JURISDICTION IDENTIFIERS

MANAGING MULTINATIONAL RESOURCES

FRANK BENNETT

LAW VIA THE INTERNET
RUTGERS UNIVERSITY, NEWARK
2017-10-21

TABLE OF CONTENTS

- In Brief
- Okay: What I Really Want
- Identifiers
- What does comparative law *do*?
- Legal Resource Registry
- Moving forward
- Thank You

IN BRIEF

Herein of the problem and what I have to say about it.

Like it says, a jurisdiction identifier is a form of metadata.

IN BRIEF

- Legal jurisdictions maintain separate electronic collections.
- Electronic collections need machine-readable identifiers.
- Multiple jurisdictions? You need *jurisdiction* identifiers.
- So we should expect them to exist.
- Because globalization.
- But they don't, although …

The logic here is pretty straightforward.

IN BRIEF

- … building a set of jurisdiction identifiers is a *relatively* simple thing to do.
- So we should.
- (coffee!)

There is more to the story, of course, but this is where it will lead us.

OKAY: WHAT I REALLY WANT

Herein of the effect of research habits on residual artifacts of scholarship.

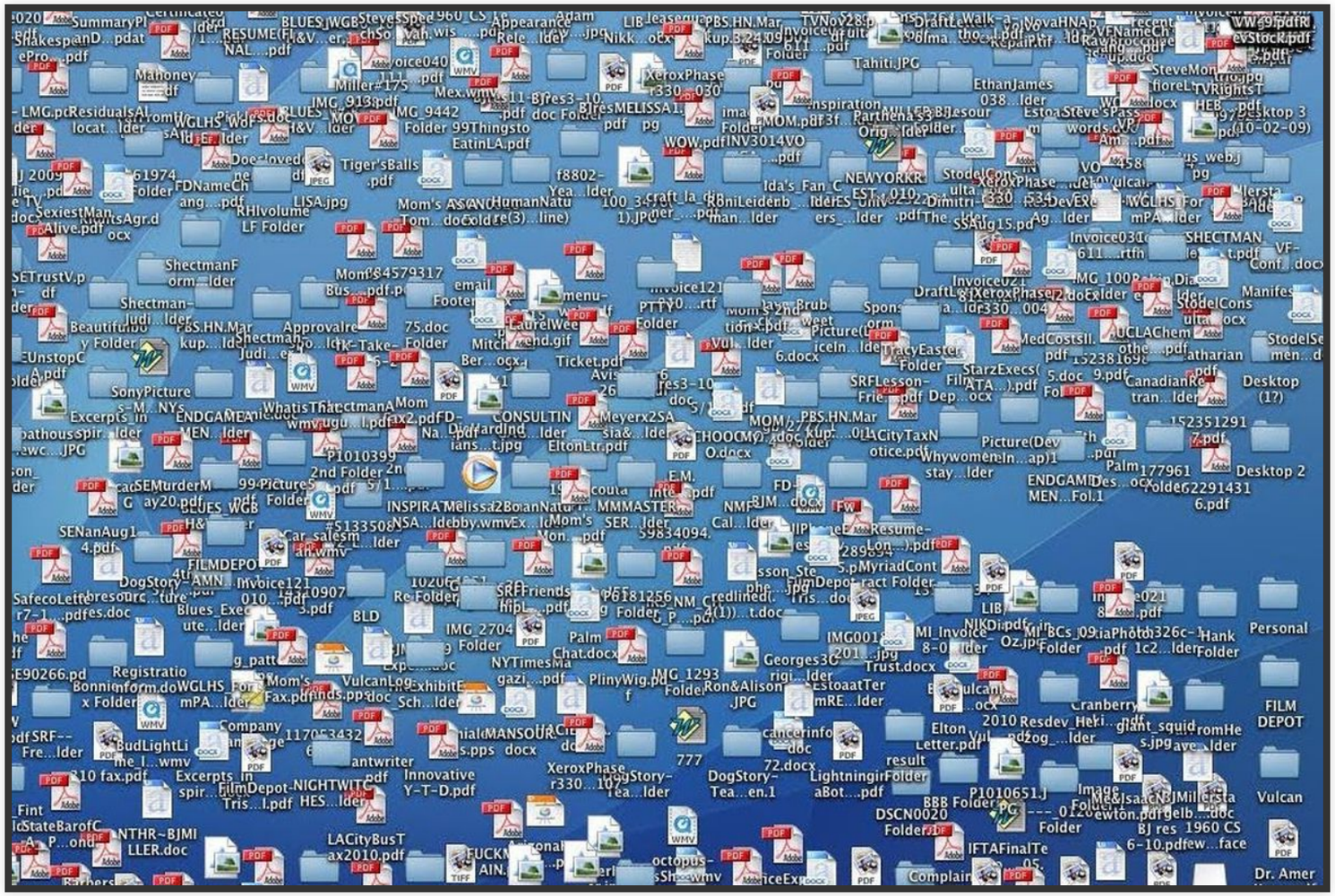
Dammit Jim, I'm a comparative lawyer, not an IT specialist!

Let's think skeptically for a sec about the impact of technology on our scholarship.

DATA AND LEGAL RESEARCH

- Where can I find stuff online?
- What resources are easiest to access?
- How can I quantify things?
- Can I bring large amounts of data into play?
- Search
- Lotsa PDF

In legal research as elsewhere, the things that can be done more easily tend to be done more often and to receive more emphasis.



Remnants of research past

When all you have is a digital hammer, things can easily end up looking like this behind the scenes.

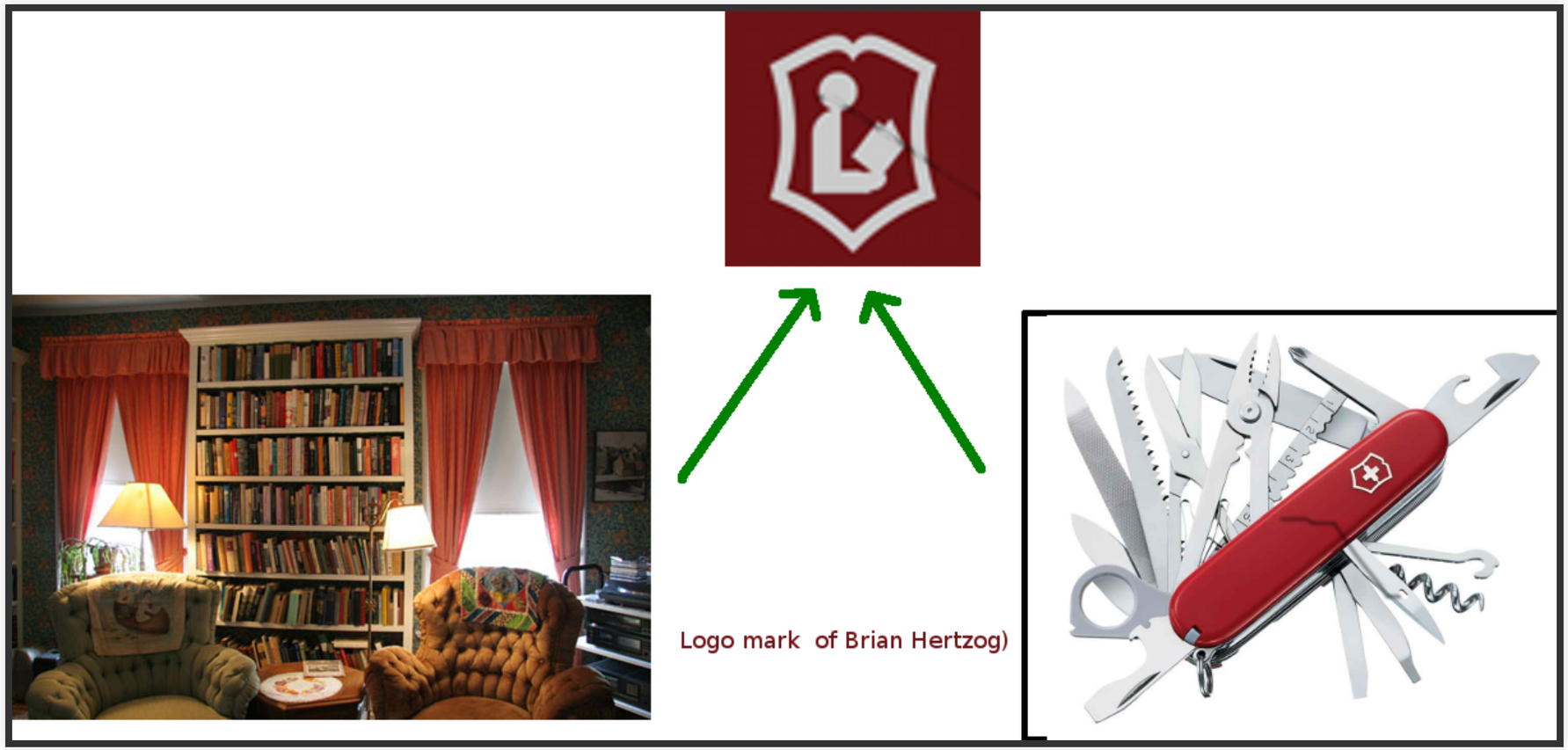
OLD-SCHOOL COMPARATIVE LAW

- A focus on specific primary sources
- Multiple national jurisdictions
- Multiple languages
- Each project a bespoke voyage of discovery

In times of higher search cost, credible comparisons of phenomena across jurisdictions began *of necessity* with systematic study of structures within the respective systems. Commentaries, outlines, ontologies, indexes: these were the raw materials that the scholar fashioned into a narrative emerging from comparison.

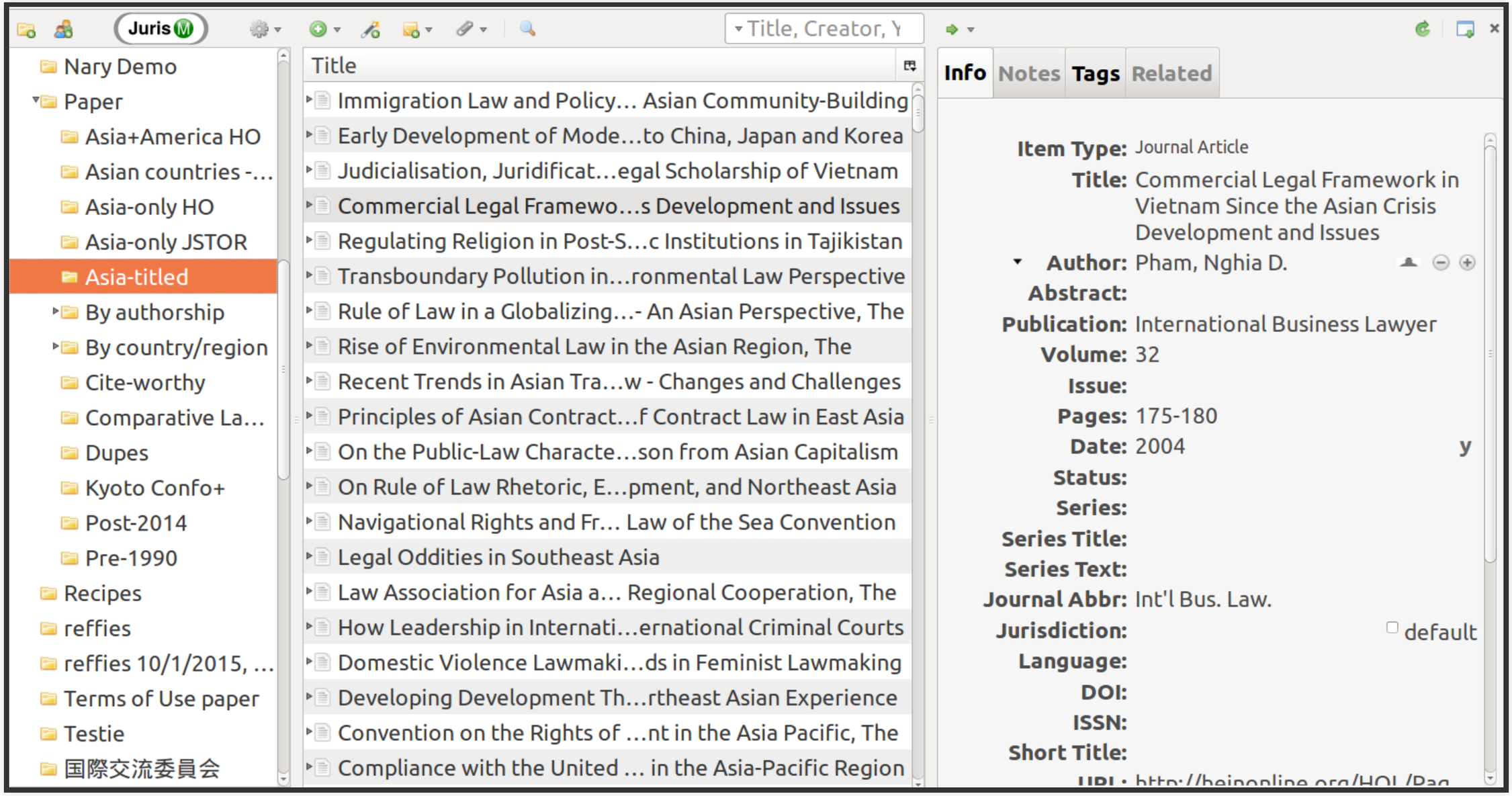
Research memories

Research habits vary, but paper-based research methods lent themselves to the production of well organized personal document collections over the course of an academic career. (Overstuffed chairs and velvet curtains are an optional extra.)



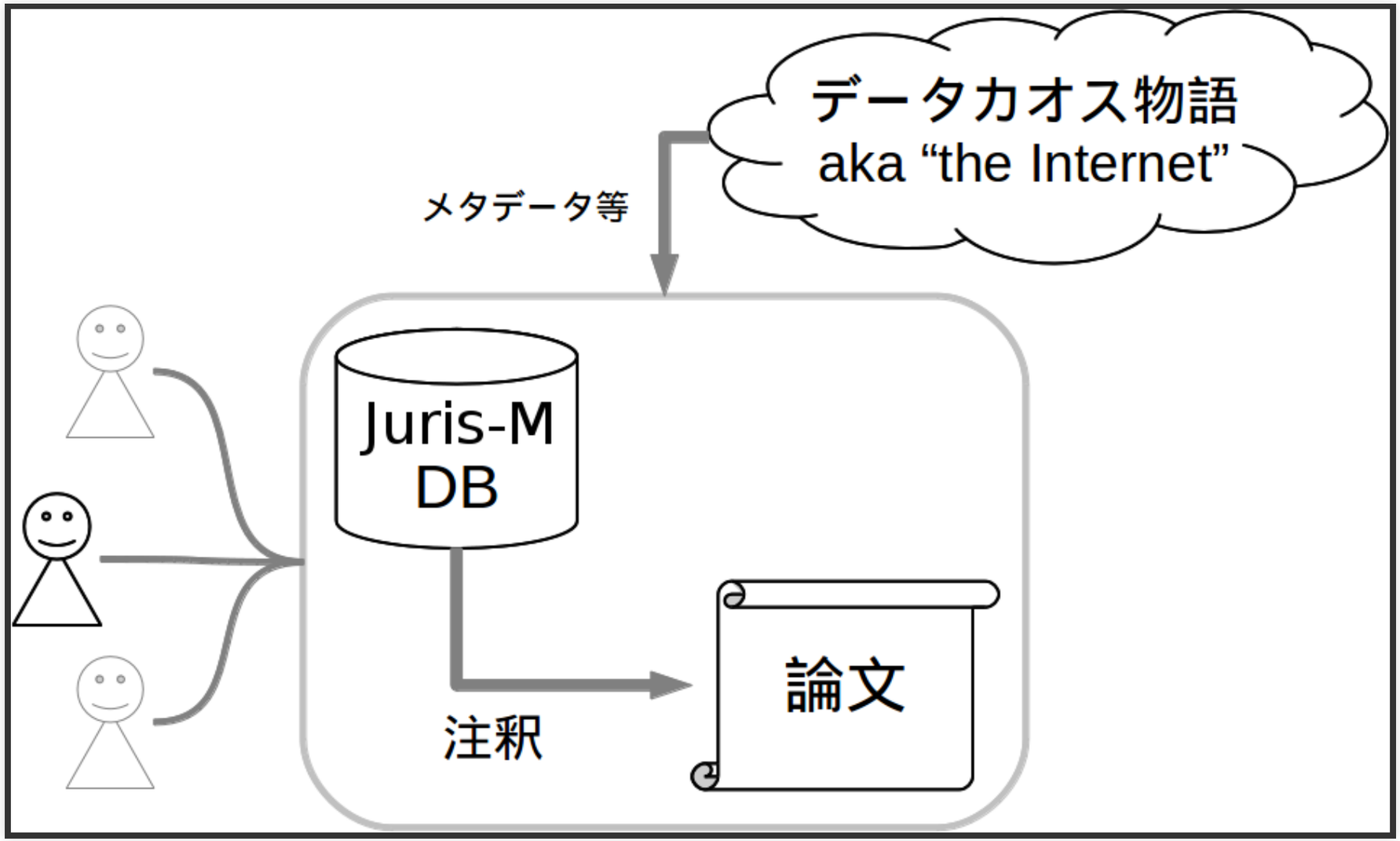
A Legal Research Multi-Tool

In an ideal world, the power of digital search and text analysis would supplement tools for managing personal libraries. This is a trend in "digital humanities" that has been late to reach legal and multilingual scholarship.



Juris-M Reference Manager

Juris-M is a variant of the Zotero reference manager built to fill this gap in the scholar's toolchest.



Housekeeping

While a body of law is a highly structured body of instrumental declarations, their publication is often an ad hoc affair. A reference manager provides a means of properly organizing documents gleaned from the "Tale of Data Chaos (aka the Internet)" for research purposes.

IDENTIFIERS

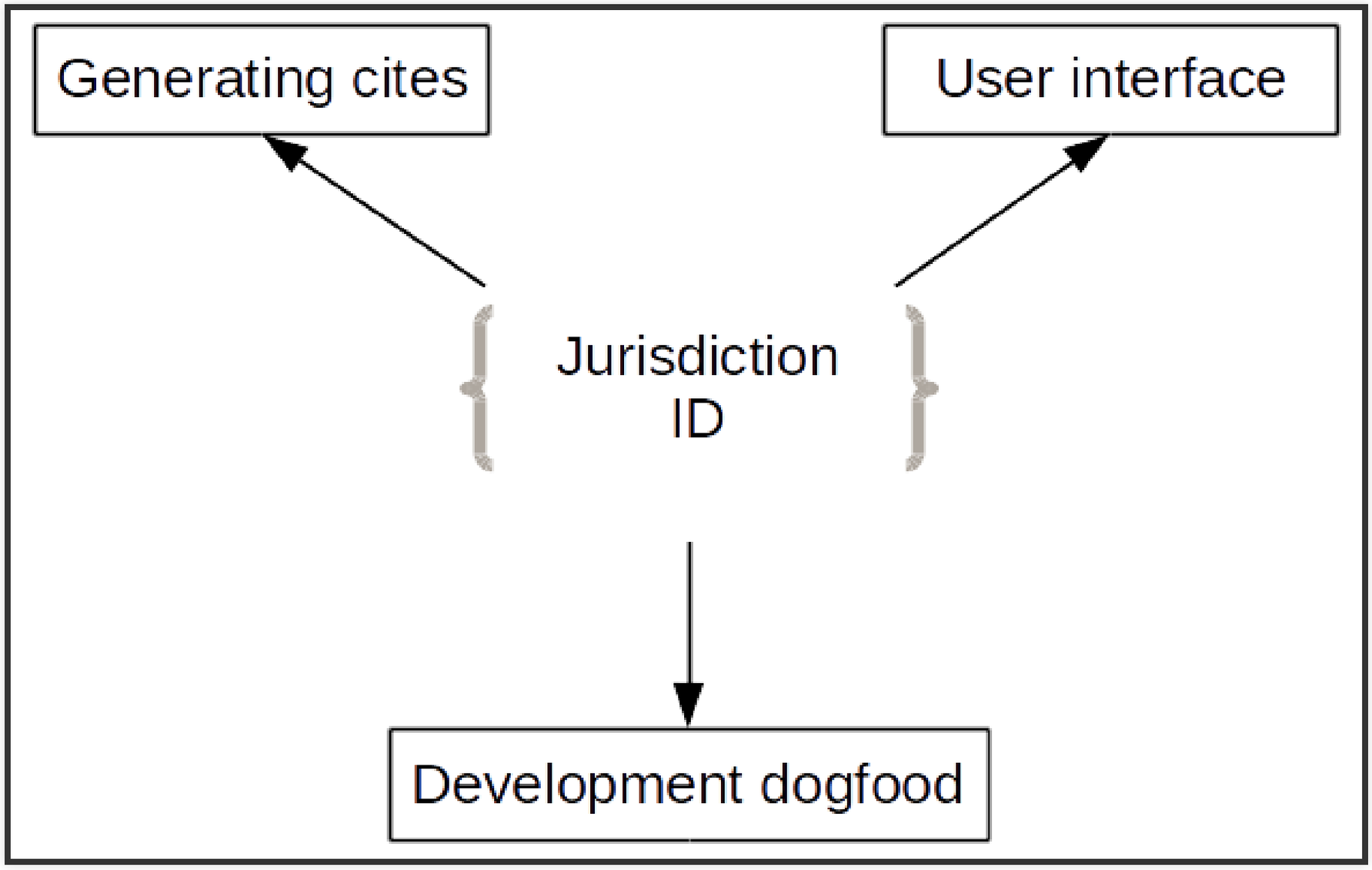
As noted at the outset, to build that ideal research platform, we're going to need identifiers.

THINGS COME IN FLAVORS

- Developers
  - Repository wranglers
  - Antique collectors
- Identifiers
  - Document identifiers
  - Jurisdiction identifiers

The complexity of an identifier system grows in proportion to its proximity to real-world artifacts.

Large-scale legal archives need unique document identifiers to manage collections under their respective umbrellas. For ease of maintenance, a lot of thought goes into crafting expressive identifiers that make sense to humans—and jurisdiction is often implicit. In a "little-data" library of eclectic resources, expressive identifiers for jurisdiction are critical, while documents can satisfactorily be identified by arbitrary slugs.



Assuming a jurisdiction identifier

Jurisdiction identifiers are useful in at least three ways.

GENERATING CITES

United States

Henley v. DeVore, 733 F. Supp. 2d 1144 (2010).

England & Wales

Rhone v. Stephens, [1994] 2 AC 310 (HL).

Japan

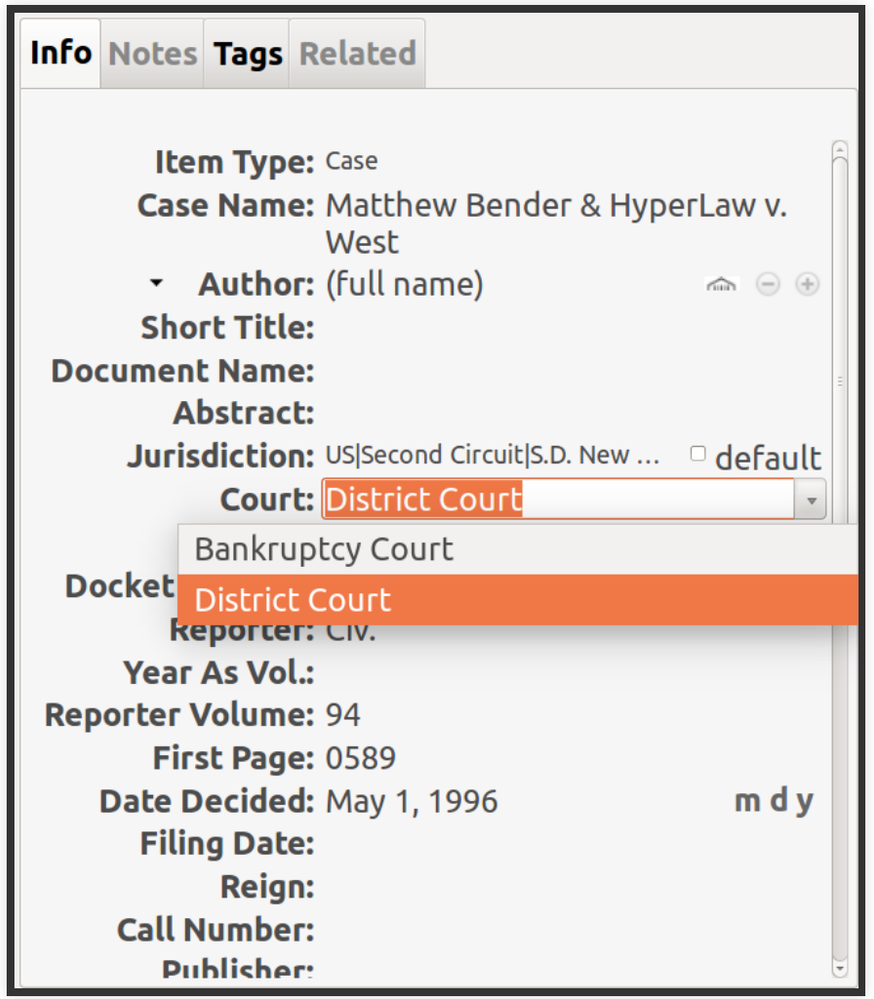
Tokyo High Court, (ne) 4593, judgment, October 4, 2007.

In cites to secondary sources, the citation format is a property of the "parent" style (Chicago, Vancouver, APA, Indigo).

But in cites to primary legal sources, the citation format is a property of the *jurisdiction*.

This adds a layer of complexity to the design of stylesheet schemata for generating human-readable references—more than once, citation technlogists have memorably characterized legal citation practices as "insane." However, the complexity is unavoidable in our time, given that legal systems are heavily invested in specific shorthand reference forms, and that citations often foreground features of the underlying archives that are unique to the jurisdiction. As a practical matter, these differences must be respected, and jurisdiction identifiers are the key to doing so.

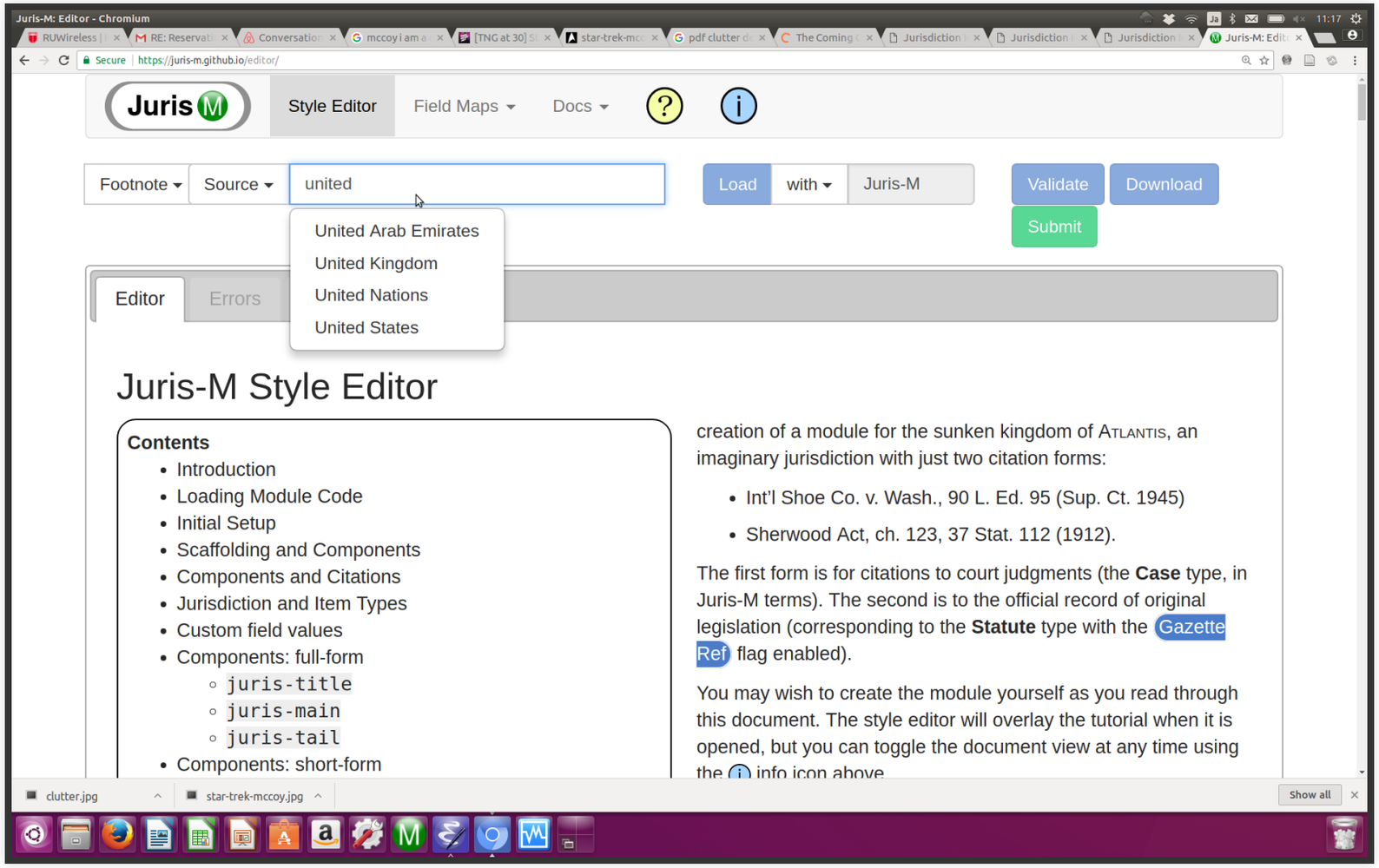
USER INTERFACE



Typeahead search of controlled lists

When machine-readable jurisdiction identifiers are known, they can be leveraged to fashion user-facing interface helpers that make for ease of use and uniformity in the underlying metadata stored in the personal library.

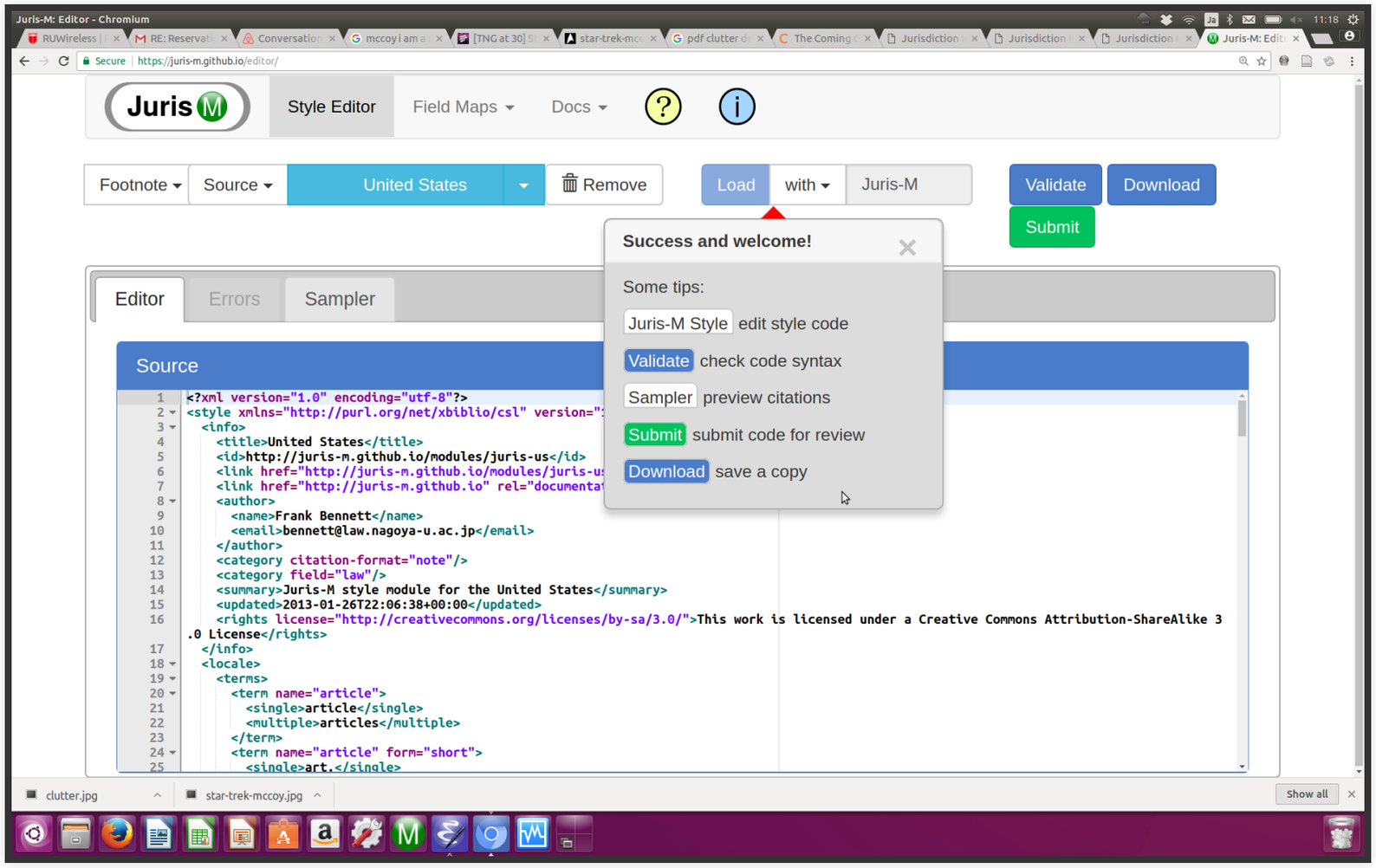
DOGFOOD FOR DEVELOPMENT (1)



Juris-M legal style editor

The Juris-M project "eats its own [identifier] dogfood." Stylesheet support for the plethora of local citation practices demands an expanding set of jurisdiction-specific legal citation modules. The Juris-M Style Editor (built atop GitHub Pages), supports per-jurisdiction selection of supplementary stylesheets for editing.

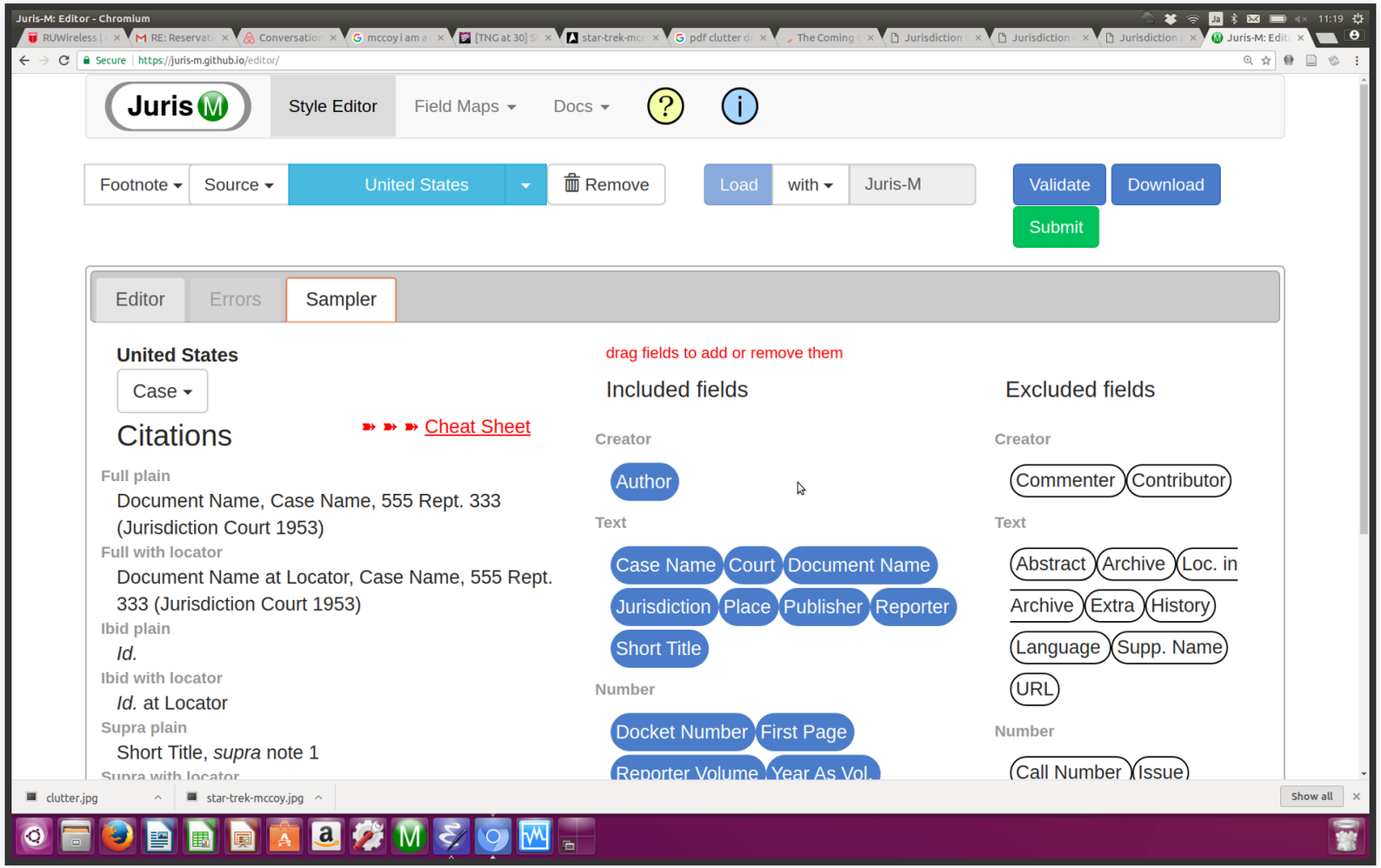
DOGFOOD FOR DEVELOPMENT (2)



Juris-M legal style editor

Loading a jurisdiction to the editor brings up its XML code.

DOGFOOD FOR DEVELOPMENT (3)



Juris-M legal style editor

Edits to the stylesheet XML can be validated through the editor, and the behavior of validated code can be explored by drag-and-drop removal, addition, and editing of item metadata input.

WHAT DOES COMPARATIVE LAW *DO*?

Herein of a preliminary empirical study of comparative law discourse in Germany and Asian countries in transition, presented at the Asian Law and Society Association conference, August 4, 2015.

In HeinOnline and JSTOR, for the years 1980 to 2014 …

```
california* AND (law* OR legal*)
```

```
german* AND (law* OR legal*)
```

```
japan* AND (law* OR legal*)
```

```
(cambodia* OR kazakh* OR lao* OR mongolia* OR tajik* OR uzbek* OR viet*) AND (law* OR legal
```

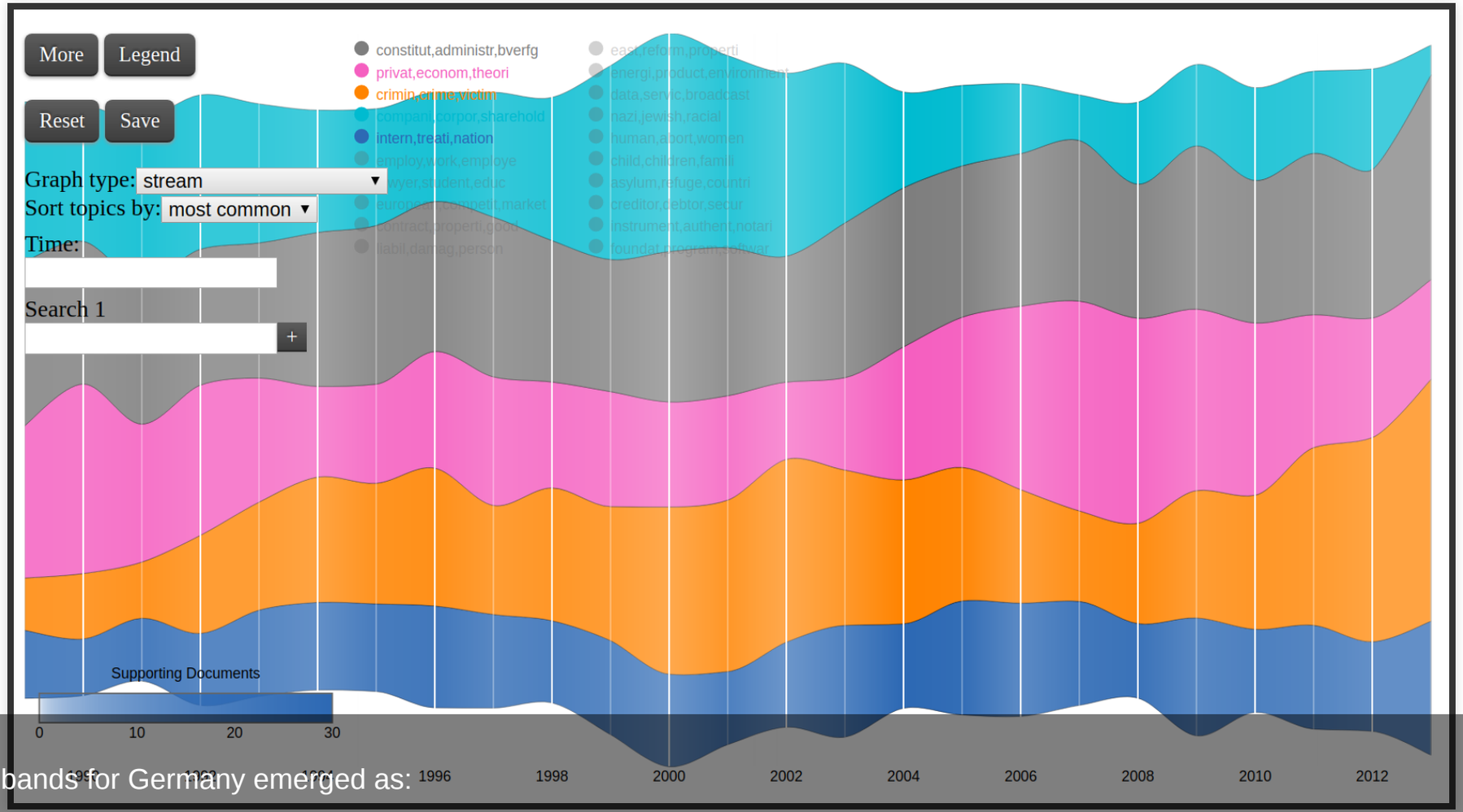
The study was based on English-language law review articles concerning several jurisdictions. We will focus here on the results from Germany and a subset of Asian countries in transition.

DATA COLLECTED

Category	HO LJL
California	317
Germany	412
Japan	396
Other Asia	201

Note: “Other Asia” includes nine articles from JSTOR.

The aim was to explore the topic patterns in collected documents via Latent Dirichlet Allocation, using MALLET and associated visualization tools. For stable results, this method relies on large document sets of 1,000 items or more. It is a point of weakness in this study that the volume of documents collected for each target was much smaller.

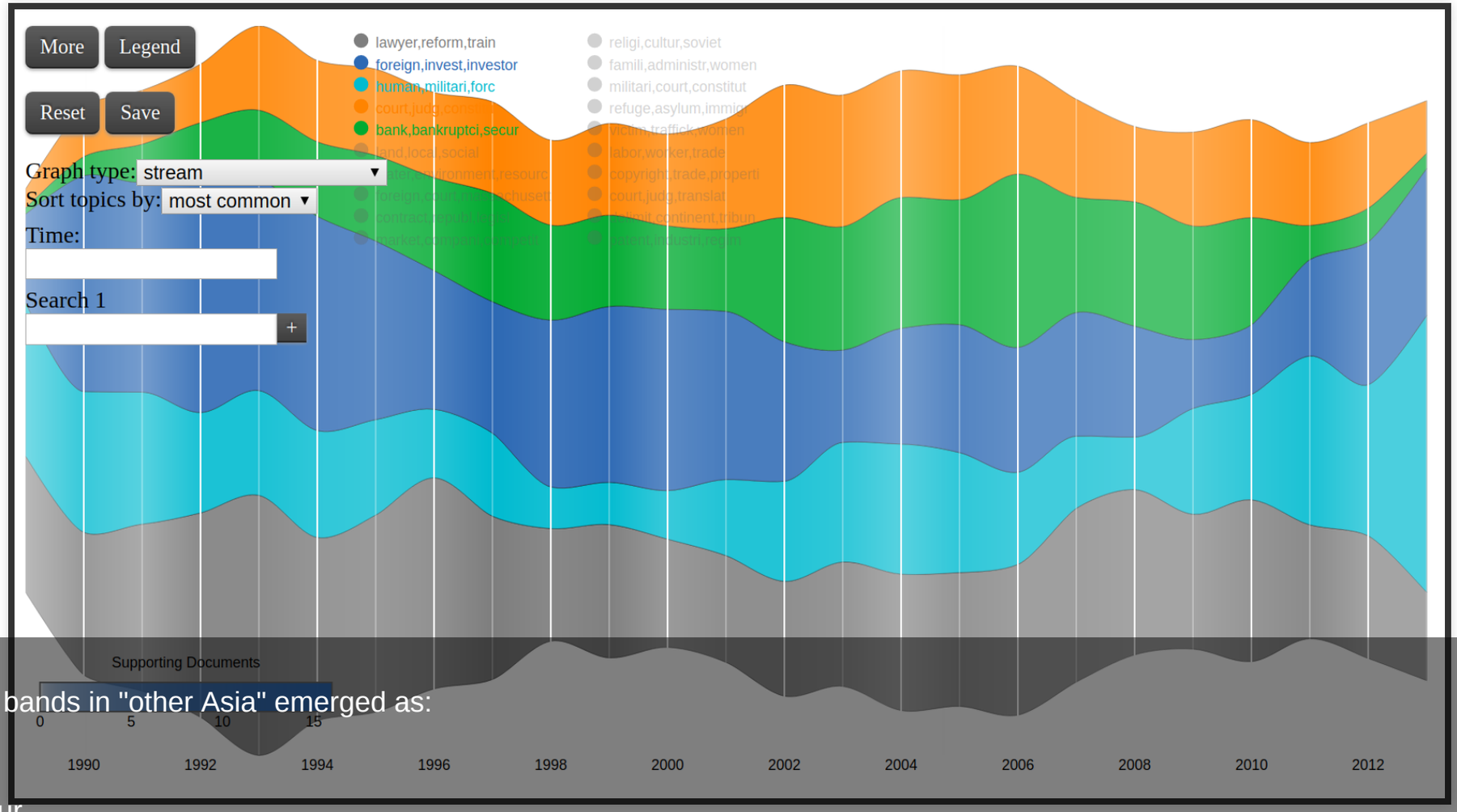


The top five LDA topic bands for Germany emerged as:

- compani,corpor,sharehold
- constitut,administr,bverfg
- privat,econom,theori
- crimin,crime,victim
- intern,treati,nation

Germany

Patterns are consistent with German participation in multi-party cross-border conversations within and beyond the European Union concerning its solutions to common problems of law and governance. This is the familiar face of "comparative law" as the term is commonly understood.



The top five LDA topic bands in "other Asia" emerged as:

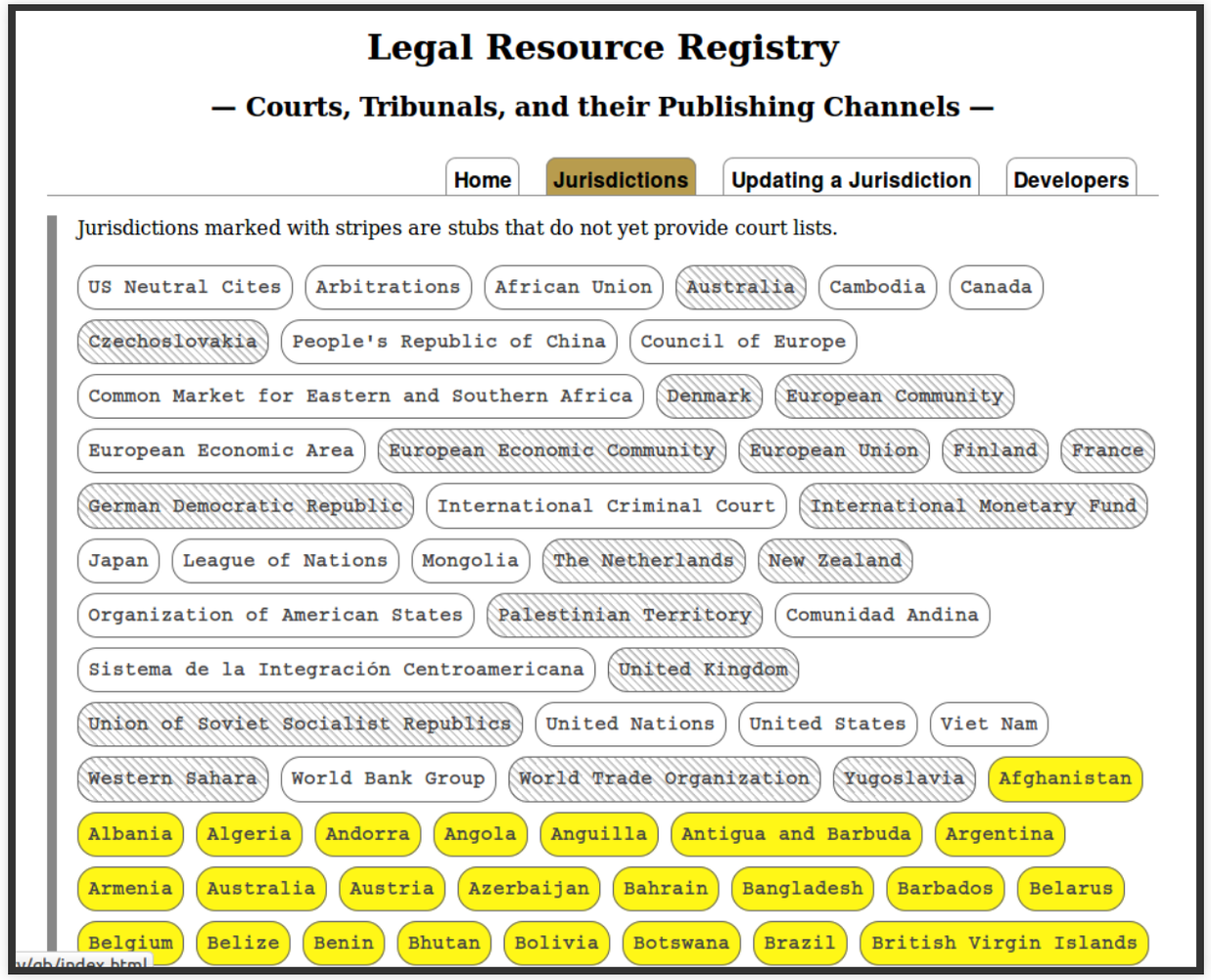
- court,judg,constitut
- bank,bankruptci,secur
- foreign,invest,investor
- human,militari,forc
- lawyer,reform,train

Asian Countries in Transition

The small sample size as well as the limitations of LDA analysis dictate caution in drawing firm conclusions from these patterns; but they do appear to reflect the exogenous preferences of remote actors, either for purposes of control or for the attraction of foreign investment. If that is correct, horizontal communication between systems about detailed aspects of law, policy and procedure is yet to emerge; and given the language barriers within Asia, legal and multilingual reference management will have a role to play in fostering such dialog.

LEGAL RESOURCE REGISTRY

Herein of the home-brew infrastructure behind Juris-M jurisdiction identifiers.



Juris-M jurisdiction identifiers are derived from the "Legal Resource Registry" (LRR). The idea for the LRR emerged during discussions of the Legal CiteM OASIS standards group chaired by John Joergensen and Fabio Vitali. The initial data set was built from IDs provided by CourtListener, and the World Law Guide collection national court information, reused with the consent of its author Harry Moers of Lexadin. Grey-stripped jurisdictions in the LRR contain court information that has not yet been verified.

Legal Resource Registry

— Courts, Tribunals, and their Publishing Channels —

Home | Jurisdictions | Updating a Jurisdiction | Developers

United States

Supreme Court · D.C. Circuit · First Circuit · Second Circuit · Third Circuit · Fourth Circuit · Fifth Circuit · Sixth Circuit · Seventh Circuit · Eighth Circuit · Ninth Circuit · Federal Circuit · Tenth Circuit · Eleventh Circuit · Alaska · Alabama · Arkansas · American Samoa · Arizona · California · Colorado · Connecticut · District of Columbia · Delaware · Florida · Georgia · Guam · Hawaii · Iowa · Idaho · Illinois · Indiana · Kansas · Kentucky · Louisiana · Massachusetts · Maryland · Maine · Michigan · Minnesota · Missouri · Northern Mariana Islands · Mississippi · Montana · Navajo Nation · North Carolina · North Dakota · Nebraska · New Hampshire · New Jersey · New Mexico · Nevada · New York · Ohio · Oklahoma · Oregon · Pennsylvania · Puerto Rico · Rhode Island · South Carolina · South Dakota · Tennessee · Texas · Utah · Virginia · U.S. Virgin Islands · Vermont · Washington · Wisconsin · West Virginia · Wyoming

The user-facing view of the LRR is a simple layer of HTML hosted via GitHub Pages.

Legal Resource Registry

— Courts, Tribunals, and their Publishing Channels —

Home | Jurisdictions | Updating a Jurisdiction | Developers

United States

D.C. Circuit

Court of Appeals for the D.C. Circuit

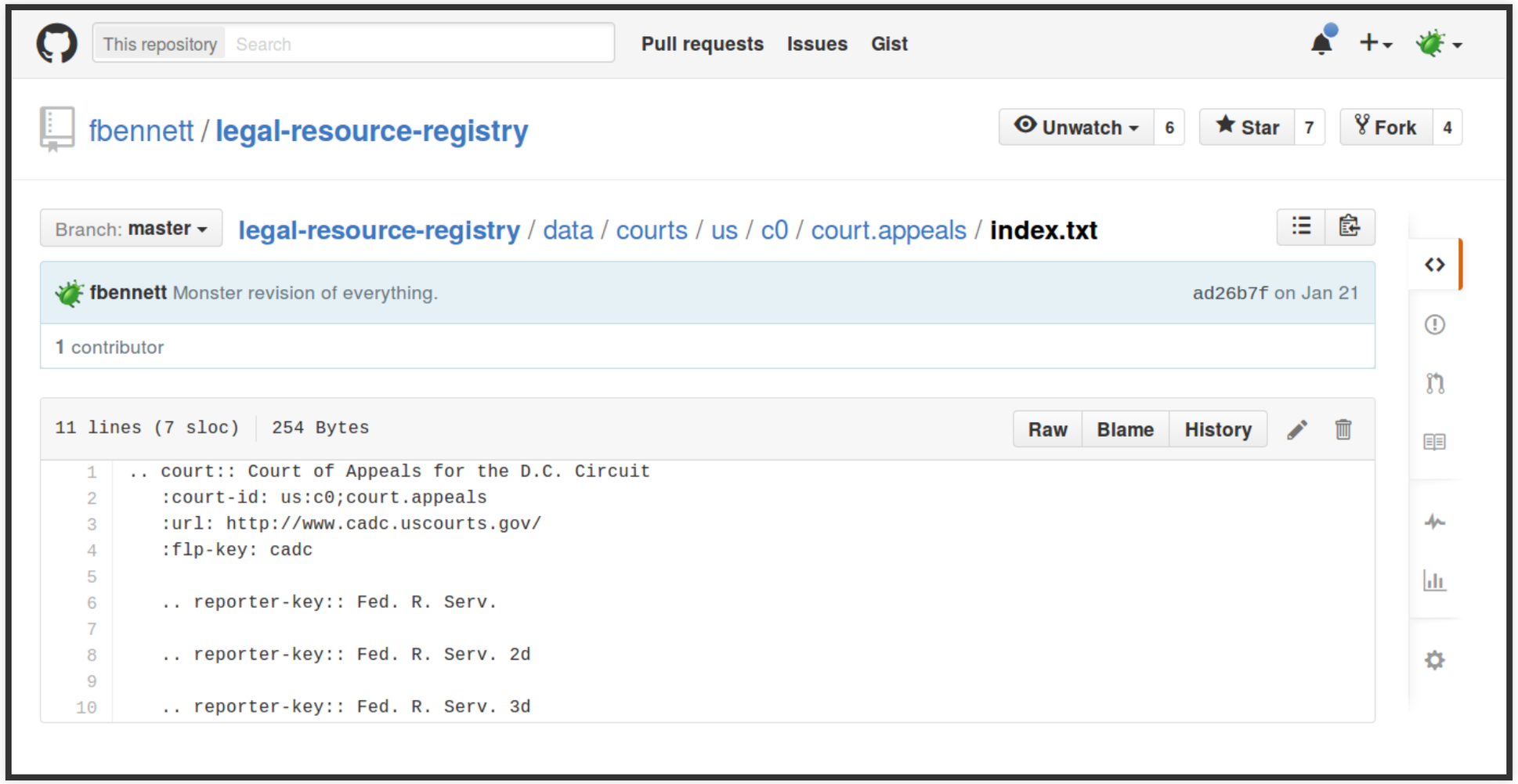
us:c0;court.appeals

Fed. R. Serv.	
From: 1938	casename: *req'd*
To: present	date-published: *req'd*
Neutral: no	jurisdiction: *req'd*
Confirmed: no	page: *req'd*
	reporter: *req'd*
	volume: *req'd*
	court-place: *optional*

Fed. R. Serv. 2d	
From: 1938	casename: *req'd*
To: present	date-published: *req'd*
Neutral: no	jurisdiction: *req'd*
Confirmed: no	page: *req'd*
	reporter: *req'd*
	volume: *req'd*
	court-place: *optional*

Fed. R. Serv. 3d	
From: 1938	casename: *req'd*
To: present	date-published: *req'd*
Neutral: no	jurisdiction: *req'd*
Confirmed: no	page: *req'd*
	reporter: *req'd*
	volume: *req'd*
	court-place: *optional*

For US courts, the LRR contains details of reporters in which decisions of each listed court might be published. This information is derived from CourtListener data (and has not been updated in some time).

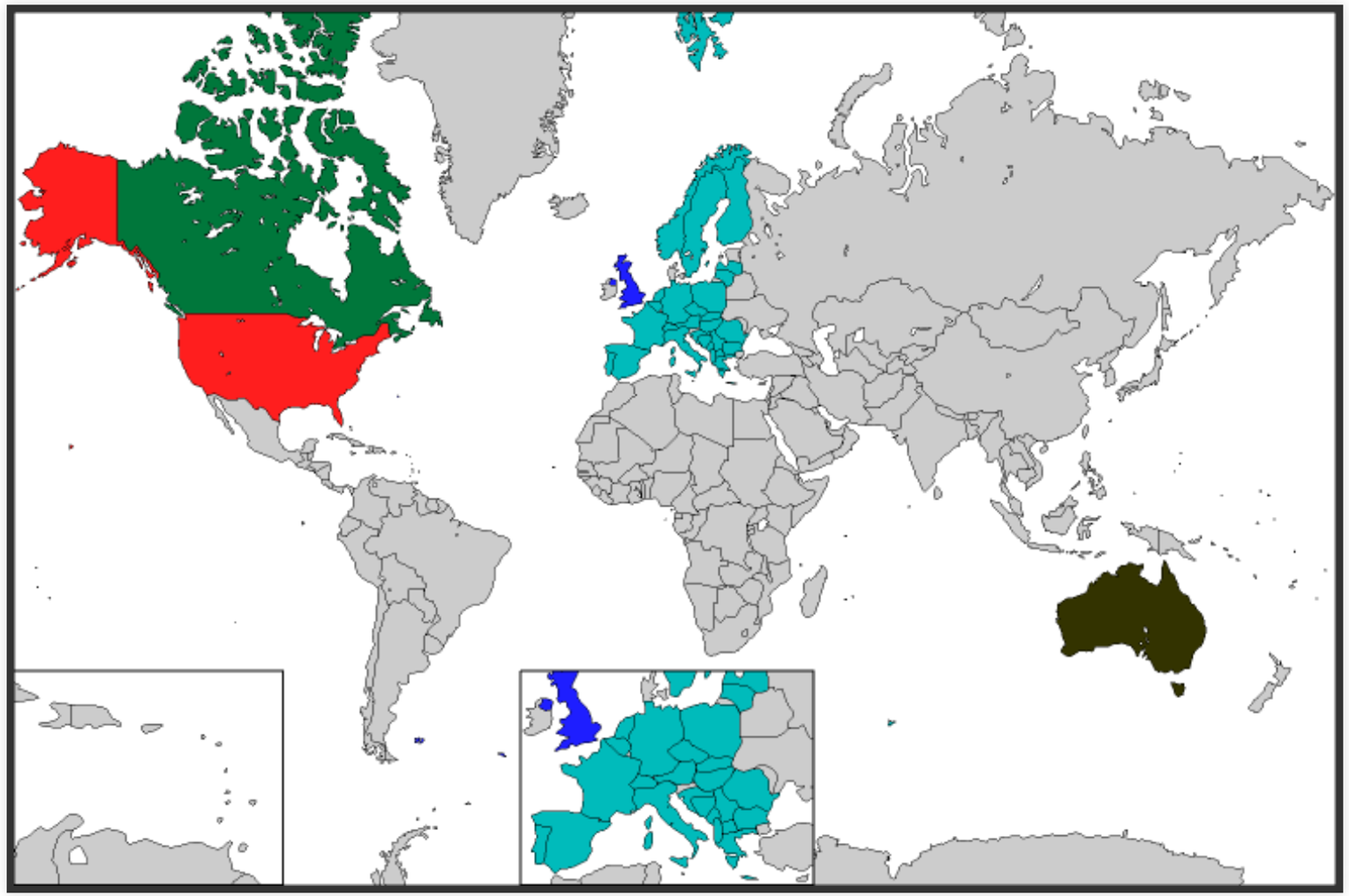


Clicking through to an identifier reveals the underlying plain-text source file, which resides in a simple file hierarchy on GitHub. The format of the identifiers themselves is defined in the IETF draft URN:Lex specification proposed by PierLuigi Spinosa, Enrico Francesconi and Caterina Lupo. Source files are set in `reStructuredText`, and the repository contains tools for grinding the data set into forms required by consuming applications.

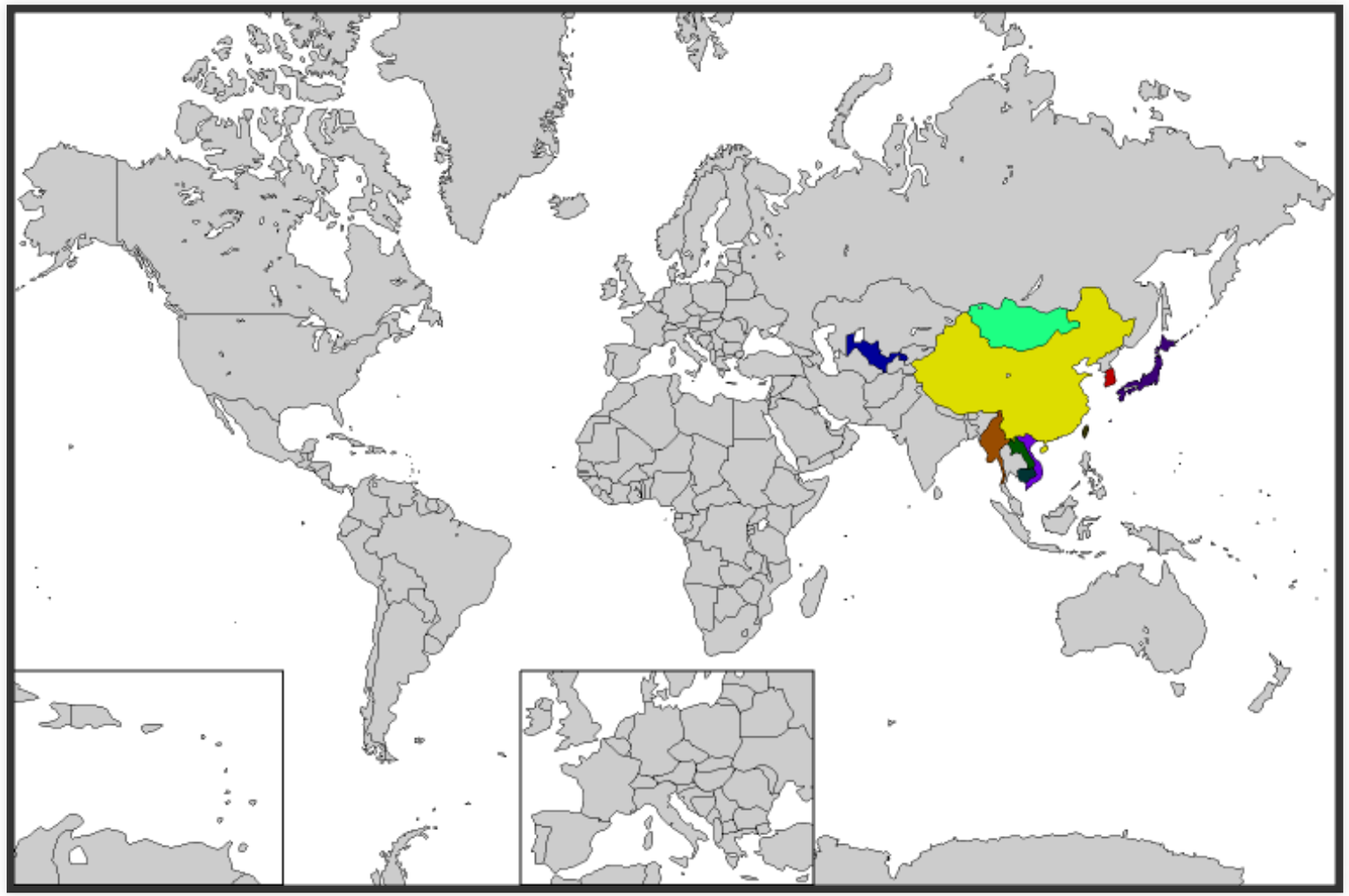
This should be converted to a database of some sort, obviously, with a proper set of front-end tools for collaborative maintenance.

MOVING FORWARD

Herein of stovepiping and the DIY challenge facing modern comparative law scholarship.



In closing, a rough observation on come of the jurisdictions where archival work has given rise to fine-grained jurisdiction identifiers. There are certainly others, but those shown above are perhaps those with the highest profile. These are federal jurisdictions, in which amalgamating resources from multiple realms is necessary in order to present a global view of legal resources. This creates an incentive for development of uniform jurisdiction identifiers, but the incentive evaporates at the border of each federal entity.



The primary Asian target jurisdictions of postgraduate research in the School of Law at Nagoya University, the home of Juris-M, do not form a federal entity. Here, and for the world at large, experience tells us that if legal researchers desire a comprehensive system for the proper organization and study of the legal instruments of the world, it is up to us, as a community of researchers, to establish its foundations.

THANK YOU

https://juris-m.github.io/

https://fbennett.github.io/legal-resource-registry

In conclusion, things to click on!

Slides by F Bennett