

AI FOR CYBERJUSTICE: COUNTING AS A HUMAN BEING IN THE ERA OF COMPUTATIONAL LAW?

Mireille Hildebrandt, FBA
PI COHUBICOL ERC ADG project

some hyperlinks for project outputs

www.cohubicol.com

www.journalcrcl.org

<https://publications.cohubicol.com/typology/>

<https://publications.cohubicol.com/vocabularies/cs/>

<https://publications.cohubicol.com/assets/uploads/cohubicol-research-study-on-text-driven-law-final.pdf>

https://www.cohubicol.com/assets/uploads/crcl23/research_study_cl_draft_15_nov_protected.pdf

COHUBICOL

Counting as a Human Being in the Era of Computational Law

Say cubicle ▸ Think Wittgenstein's cube

[Learn more](#)

**It would be nice if all of the data which sociologists require could be enumerated because then we could run them through IBM machines and draw charts as the economists do.
However, not everything that can be counted counts, and not everything that counts can be counted
– William Cameron, *Informal Sociology* (1963)**



About the Journal

The *Journal of Cross-disciplinary Research in Computational Law (CRCL)* invites excellence in law, computer science and other relevant disciplines with a focus on two types of 'legal technologies': (1) **data-driven** (e.g. predictive analytics, 'intelligent' search) and (2) **code-driven** (e.g. smart contracts, algorithmic decision-making (ADM), legal expert systems), and (3) their hybrids (e.g. **code-driven decision-making based on data-driven research**).

Legal practice is where computational law will be resisted, used or even fostered. CRCL wishes to raise questions as to (1) when the introduction of legal technologies should be resisted and on what grounds, (2) how and under what conditions they can be integrated into the practice of law and legal research and (3) how their integration may inform, erode or enhance legal protection and the rule of law.

Please [subscribe](#) for updates on upcoming articles and issues.

BROWSE

Categories

[Legal theory](#)

[Philosophy](#)

[Computer science](#)

CURRENT ISSUE

ATOM 1.0

RPT 1.0

RPT 1.0

USAGE STATISTICS INFORMATION

We log anonymous usage statistics.

Please read the [privacy information](#) for details.

CRCL23: Computational 'law' on edge

Photo by Evan Provan

CROSS-DISCIPLINARY RESEARCH IN
COMPUTATIONAL LAW

COMPUTATIONAL 'LAW' ON EDGE


The 2nd international conference organised by [COHUBICOL](#) in collaboration with [CRCL](#)

General Co-Chairs:

[Katie Atkinson](#), [Mireille Hildebrandt](#), [Frank Pasquale](#),
[Laurence Diver](#)

20 - 21 November 2023 in Brussels

Hybrid • attendance free of charge

[Registration now open](#) 

[Programme](#)

[Programme committee](#)

[Format](#)

[Call for abstracts](#)

[Conference ethos](#)

[Deadlines](#)

[Reviewers](#)

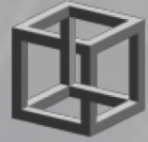


COHUBICOL

Research Study on Text-Driven Law

Laurence Diver, Tatiana Duarte, Gianmarco Gori,
Emilie van den Hoven and Mireille Hildebrandt

September 2023



COHUBICOL

Research Study on Computational Law

Pauline McBride and Laurence Diver

[DRAFT]

PROJECT PUBLICATIONS

[Home](#)

[Get in touch](#)

VOCABULARIES

[Legal vocabulary](#)

[Computer Science vocabulary](#)

RESEARCH STUDIES

TYPOLOGY OF LEGAL TECH

The Typology

[How to use](#)

[FAQs & methodology](#)


[Teaching & training resources](#)

[Outreach activities](#)

Typology of Legal Technologies

A Method – A Mindset

The Typology is a curated set of legal technologies (applications, scientific papers, and datasets) that we handpicked to demonstrate the potential impact on *legal effect* of different types of 'legal tech'. To understand how and why we created this, see the [FAQs & methodology](#) page.

- **Use the filters below** to find legal techs you are interested in. Click a system to view its full profile.
- **Compare systems** by clicking  on one or more systems (view the comparison at the bottom of this page).

SHOWING 30 TECHS

[RESET FILTERS](#)

END-USERS	FUNCTIONALITY	CODE/DATA-DRIVEN	TYPE OF SYSTEM	
Any	Any	Either	<input checked="" type="radio"/> Any	<input type="radio"/> App <input type="radio"/> Dataset <input type="radio"/> Paper
Akoma Ntoso	Automatic Catchphrase Identification from Legal Court Case Documents (Mandal et al. 2017)	Blawx		
Legislation Search	Litigation Search	Legislation		
Casetext	Catala	Chinese AI and Law dataset (CAIL2018)		
Litigation Search	ADM Legislation	Litigation		
Contract Understanding Atticus Dataset (CUAD)	DataLex	Della		
Contract	Legal Contract	Contract		

What's Next?

- AI for Cyberjustice and the Digital Transformation
 - What are we talking about
- A Typology of Legal Technologies
 - What, Why, How, by Whom and for Whom
- Question Zero
 - Think before you invest
- Counting as a human being
 - Not everything that can be counted counts,
 - Not everything that counts can be counted

What's Next?

- AI for Cyberjustice and the Digital Transformation
 - What are we talking about
- A Typology of Legal Technologies
 - What, Why, How, by Whom and for Whom
- Question Zero
 - Think before you invest
- Counting as a human being
 - Not everything that can be counted counts
 - Not everything that counts can be counted

AI for Cyberjustice the digital transformation

- What is AI in this context?
 - Generative AI (LLM)?
 - Rule-based systems (ADM)?
 - Predictive AI (recidivism, fraud detection in tax, social benefits, immigration)?
- Where would Judicial e-Auctions fit?

AI for Cyberjustice the digital transformation

EP version of the AI Act, staying close to the OECD definition:

Art. 3(1) “artificial intelligence system’ (AI system) means

- a machine-based system
- that is designed to operate with varying levels of autonomy and
- that can, for explicit or implicit objectives,
- generate outputs such as predictions, recommendations, or decisions,
- that influence physical or virtual environments;”

AI for **Cyberjustice** the digital transformation

- What is cyberjustice in this context?
 - Justice in cyberspace?
 - The digital transformation of the administration of justice?
 - the judiciary, the public prosecutor's office, private law enforcement procedures (e.g. the judicial e-auction), public administration (e.g. tax decisions), the legislature
 - Justice in the context of the digital transformation of the administration of justice?

AI for Cyberjustice

the digital transformation

- What is the digital transformation of the administration of justice?
 - Electronic instead of paper files?
 - ADM?
 - Digital-ready legislation (EU LEOS project)?

What's Next?

- AI for Cyberjustice and the Digital Transformation
 - What are we talking about
- A Typology of Legal Technologies
 - What, Why, How, by Whom and for Whom
- Question Zero
 - Think before you invest
- Counting as a human being
 - Not everything that can be counted counts
 - Not everything that counts can be counted

<https://www.cohubicol.com>

PROJECT PUBLICATIONS

[Home](#)

[Get in touch](#)

VOCABULARIES

RESEARCH STUDIES

TYPOLOGY OF LEGAL TECH

The Typology

[How to use](#)

[FAQs & methodology](#)


[Teaching & training resources](#)

[Outreach activities](#)

Typology of Legal Technologies

A Method – A Mindset

The Typology is a curated set of legal technologies (applications, scientific papers, and datasets) that we handpicked to demonstrate the potential impact on *legal effect* of different types of 'legal tech'. To understand how and why we created this, see the [FAQs & methodology](#) page.

- **Use the filters below** to find legal techs you are interested in. Click a system to view its full profile.
- **Compare systems** by clicking  on one or more systems (view the comparison at the bottom of this page).

SHOWING 30 TECHS

[RESET FILTERS](#)

END-USERS	FUNCTIONALITY	CODE/DATA-DRIVEN	TYPE OF SYSTEM	
Any	Any	Either	<input checked="" type="radio"/> Any	<input type="radio"/> App <input type="radio"/> Dataset <input type="radio"/> Paper
Akoma Ntoso	Automatic Catchphrase Identification from Legal Court Case Documents (Mandal et al. 2017)	Blawx	Legislation	Search
Casetext	Catala	Chinese AI and Law dataset (CAIL2018)	Litigation	Search
Contract Understanding Atticus Dataset (CUAD)	DataLex	Della	Contract	Search
			ADM	Legislation
			LES	Contract
			Legislation	Compliance
			Contract	Contract

A Typology of Legal Technologies

- What?
 - A hand-picked, curated set of 30 tokens (legal techs)
 - That **typify** various types/categories of legal tech (e.g. data- and/or code-driven)
- Not a taxonomy
 - No attempt to be complete in terms of tokens
 - No attempt to define mutually exclusive categories
 - Arguably reasonably comprehensive in terms of types

A Typology of Legal Technologies

- What?
 - A resource
 - A method
 - A mindset
- See our research blogs to see how we 'apply' the methodology to new/other legal techs
 - 'Juridische Tracker' : <https://www.cohubicol.com/blog/caselaw-revisited-rechtln-case-law-tracker-assessed-with-the-typology-of-legal-technologies/>
 - 'Casetext's Co-Council' : <https://www.cohubicol.com/blog/casetext-cocounsel-openai-typology/>

Caselaw revisited: Recht.nl's case law tracker assessed with the Typology of Legal Technologies

🕒 10 minute read

Photo by John So

[Home](#) / [Blog](#) / Caselaw revisited: Recht.nl's case law tracker assessed with the Typology of Legal Technologies



Emilie van den Hoven
PhD candidate (law)

✉ Email

🐦 Twitter

23/11/2024

In a [recent news item](#), the important Dutch legal knowledge platform Recht.nl announced that they have launched a new case law tracker ('Jurisprudentie Tracker').

Until recently, the platform focused on providing legal news and reporting developments of note across legal domains, gathering and mapping relevant legal scholarship in those domains, and functioned as a go-to place for legal vacancies across sectors. Recht.nl claims to have 20,000 subscribers and 4.5 million pageviews per year. Many can access the service for free with an institutional subscription, for instance through a university licence. A personal subscription currently costs €125 per year. Jurisprudentie Tracker is available to all subscribers, providing a quick and easy overview of published case law in the

Casetext's CoCounsel through the lens of the Typology

🕒 14 minute read

[Home](#) / [Blog](#) / **Casetext's CoCounsel through the lens of the Typology**



Pauline McBride

Postdoctoral researcher (law)

✉ Email


🐦 Twitter

Everyone is talking to OpenAI's ChatGPT.¹ Meanwhile, OpenAI has been talking to developers of legal technologies.² In March, Casetext announced a collaboration with OpenAI and launched CoCounsel, Casetext's latest product offering.³ CoCounsel is built on GPT-4,⁴ OpenAI's most advanced [Generative Pre-trained Transformer](#) to date and the first [large language model](#) to be credited with a passing score in a simulated bar exam.⁵ Casetext describes CoCounsel as 'groundbreaking' and '[t]he legal AI you've been waiting for'.⁶ It invites lawyers to 'delegate substantive work to ... [CoCounsel] and trust the results.'⁷ Should they?

Casetext is no rookie start-up. The US-based company, which has just entered into an agreement to be acquired by Thomson Reuters for \$650m,⁸ was an early adopter of advanced machine learning technologies. Its product offerings include Casetext Research, now simply 'Research' (a legal research

A Typology of Legal Technologies





- Why?
 - Objectives:
 - Mapping
 - Comparing
 - Assessing

- **Use the filters below** to find legal techs you are interested in. Click a system to view its full profile.
- **Compare systems** by clicking  on one or more systems (view the comparison at the bottom of this page).

SHOWING 4 TECHS




 RESET FILTERS

END-USERS	FUNCTIONALITY	CODE/DATA-DRIVEN	TYPE OF SYSTEM	
<input type="text" value="Any"/>	<input type="text" value="Any"/>	<input type="text" value="Either"/>	<input type="radio"/> Any	<input checked="" type="radio"/> App
			<input checked="" type="radio"/> Dataset	<input type="radio"/> Paper

Chinese AI and Law dataset (CAIL2018) Litigation  <input checked="" type="checkbox"/>	Contract Understanding Atticus Dataset (CUAD) Contract  <input checked="" type="checkbox"/>	LexGLUE Search  <input checked="" type="checkbox"/>
Statutory Article Retrieval Dataset (BSARD) Search  <input checked="" type="checkbox"/>		

Compare legal techs

 RESET COMPARISON

	Statutory Article Retrieval Dataset (BSARD) 	LexGLUE 	Chinese AI and Law dataset (CAIL2018) 
Intended users	<ul style="list-style-type: none"> ▪ Natural persons ▪ Software developers 	<ul style="list-style-type: none"> ▪ Academics ▪ Software developers 	<ul style="list-style-type: none"> ▪ Academics ▪ Software developers
Code- or data-driven			
Form	<ul style="list-style-type: none"> ▪ Dataset (off-the-shelf) 	<ul style="list-style-type: none"> ▪ Dataset (off-the-shelf) 	<ul style="list-style-type: none"> ▪ Dataset (off-the-shelf)
Automation or support	<ul style="list-style-type: none"> ▪ Legal research strategy 	<ul style="list-style-type: none"> ▪ Legal decision support ▪ Legal research strategy ▪ Legal strategy support 	<ul style="list-style-type: none"> ▪ Legal decision support ▪ Legal research strategy ▪ Legal strategy support
In use?	Unknown	In current use	Unknown
Creators	<ul style="list-style-type: none"> ▪ Academics 	<ul style="list-style-type: none"> ▪ Academics 	<ul style="list-style-type: none"> ▪ Academics
Access	<ul style="list-style-type: none"> ▪ Free download/web application 	<ul style="list-style-type: none"> ▪ Free download/web application 	<ul style="list-style-type: none"> ▪ Free download/web application

See our [methodology](#) for field definitions.

A Typology of Legal Technologies

■ Why?

- Setting the stage for **further research** into legal technologies
- Offering a strategy **to evaluate** different types of legal tech
- Providing a methodology **to compare** legal techs
- Ensuring that both lawyers and developers can **navigate and understand**

A Typology of Legal Technologies

- How?
 - Mapping in terms of:
 - Papers, datasets, applications
 - Claimed functionality
 - Background of the developers (law, CS; jurisdiction)
 - Claimed target jurisdiction
 - Access (licence, open source)

Akoma Ntoso

Legislation: drafting Legislation: representation Search: case law Search: legislation

akomantoso.org

Main research: March 2022

CONTENTS

- [What does it claim to do?](#)
- [Substantiation of claims & potential issues](#)
- [How might the end-user assess effectiveness?](#)
- [What form does it take?](#)
- [Is it currently in use?](#)
- [The creators](#)
- [Jurisdiction](#)
- [License](#)

▼ What does it claim to do?

Akoma Ntoso (AKN) is an open standard [markup language](#) that defines a method for 'tagging' the structure and content

of legal documents such that they are machine readable. The goal is to enable other systems to process such

documents in more sophisticated ways than is possible with standard word processor files.

Claimed essential features

- Makes the structure and meaning of legal documents machine-readable.
- Facilitates interchange of documents across institutions and jurisdictions.
- Allows precise citation and cross-referencing of documents.
- Allows identification of the content of the law at a given point in time.

▶ [RELEVANT QUOTES](#)

▼ AT A GLANCE ?	
Intended users	<ul style="list-style-type: none">▪ Legislators▪ Public administration▪ Software developers
Code- or data-driven	Code-driven <ul style="list-style-type: none">▪ Rules as Code
Form	Component (combination) Details
Automation or support	–
In use?	In current use
Creators	Academics Details
Access	<ul style="list-style-type: none">▪ Free download/web application

See our [methodology](#) for field definitions.

A Typology of Legal Technologies

- By whom:

- L. Diver, P. McBride, **M. Medvedeva**, A. Banerjee, **E. D'hondt**, T. Duarte, D. Dushi, G. Gori, E. van den Hoven, **P. Meessen**, M. Hildebrandt, 'Typology of Legal Technologies' (COHUBICOL, 2022), available at <https://publications.cohubicol.com/typology>
 - In-depth collaboration of lawyers and computer scientists

A Typology of Legal Technologies

- For whom (e.g.):
 - Practicing lawyers
 - Bar Associations
 - Legislatures
 - Judiciary
 - Developers
 - Legal researchers

A Typology of Legal Technologies

The main investigation concerns:

- **claims made on behalf of** AI systems in law
- **the substantiation of such claims**
 - Mathematical verification, empirical validation
 - Impact on the domain: gaps between requirements and specifications
 - Real-world impact (gap between specification and real-world goal)

What's Next?

- AI for Cyberjustice and the Digital Transformation
 - What are we talking about
- A Typology of Legal Technologies
 - What, Why, How, by Whom and for Whom
- Question Zero
 - Think before you invest
- Counting as a human being
 - Not everything that can be counted counts
 - Not everything that counts can be counted

Question 0

To deploy AI in law, we need to make certain assumptions about the computability of the law.

- For data-driven law (using machine learning) we probably need to
 - **frame law as a corpus of legal training data**
- For code-driven law (using programming languages to 'render' the law) we need to
 - **frame law as a closed system of rules or algorithms**

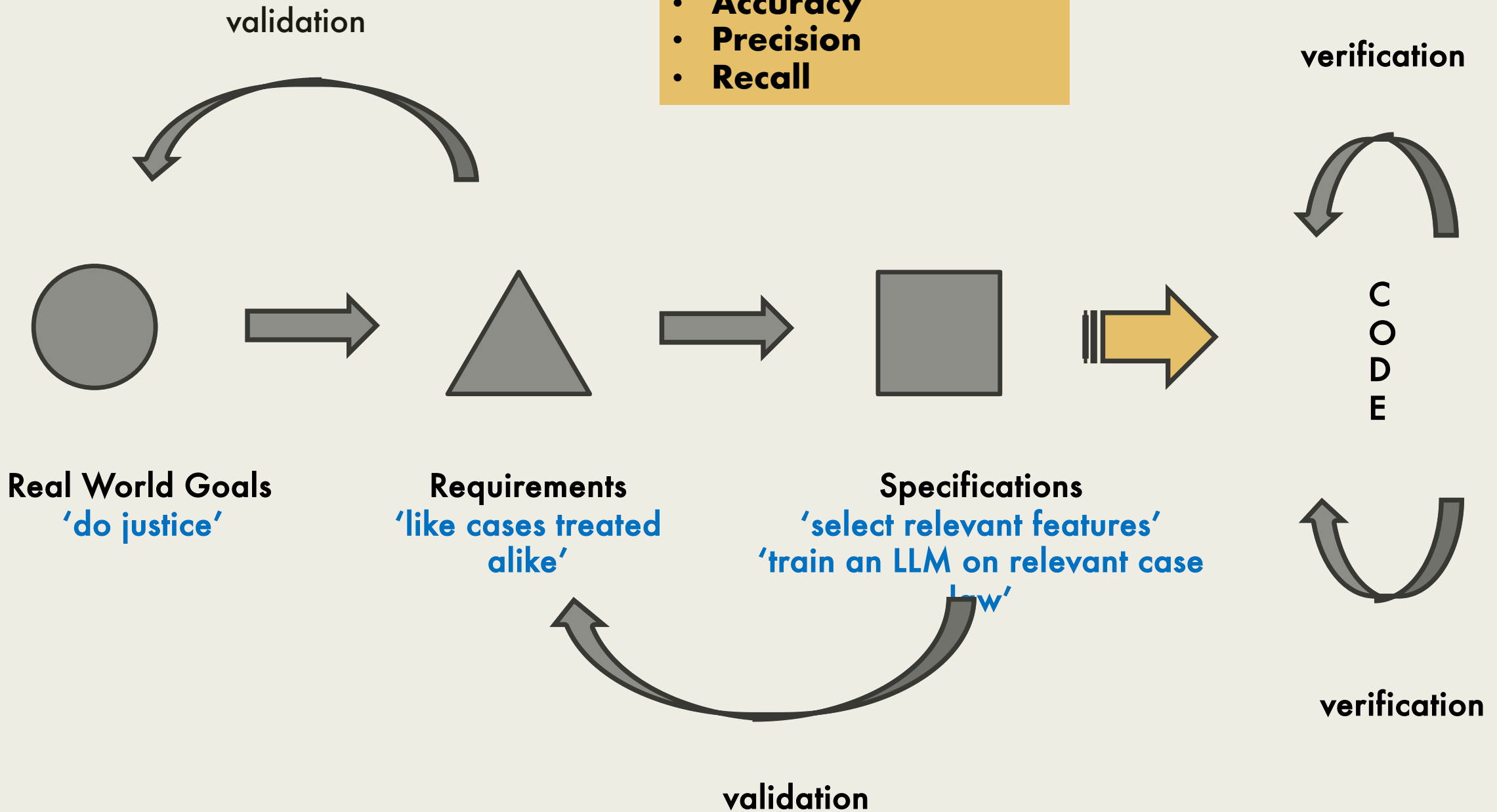
Question Zero

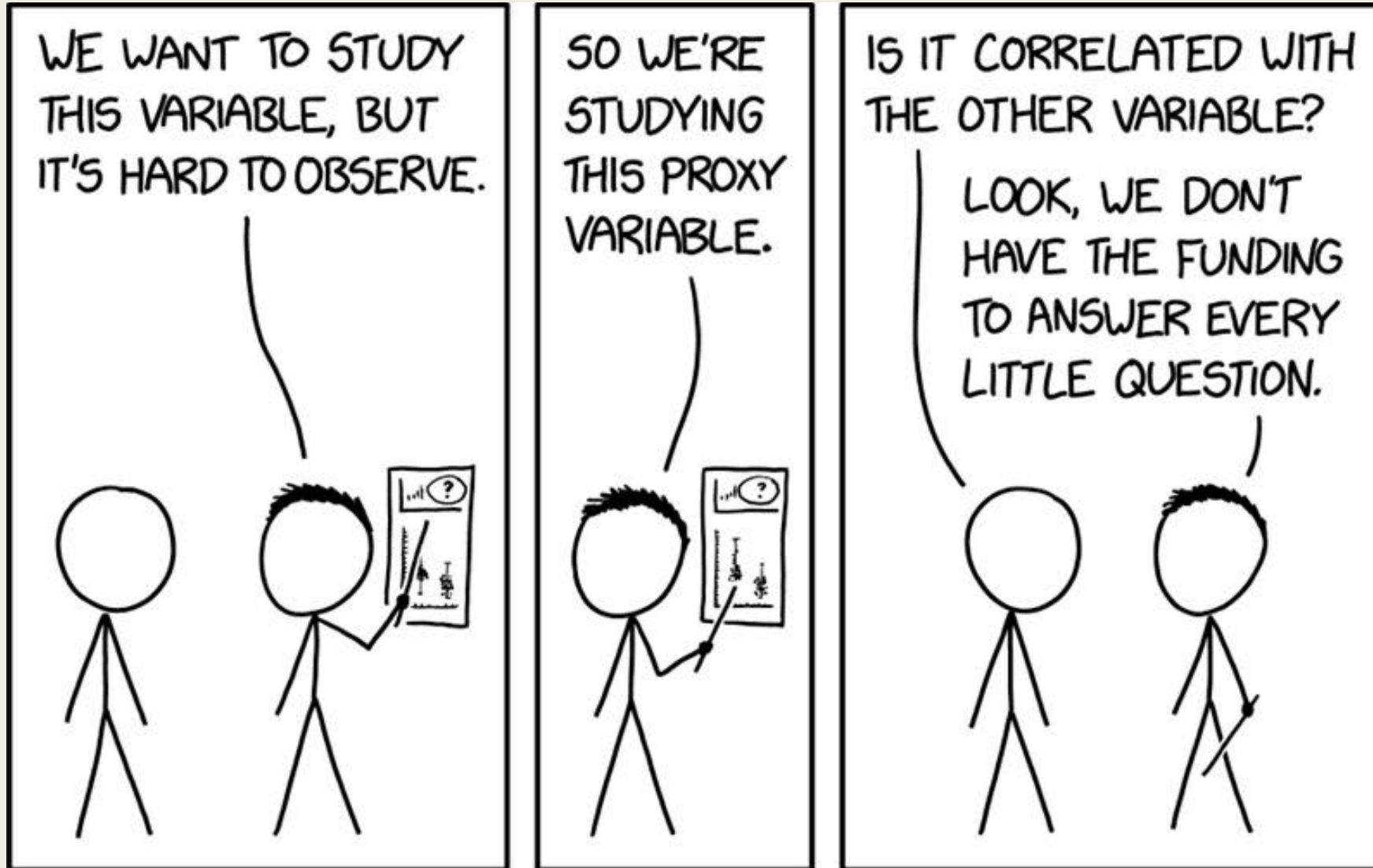
- What matters is not computable
- It can, however, be made computable
- This can always be done in different ways
- And those differences matter

- Key issues here are the selection and construction of the proxies
 - Training data, feature selection, hypothesis space, goals
- And, in the case of RLHF, the prompts provided to achieve alignment

ML output-testing:

- Accuracy
- Precision
- Recall







Before deciding on a risk/benefits analysis, always first ask:

- 1. What problem(s) does this legal tech solve?**
- 2. What problem(s) does it NOT solve?**
- 3. What problem(s) does it create, in the longer run?**

This – of course – concerns ‘real world’ problems

some hyperlinks to key papers on the problems with predictive AI

- https://www.cohubicol.com/assets/uploads/crcl23/kapoor_henderson_narayan_anan_position_paper_crcl23.pdf
- <https://link.springer.com/article/10.1007/s10506-021-09306-3>
- https://masha-medvedeva.github.io/papers/MedvedevaMcBride_LegalJudgmentPrediction.pdf

What's Next?

- AI for Cyberjustice and the Digital Transformation
 - What are we talking about
- A Typology of Legal Technologies
 - What, Why, How, by Whom and for Whom
- Question Zero
 - Think before you invest
- Counting as a human being
 - Not everything that can be counted counts,
 - Not everything that counts can be counted

Down to earth integration in legal research and practice

Legal problems cannot be solved by computational technologies

- Legal problems concern 'matters of law':
 - What complex legal norm applies?
 - How should it be interpreted with an eye to the case at hand?
 - What facts are relevant?
 - How should they be interpreted and qualified in light of the legal norm?
- This involves **speech acts** that require **judgement** rather than **logic or statistics**

Down to earth integration in legal research and practice

Legal problems cannot be solved by computational technologies

- Digital technologies may contribute to **the administration of justice**:
 - Logistics (the pipe line of court documents)
 - Access to the relevant documents (parties)
 - Remote testimony, hearing, pleading?
 - Access to judgments
 - those who share jurisdiction (internal transparency)
 - 'the people' (external transparency)

Down to earth integration in legal research and practice

Legal problems cannot be solved by computational technologies

- Software development may contribute to **the making of law**:
 - Domain specific programming languages
 - Revealing the issues of deliberate and the role of unintended ambiguity
 - Logic programming
 - Revealing the issue of the law's dedication to double negations
 - Beware of scaling the past and freezing the future

Down to earth integration in legal research and practice

Legal problems cannot be solved by computational technologies

- **Distant reading technologies** may contribute to legal research and practice:
 - Legal search (case law, legislation, doctrine, principles, custom)
 - but note the key decisions this requires:
 - adding metadata, working with preconceived templates
 - labelling, defining relevance
 - Beware of the proxies and the legal interpretation they imply
 - Distant reading should prepare for close reading

Down to earth integration in legal research and practice

Legal problems cannot be solved by computational technologies

- **Distant reading technologies** may contribute to legal research and practice:
 - Argumentation mining and legal reasoning
 - As input for qualitative research (close reading)

Down to earth integration in legal research and practice

Legal problems cannot be solved by computational technologies

- **Generative technologies** may contribute to legal research and practice (RLHF):
 - Depending on their further development
 - To detect their provenance (art. 28b and 52 AI Act)
 - To prevent false references (art. 28b AI Act)
 - To enable checking the reliability (art. 28b AI Act)
 - Note the many unresolved issues of alignment
 - And the many security issues with prompting



m.e.menair