

THE RESEARCH UNIVERSITY PRESS

International
Computational Law Journal

VOLUME 2 / ISSUE 5 / CURATED SOURCE EDITION

Deriving Insights from Legal Data

Robert Mahari · Dazza Greenwood

MIT Computational Law Report · Originally published 2022-02-25

Independent reader edition under CC BY 4.0. This attribution cover was added; original source PDF pages follow unchanged. Original authorship and source identity are retained. MIT, Stanford University, Stanford Law School and CodeX have not sponsored or endorsed this republication.

Source, license and version evidence accompany the online article.

MIT Computational Law Report • Media

Deriving Insights from Legal Data

Robert Mahari, Dazza Greenwood

Published on: Feb 25, 2022

License: Creative Commons Attribution 4.0 International License (CC-BY 4.0)

IdeaFlow 11 - Deriving Computational Insights from Legal Data

Visit the web version of this article to view interactive content.

law.MIT.edu's IdeaFlow Episode 11: Deriving Computational Insights from Legal Data

- *To signup for future IdeaFlow episodes:* https://forms.gle/aZbxfsEHkM7egeyf8
- For more episodes of IdeaFlow, see: https://law.mit.edu/media

This episode explores two perspectives on deriving computational insights from legal data.

Host: Dazza Greenwood, Executive Director, law.MIT.edu

Guest: Robert Mahari is pursuing a joint JD/PhD degree at Harvard Law School and the MIT Media Lab's Human Dynamics Group

Date: Friday, February 25th, 12:00 PM ET

First, we will discuss the law as a knowledge system that grows by means of citations. We will compare the citation networks in law and science by leveraging tools from "science-of-science". We will explore how, despite the fundamental differences between the two systems, the core citation dynamics are remarkably universal, suggesting that the citation dynamics are largely shaped by intrinsic human constraints and robust against the numerous factors that distinguish the law and science.

Second, we will explore how legal citation data can be used to build sophisticated NLP models that can aid in forming legal arguments by predicting relevant passages of precedent given the summary of an argument. We will discuss a state-of-the-art BERT model, trained on 530,000 examples of legal arguments made by U.S. federal judges, which predict relevant passages from precedential court decisions given a brief legal argument. We will highlight how this model performs well on unseen examples (with a top-10 prediction accuracy of 96%) and how it handles arguments from real legal briefs.

After the segment on deriving computational insights from legal data, we will also provide discussion time for the upcoming special release of the MIT Computational

Law Report on "Composable Governance" (https://law.mit.edu/composablegovernance) and how you can get involved.

Background on Robert Mahari:

> *LinkedIn*
>
> 
>
> *Robert Mahari is pursuing a joint JD/PhD degree at Harvard Law School and the MIT Media Lab's Human Dynamics Group. His work focuses particularly on how computational legal systems can increase the efficiency of legal processes and promote access to justice. To this end, Robert leverages tools from machine learning, game theory, and network science to research and prototype computational legal solutions in collaboration with private and public entities around the world.*