Open Legal Data and AI: Past, Present & Future

Presented by: William E. Palin, Esq.

Head of Case Law and Developer @ Free Law Project

bill@free.law

@flooie

https://free.law



Glad to be back!







Free Law Project

Free Law Project, a 501(c)(3) non-profit, uses technology, data, and advocacy to democratize access to legal information, regardless of socioeconomic status. We strive to make the law understandable and accessible to all, using advanced technology and data analysis to break down barriers.









Open Legal Data and AI: Past, Present and Future

If you think AI is going to change the future of the legal system it is important to understand that it has already been changing the legal landscape for years.



See and Hear and Speak (or write)



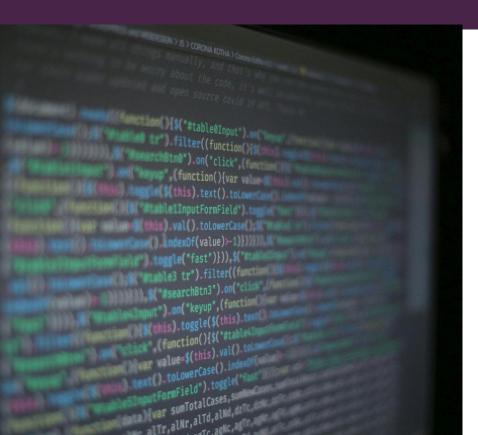
Al to: See

Hear

Speak

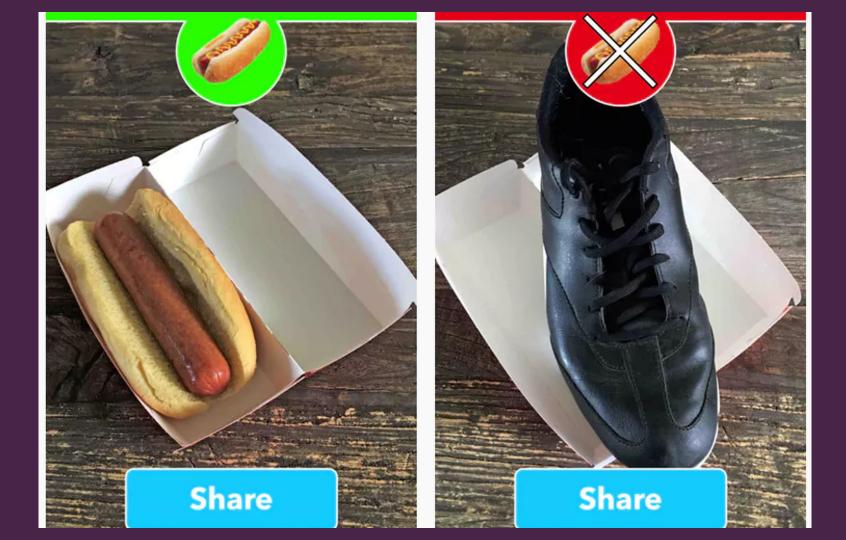


FLP Loves Machine Learning



- Image Classification
- Object Detection
- Text Classification
- Speech to Text
- And building LLM datasets





Explain boy you will resolve any potential conflict of interest, including the procedure you will follow in determining these was a superior of the procedure of a litigation and financial arran, identify the categories of litigation and financial arran, identify the categories conflicts of the procedure of the

Although I don't anticipate any conflicts, I will follow the guidedname created under the code of Judicial Common should natter as the code of Judicial Common should natter the code of the code before court. In case the code of interest, I will seek appropriate guidance and cities roccae or dreet, as needed.

Do you have any plans, commitments, or agreements to pursue outside employment, with or without compensation, during your service with the court? If so, explain.

Board of Trustees, College of the Holy Cross

Board of Advisors, Gordon Public Policy Center Brandeis University, Waltham, MA 02254

List sources and amounts of all income received during the List sources and amounts of all income received during the calendar year preceding your part of the current calendar year, including all salaries, fees, dividends, interest, gitts, rents, rough prints, honoraris, and concern items exceeding 500 or more Hints, honoraris, and concern items exceeding 500 or more interest, gitter of see, occurrently the financial disclosure report prefer to de so, occurrently the financial disclosure report interest by the think of the financial disclosure reports.

Salary: \$22,500.00 (gross pay) galary: \$22,500.00 (gross pay) approximately Calendar Year 1989 (to date): Please complete the attached net worth statement in datail (add schedules as called for). Maye you ever hald a position or played a role in a political campaign? If so please issuitly the particulars of the campaign including the cambaiate, dates of campaign your title and responsibilities.

242

August, 1988 - Member of Caucus Team at Empublican National Convention - bush for President Campaign. October 23, 1988 Black Press Ball of Fame, Baltimore, B October 29, 1988 - People of Color Rally - Cincinnati, Ohio Bush/Chayle Caspaign - Speaker

FINANCIAL STATEMENT MET WORTH

Share

TESTIMONY OF ANITA F. HILL, PROFESSOR OF L. UNIVERSITY OF OKLAHOMA, NORMAN, OK

The CHAIRMAN. Professor Hill, please make whatever i you would wish to make to the committee.

The CHARMAN. Excuse me. I instruct the officers not anyone in or out of that door while Professor Hill is making

Ma. Hill. Mr. Chairman, Senator Thurmond, members of the committee, my name is Anita F. Hill, and I am a professor of law at the University of Oklahoma.

I was born on a farm in Okmulgee County, OK, in 1956. I am the t was born on a farm in Ukmuigee County, UK, in 1996. I am the youngest of 13 children. I had my early education in Okmuigee County. My father, Albert Hill, is a farmer in that area. My mother's name is Erma Hill. She is also a farmer and a housewife.

My childhood was one of a lot of hard work and not much money, but it was one of solid family affection as represented by my parents. I was reared in a religious atmosphere in the Baptist faith, and I have been a member of the Antioch Baptist Church, in Tulsa, OK, since 1983. It is a very warm part of my life at the

For my undergraduate work, I went to Oklahoma State Universitor my undergraduate work, I went to Okuanoma State University, and graduated from there in 1977. I am attaching to the statement a copy of my resume for further details of my education. The CHAIRMAN, It will be included in the record.

Ms. Hill. Thank you.

I graduated from the university with academic honors and proceeded to the Yale Law School, where I received my J.D. degree in

Upon graduation from law school, I became a practicing lawyer with the Washington, DC, firm of Wald, Harkrader & Ross. In 1981, I was introduced to now Judge Thomas by a mitual friend. Judge Thomas told me that he was anticipating a political appointment and asked if I would be interested in working with him. He ment and asked it I would be interested in working with him. Lie was, in fact, appointed as Assistant Secretary of Education for Civil Rights. After he had taken that post, he asked if I would become his assistant, and I accepted that position.

In my early period there, I had two major projects. First was an article I wrote for Judge Thomas' signature on the education of minority students. The second was the organization of a seminar on high-risk students, which was abandoned, because Judge Thomas transferred to the EEOC, where he became the Chairman of that

During this period at the Department of Education, my working relationship with Judge Thomas was positive. I had a good deal of responsibility and independence. I thought he respected my work

responsibility and independence. I models as responsibility and that he trusted my judgment.

After approximately months of working there, he asked me to go out socially with working the ment and telling the good socially with working the ment and telling the good social with the ment difficult things, experiences of my world about it are the two most difficult things, experiences of my ment and a second process of my ment and a second proces life. It is only after a great deal of agonizing consideration and a number of sleepless nights that I am able to talk of these unpleasant matters to anyone but my close friends.

For my first months at the EEOC, where I continued to be a substant to Judge Thomas to Indian Theorem 1 For my first months at the EDOV, where I continued to be assistant to Judge Thomas, there were no excual conversation overtures. However, during the fall and winter of 1982, these because of the property of overtures. However, during the fall and winter of 1983, these be again. The comments were random, and ranged from pressing about thy 1 didn't go out with him, to remarks about my personal processing the state of the comments are the state of the comments. about why I didn't go out with him, to remarks about my person appearance. I remember him saying that "some day I would he to tall him the real "secon that I wouldn't go out with him."

He began to about displacance in the team and would would he

to tell him the real reason that I wouldn't go out with him.

He began to show displeasure in his tone and voice and his can meaning in his down the comment of an explanation. He commended on what I was a superstance of the commended on the com

Thomas was d at which we looked at the

the summer recall, the last day of my employ work directly from bork to last day of my employ the last day of my employ the last from bork to law of my employ the last from bork to law end are season that the last from the las On other being larger late 1982, I F cerned that legrading ought the apped be

socially with him, and exould jeopardize what at the time king relationship. I had a normal of the office. I believed then, as hip with a person who was suised. I was very uncomfortable

nd explaining my reasons, my suggestions. However, to my continued to ask me out on ustify my reasons for saying ace in his office or mine. They tions which would not have

more strained when Judge discuss sex. On these occareports on education that because of the time a government cafeteria. turn the conversation to sations were very vivid.

had seen in pornographic films in-nen having sex with animals, and films ampass or rape scenes. He talked about pernographic depicting individuals with large penises, or large breasts invested especting individual with large peniss, and films invested in various sex acts.

In several consumers are acts of the penisses, or large penisses, or large

Share

ulsa. On one occa We had worked Andoubtedly, I passed from the Thomas There were a successioned by a work as conference The group in the third my assistance for the conference of the confe

to no effect and finally ad call directly. She did so itation, Ms. Cahall wante

and some research was information and did attempted another possible conference

ML Models on Images

	794	
AG-10 FINANCIAL DISCLOSU	URE REPORT Annual Region Due by May 15 certain Judicial Employees (28	from Judicus Officers and USCA App. 1, 91 301-69)
Smith, Fern M.	U.S. Dist. Ct. N. D. Cal.	5/10/88
Ludge	Does of Energy/Nomination/Termination (only if minut or float report) May 9, 1988	11/57
Home or offer where Superior Cou Hall of Justice San Francis	nt - Dept 25 2 - 850 Bryant St. Co, Cal. 94108	\$/30/28
BMFORTANT NOTES: Please read the information persishing to your spouse an identifying each attachment by showing vox- Complete all sections, checking the NON Compare and recording the report with la- prist clearly. Sign on last page.	instructions accompanying this form. The report of dependent children, if any Atlanch additional ur name, the date of the report, and the section(s IE) but for each section where you have an report at you'r and this tiems in the same order as let	rt should include sheets if needed,) being completed, ertable information, st year. Type or
POSITIONS. (Reporting individual only; see POSITION	e pp. 15-16 of Instructions.) NAME OF ORGANIZATION/E	NULLA
NONE (No reportable positions)		
Judge	Superior Court. San	Francisco
Trustee	Fern M. Smith Tru	
Annual instal	De M. P. To districtions) PARTIES AND TERMS L. And Bes as a Resease of the bear of the b	
DATE NONE (to reprode agreemen)	PARTIES AND TERMS L. and Breasen Breasen & 18th Ke are as of Sept 1916 to be Martin above System Berlind. Martin above for pour, we pp. 18-20 of humoions;	made in
EATE NONE (the reproduct against set) Sept. 1985 Fr. a. M. Smith Gotarskip Sh Annual Insylul III. NONINVESTMENT INCOME (the (the continue) NONE (the reproduct and incomments are continued)	PARTIES AND TERMS L. And Beaute n Francis n & IR. K. acc as of Sept 1986 to be Acc as of Sept 1986 to be Acc as of Sept 1986 to be Considerated as the Sept 1988 of Inductions (CC. AND TITE	GROSS INCOME
DATE NONE (On-myndatagement) Sept. 1986 Fe C. M. Sm. t. Sect. 1986 Fe C. M. Sm. t. SOUE NONE (On-myndatagement) Stude of C. M. Sm. t. Stude of C. M. Sm. t.	PARTIES AND TERMS L. and Booaco A. Broade A. J. M. K. and Booaco A. Broade A. J. M. K. and Booaco A. Sept. 1816 to k. (Machine And Syear Benedicture) KI AND TITE (Septime - Wages)	GROSS INCOME (DOWN, not spouse s)
MONE (Personal Section 19 South	PARTIES AND TERMS L. and Because a Because of IR-Ke ance at or Sept. 1916 to be ("mant's BAN" Syear secret ("mant's BAN" Syear secret ("mant's BAN" Syear secret ("mant's BAN" Syear secret ("mant's BAN" TITLE ("spenia" Wage: a) Seafon Millo Wage:	GROSS INCOME (DOWN, not spouse s)
DATE NONE (proposed agreement) Suppl 1886 Sec. M. Se	PARTIES AND TERMS L. and Booaco A. Broade A. J. M. K. and Booaco A. Broade A. J. M. K. and Booaco A. Sept. 1816 to k. (Machine And Syear Benedicture) KI AND TITE (Septime - Wages)	GROSS INCOME (DOWN, not spouse s)

Pag	ANCIAL DISCLOSURE REPORT te 5 of 39	Name of Person Reporting Stein, Sidney H							Date of Report 8/16/2005	
VII. INVESTMENTS and TRUSTS - incores, whose, transcardious (includes fuence of the spoune and dependent children. See pp. 14-57 of filing instructions.) A. B. C. D.										
	Description of Assets (including trust assets)	Income during reporting period		Gross value at end of reporting period		Transactions during re			porting period upt from disclosure	
Place "(X)" after each asset energe from prior disclosure	(1) Amount Code I (A -H)	Type (e.g. div. rest. or int.)	Value Code 2 (I-P)	Value Method Code 3 (Q-W)	(1) Type (e.g. buy, sell, merger, redemption)	(2) Date: Month- Day	(3) Value Code 2 (J-P)	(4) Guin Code I (A-H)	(3) Identity of buyerhelle (if private transaction	
73.	CHINA TELECOM CORP. LTD.									
74.	CHINA TELECOM CORP. LTD.								7	
75.	CIGNA CORP.									
76.	CISCO SYSTEMS		-			-			-	
77.	CISCO SYSTEMS									
78.	CONOCO PHILLIPS									
79.	CONOCO PHILLIPS									
80.	CONSTELLATION ENERGY									
81.	COOPER INDUSTRIES CL. A									
82.	DELL COMPUTERS									
83.	EMC CORP. MASS.	7								
34.	EMC CORP. MASS.									
35.	EBAY INC.									
36.	EBAY INC.									
87.	EBAY INC.									
38.	FOREST LABS CL. A									
39.	POREST LABS CL. A									

14. J 89/21: S. HRG. 162-505, et. 5

S. HRG. 102-505, Pr. 5

CONFIRMATION HEARINGS ON FEDERAL APPOINTMENTS

HEARINGS

BEFORE THE

COMMITTEE ON THE JUDICIARY UNITED STATES SENATE

ONE HUNDRED SECOND CONGRESS

FIRST SESSION

ON

CONFIRMATION HEARINGS ON APPOINTMENTS TO THE FEDERAL JUDICIARY

JULY 24, 29, 31, AND SEPTEMBER 26, 1991

Part 5

Serial No. J-102-7

Printed for the use of the Committee on the Judiciary



You can do it too!





- 1. Go to our github repo
- 2. Follow instructions on how to build and train the model yourself

https://github.com/freelawproject/disclosure-extractor



Inside the Data:

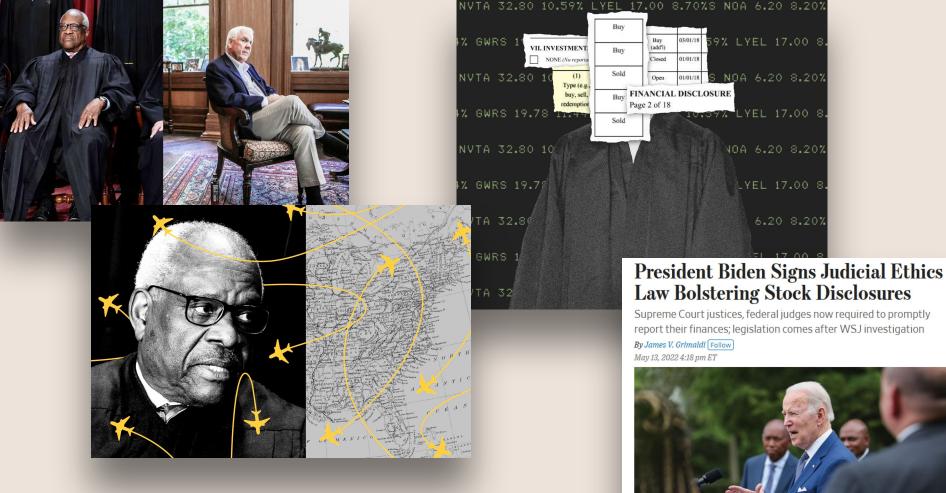
Travel Reimbursements to over 600 locations across the globe.





Millions of dollars in art investments





PROPUBLICA THE WALL STREET JOURNAL.

Using ML to Improve Case Law





- Enhanced over 1 million case law opinions from the case law access project
- Trained models to improve data structure
- Which allows us to enhance the data in our models and lets our users do the same





So What is next?



Train & Build the Next Generation of LLMs

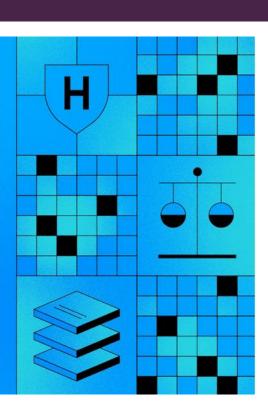
COLD

Collaborative Open Legal Data

CASES



COURT LISTENER







OpenAl Data Partnerships

Working together to create open-source and private datasets for Al training.

Submit interest





"Including your content can make AI models more helpful to you by increasing their understanding of your domain. We're already working With many partners who are eager to represent data from their country or industry. For example, we recently partnered with the Icelandic Government and Miðeind ehf to improve GPT-4's ability to speak Icelandic by integrating their curated datasets. We also partnered with non-profit organization Free Law Project, which aims to democratize access to legal understanding by including their large collection of legal documents in Al training. We know there may be many more who also want to contribute to the future of AI research while discovering the potential of their unique data."



Advancing AI Through Partnerships:



Clear brief













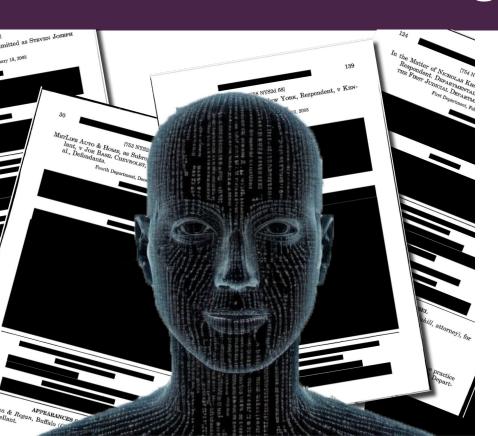
Present(ish): Speech-to-Text

Over 2,000 days of continuous oral argument audio from over thirty thousand cases.





Future: A new scanning project for the future



Our upcoming scanning project will leverage AI to be faster, more successful, and improve the quality of the scans.

Al will also be used to create REDACT, a tool to remove any information that has copyright restrictions.



Thank you!



William E. Palin, Esq.

Free Law Project Case Law Lead

bill@free.law

Socials: @flooie https://free.law



Sign up for our _ newsletter!

