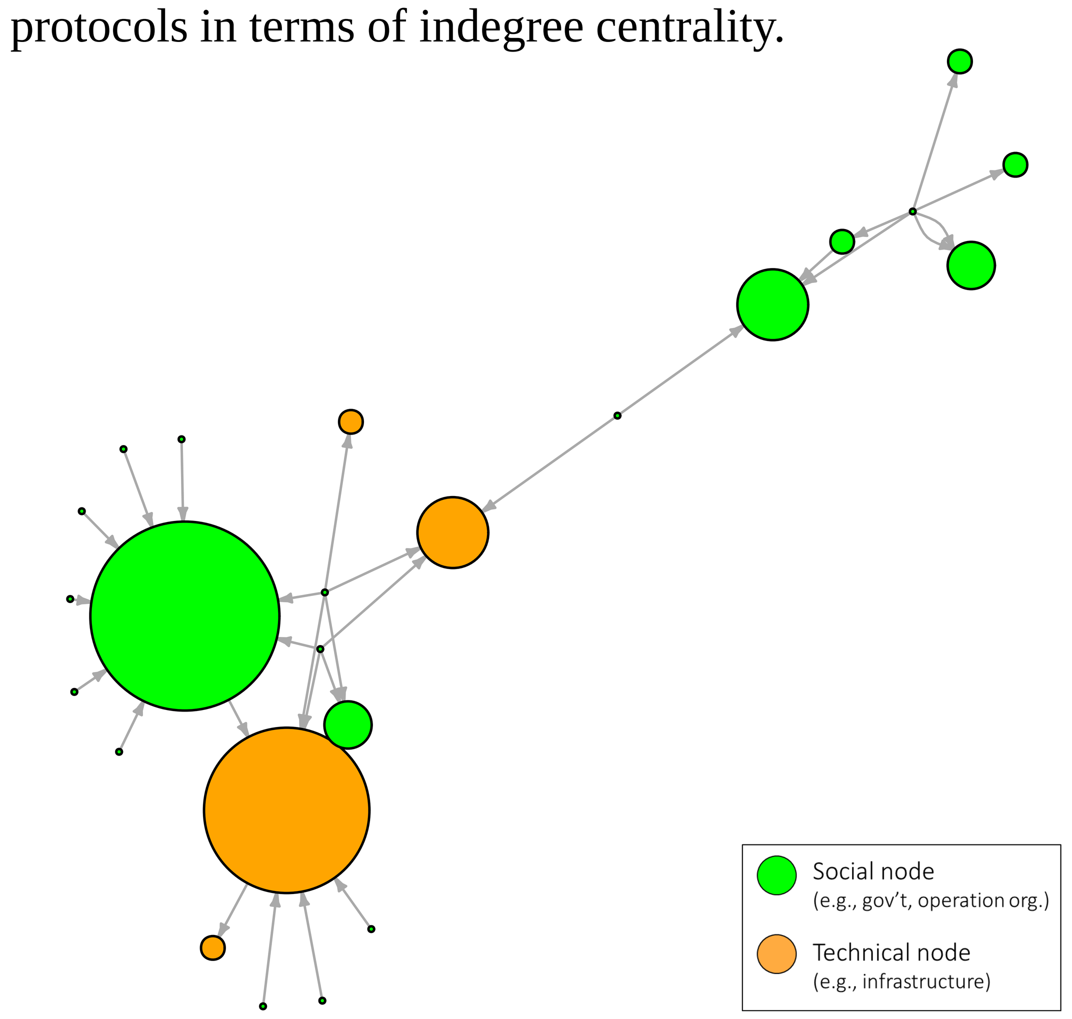


Figure 5. Rule protocols of Coyote Valley Dam (Scaled by indegree centrality)

5. Discussion and Conclusion

Our automated text mining and network visualization tool achieved acceptable accuracy (overall 74%) with a relatively small training set compared to manual coding. Our tool can facilitate understanding how a rule change might propagate and affect the overall social-technical network. For future study, more training data and better preprocessing is needed to increase the classification accuracy. This tool has broader implications, where a civil engineering project is needed to compare and analyze the expected policy outcome in comprehensive ways.

Acknowledgement

This work has been supported by a National Science Foundation grant (#1913665) and the Doctoral Fellowship provided by the Purdue Systems Collaboratory.

References

- Crawford, S. E., & Ostrom, E. (1995). A grammar of institutions. American political science review, 89(3), 582-600.
- Rice, D., Siddiki, S., Frey, S., Kwon, J. H., & Sawyer, A. (2021). Machine coding of policy texts with the Institutional Grammar. Public Administration, 99(2), 248-262.
- Rinaldi, S. M., Peerenboom, J. P., Kelly, T. K., Rinaldi, B. S. M., Peerenboom, J. P., & Kelly, T. K. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies. IEEE Control Systems, 21(6), 11–25.
- Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H., & Radev, D. R. (2010). How to analyze political attention with minimal assumptions and costs. American Journal of Political Science.

* Contact: Samuel Park (park1029@purdue.edu)