

Internet Law's Interdisciplinary Gap?

Dan Bateyko
Cornell University



Motivation

Internet Law scholars must meaningfully engage with technology.
How do they perform this commitment?

Citation patterns might offer a glimpse. I propose three measures for evaluating Internet Law's interdisciplinary work: (1) citation mix, (2) citation distribution, and (3) citation timeliness. I aim to identify and showcase archetypes or styles of Internet Law writing across a large body of Internet Law scholarship.

The View from Computer Science

CS expertise can used for...

1. Rebuttal
2. Synchedoche
3. Diagnostic
4. Formalizer

The View from Computer Science

CS expertise can used for...

- 1. Rebuttal**
- 2. Synchedoche**
3. Diagnostic
4. Formalizer

Computing as Rebuttal

“Computing can clarify the limits of technical interventions, and of policies premised on them.”

“

...we encountered misconceptions about software, APIs, and compatibility. We suspect that Oracle’s lawyers introduced most of these misconceptions for strategic reasons...

— JOSHUA BLOCH AND PAMELA SAMUELSON (2022)

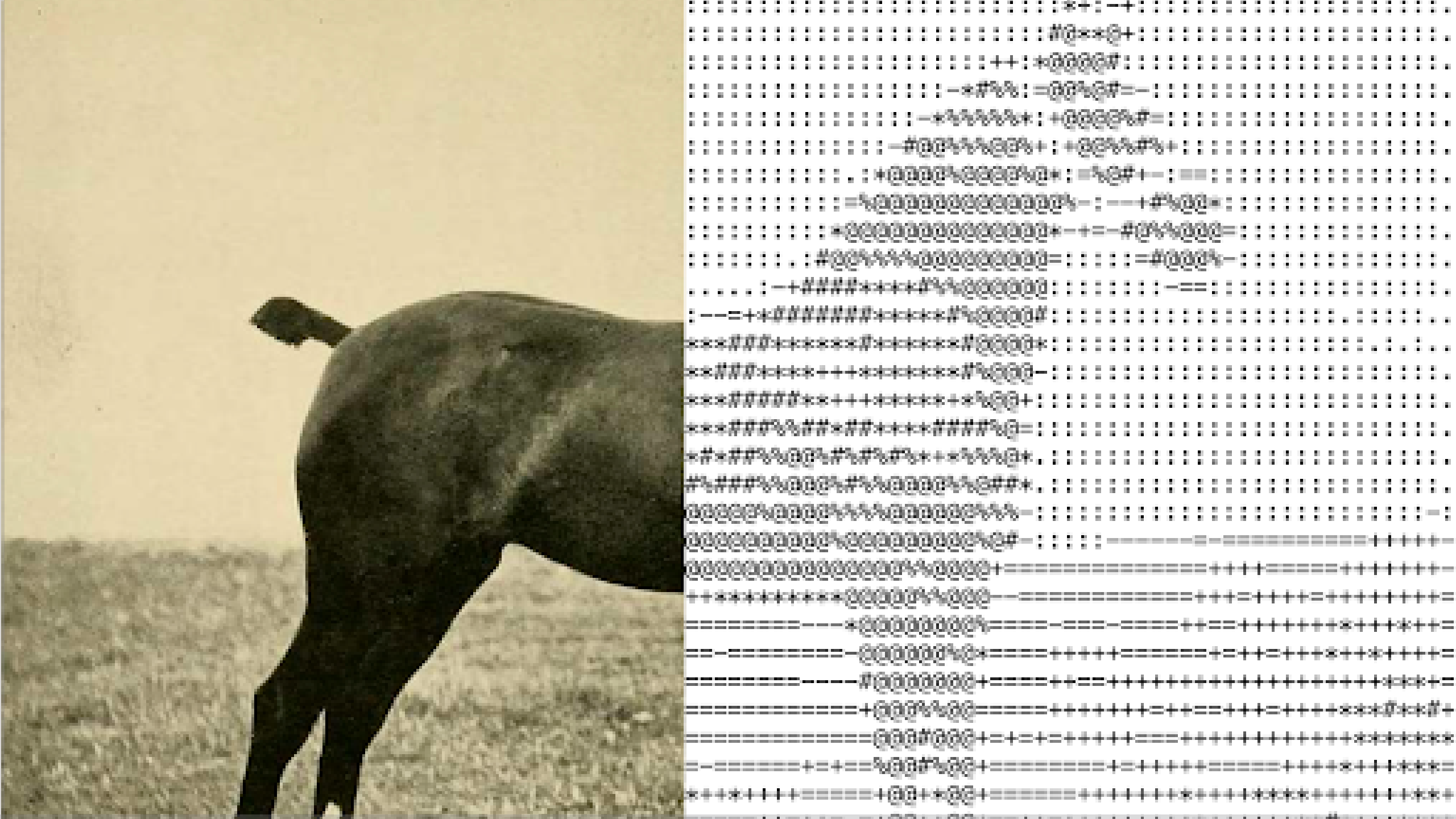
“

Doomed to be shallow...the worst of both worlds...the cross-sterilization of ideas."

— JUDGE FRANK EASTERBROOK

Computing as Synecdoche

“Computing can foreground long-standing social problems in a new way.”



“

The trouble with technology—and hence law and technology scholarship—is that truly novel affordances tend to invite reexamination of how we live.

— RYAN CALO

Can we capture these approaches
bibliometrically?



Method



“

U.S. legal scholarship is weird. It is weird in many ways: articles are very long, footnotes are innumerable, and the article selection process is unlike any other.

—JONATHAN I. TIETZ AND NICHOLSON PRICE II (ACKNOWLEDGEMENTS AS A WINDOWS INTO LEGAL ACADEMIA)

Dataset by the numbers

183

Unique Law Reviews

62,031

Articles

313,874,435

Words in the footnotes

Three-stage LLM pipeline

1. Prompt to classify article's subfield

- a. Washington And Lee Subject Categories for Law Reviews (e.g. Science, Technology, and Computing Law)

2. Prompt to split footnotes into citations

- a. Evaluation (WIP): 100% accuracy on 30 test cases

3. Prompt to classify citations

- a. Classify citations into 13 standardized categories using detailed type-specific recognition rules



Analysis

1. Citation Mix

- a. An article's citations to law, computer science, gray literature, news, etc

2. Citation Distribution

- a. Position-based metric of *where* the citation appears in the article

3. Citation Timeliness

- a. Compare citation years to year of article



Preliminary Results



Full Citation	Type	Year	Author(s)
Julie E. Cohen, <i>Between Truth and Power</i> , Oxford University Press (2019).	Book	2019	Julie E. Cohen
Jonathan L. Zittrain, <i>The Generative Internet</i> , 119 Harv. L. Rev. 1974 (2006).	Law Review	2006	Jonathan L. Zittrain
Allison Koenecke (Cornell University, USA), Anna Seo Gyeong Choi (Cornell University, USA), Katelyn X. Mei (University of Washington, USA), Hilke Schellmann (New York University, USA), and Mona Sloane (University of Virginia, USA), <i>Careless Whisper: Speech-to-Text Hallucination Harms</i> .	Computer Science Article	2024	Allison Koenecke, Anna Seo Gyeong Choi, Katelyn X. Mei, Hilke Schellmann, Mona Sloane

Beginning				Middle				End		
	1	2	3	4	5	6	7	8	9	10
Citation Type	academic-paper	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	blog-post	0.05	0.02	0.03	0.03	0.03	0.04	0.02	0.01	0.02
	books-and-treatises	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02
	case-law	0.08	0.04	0.04	0.03	0.02	0.02	0.02	0.02	0.02
	computer-science	0.10	0.15	0.16	0.13	0.11	0.16	0.18	0.15	0.17
	cross-reference	0.02	0.01	0.01	0.02	0.03	0.01	0.01	0.01	0.02
	gray-literature	0.26	0.38	0.37	0.39	0.44	0.45	0.44	0.46	0.42
	law-review	0.09	0.06	0.05	0.07	0.07	0.07	0.05	0.05	0.07
	legal-authority	0.13	0.10	0.06	0.08	0.07	0.05	0.05	0.08	0.07
	news-article	0.08	0.13	0.14	0.12	0.10	0.10	0.11	0.11	0.11
	other	0.14	0.08	0.10	0.10	0.09	0.06	0.07	0.07	0.09
	social-media	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	working-paper	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.01

Average Citation Distribution from 50 Internet Law Articles

Next steps



Next steps

How do Internet Law articles compare to other interdisciplinary law fields on these measures?

Are there trends over time?

What might be better measures for interdisciplinarity?

Questions?

Dan Bateyko
Contact: drb348@cornell.edu



Which Law Articles “read” the closest to CS?

- Compare a corpus of Internet Law; CS&Law; arXiv
- Embedding arithmetic (e.g. King - Man + Woman = Queen) but for citation networks (vCS-vLaw)

Journal A → Journal X → Journal C

Journal A → Journal Y → Journal C

Journals that function similarly in a trail will have similar representations, measured by cosine similarity.

