Simon Willison's Weblog

## **July 2022**

Search posts from July 2022

Search

31 posts: 4 entries, 25 links, 2 quotes

#### July 6, 2022

<u>Defensive CSS</u> (<u>via</u>) Fantastic new site by Ahmad Shadeed describing in detail CSS patterns which can help build layouts that adapt well to unexpected content—things like overly long titles or strange aspect ratio images, common when you are designing against UGC.

# 5:16 pm / css, frontend, ahmad-shadeed

<u>Bun</u>. "Bun is a fast all-in-one JavaScript runtime"—this is very interesting. It's the first project I've seen written using the Zig language, which I see as somewhat equivalent to Rust. Bun provides a full Node.js-style JavaScript environment plus a host of packaged tools—an npm install client, a TypeScript transpiler, bundling tools—all wrapped up in a single binary. The JavaScript engine itself extends JavaScriptCore. Bun also ships with its own wrapper for SQLite.

# 5:24 pm / javascript, sqlite, npm, typescript, zig

Helpful 404s for Jekyll (and GitHub Pages). Neat trick from Ben Balter: JavaScript that runs on your 404 page, fetches the sitemap.xml, parses all of the URLs out of it and then uses a levenshtein edit-distance comparison to find the closest URL to the one that you landed on and suggests that as a "Perhaps you're looking for?".

# 5:31 pm / javascript

How to Temporarily Disable Face ID or Touch ID, and Require a Passcode to Unlock Your iPhone or iPad. Hold down the power and volume up buttons for a couple of seconds, and your iPhone will no longer allow you to use FaceID to unlock it without first entering your passcode.

# 5:38 pm / iphone, security

## July 7, 2022

SOC2 is about the security of the company, not the company's products. A SOC2 audit would tell you something about whether the customer support team could pop a shell on production machines; it wouldn't tell you anything about whether an attacker could pop a shell with a SQL Injection vulnerability.

— Thomas Ptacek

# 8:31 pm / security, sql-injection, thomas-ptacek, fly

## **July 9, 2022**

## Using GPT-3 to explain how code works

SQLite create a table that has a compound primary key of type, id and a content column.

CREATE TABLE pages (type TEXT, id TEXT, content TEXT, PRIMARY KEY (type, id));

CREATE TABLE pages\_done (type TEXT, id TEXT, PRIMARY KEY (type, id));

SQL query returning rows in the pages table that do not have corresponding records in the pages\_done table:

SELECT type, id, content FROM pages WHERE (type, id) NOT IN (SELECT type, id FROM pages\_done)

That query but using a join:

SELECT p.type, p.id, p.content FROM pages p LEFT JOIN pages\_done pd ON p.type = pd.type AND p.id = pd.id WHERE pd.type IS NULL;

Which of those two queries is more efficient and why?

The first query is more efficient because it uses a WHERE clause with a <u>subquery</u>. The second query is less efficient because it uses a LEFT JOIN.

Why is a subquery more efficient?

A <u>subquery</u> is more efficient because it only returns the rows that meet the conditions specified in the WHERE clause. A LEFT JOIN returns all rows from the left table, even if there is no match in the right table.

One of my favourite uses for the GPT-3 Al language model is generating explanations of how code works. It's shockingly effective at this: its training set clearly include a vast amount of source code.

[... 1,983 words]

3:19 pm / ai, gpt-3, openai, prompt-engineering, generative-ai, llms, ai-assisted-programming

## July 10, 2022

<u>GPT-3 prompt for spotting nonsense questions</u> (via) In response to complaints that GPT-3 will happily provide realistic sounding answers to nonsense questions, rictic recommends the following prompt:

I'll ask a series of questions. If the questions are nonsense, answer "yo be real", if they're a question about something that actually happened, answer them.

# 4:33 am / gpt-3, openai, prompt-engineering, generative-ai, llms

## July 11, 2022

# sqlite-comprehend: run AWS entity extraction against content in a SQLite database

I built a new tool this week: <u>sqlite-comprehend</u>, which passes text from a SQLite database through the AWS Comprehend entity extraction service and stores the returned entities.

[... 1,146 words]

9:31 pm / aws, git, github, projects, sqlite, datasette, weeknotes

#### **July 13, 2022**

Bringing page transitions to the web (via) Jake Archibald's 13 minute Google I/O talk demonstrating the page transitions API that's now available in Chrome Canary. This is a fascinating piece of API design—it works by effectively creating a static image screenshot of the before and after states of the transition, then letting you define CSS animations that animate a transition between the two static images. By default the screenshot encompasses the full viewport, but you can instead define multiple elements within the page and apply separate transitions to them. It's only available for SPAs right now but the final design should include support for multi-page applications as well—which means transitions with no JavaScript needed at all!

# 4:26 pm / css, javascript, google-io, jake-archibald, view-transitions

#### July 14, 2022

The DALL·E 2 Prompt Book (via) This is effectively DALL-E: The Missing Manual: an 81 page PDF book that goes into exhaustive detail about how to get the most out of DALL-E through creative prompt design.

# 11:26 pm / ai, openai, dalle, prompt-engineering, generative-ai

## **July 15, 2022**

Datasette Discord community (via) I started a Discord chat community for Datasette. 57 people have joined up already!

# 3:17 am / datasette, discord

## **July 19, 2022**

<u>Soft Deletion Probably Isn't Worth It</u>. Brandur argues that soft deletion—where you delete records by populating a "is\_deleted" or "deleted\_at" column in your table—isn't worth the additional complexity and risk it adds to other database queries. Instead, he suggests having a separate deleted records table which records the deleted data in a JSON blob—allowing you to review and recover it manually if necessary, and giving you an easy way to expire deleted records that have exceeded your retention policy.

# 8:40 pm / databases, brandur-leach

## July 20, 2022

The Checkered Flag Diagram for visualizing SQL joins. I really like this alternative to Venn diagrams for showing the difference between different types of SQL join (left join, right join, cross join etc).

# 1:16 pm / sql

<u>Visual Studio Code: Development Process</u> (<u>via</u>) A detailed description of the development process used by VS Code: a 6-12 month high level roadmap, then month long iterations that each result in a new version that is shipped to users. Includes details of how the four weeks of each iteration are spent too.

# 4:34 pm / microsoft, software-engineering, vs-code

#### Weeknotes: Datasette, sqlite-utils, Datasette Desktop

A flurry of releases this week, including a new Datasette alpha and a fixed Datasette Desktop.

[... 1,113 words]

11:13 pm / datasette, weeknotes, datasette-cloud, sqlite-utils, annotated-release-notes, datasette-desktop

How John Wiseman tracks worldwide GPS interference (via) Part of the ADS-B signals broadcast by commercial aircraft include a measure of GPS accuracy. By collecting global data every day, John is able to generate a map of areas that are experiencing higher than expected GPS interference, which generally corresponds to military jamming technology. He sometimes posts the resulting maps on Twitter—he just picked up increasing jamming activity around Moscow.

# 11:45 pm / gps, john-wiseman

## **July 22, 2022**

**Promise Maps**. Egbert Teeselink describes a neat JavaScript caching pattern: instead of caching key:value cache key:promise-that-resolves-to-value—doing this gives you dog piling prevention for free, because the first lookup of a value trigers the computation to fetch it while subsequent lookups wait on the same promise to resolve—or resolve instantly if the computation has completed.

# 3:52 pm / dogpile, javascript

## **July 24, 2022**

I discovered a while ago that all those errors and bugs that only appear when you demo something to an audience also magically appear when you record yourself demoing it to nobody. Maybe narrating a feature to a pretend audience takes the blinders off enough that you notice little mistakes you wouldn't have otherwise.

— karaterobot

# 8:59 pm / testing

You should take more screenshots (via) Alex Chan suggests saving screenshots of your work, since they may well last a lot longer than the projects themselves. I try to do that these days but I have SO many projects from the past that I didn't capture in this way, and that I really regret not keeping a better visual record of.

Sqitch tutorial for SQLite (via) Sqitch is an interesting implementation of database migrations: it's a command-line tool written in Perl with an interface similar to Git, providing commands to create, run, revert and track migration scripts. The scripts the selves are written as SQL in whichever database engine you are using. The tutorial for SQLite gives a good idea as to how the whole system works.

# 11:44 pm / databases, migrations, sqlite

#### July 25, 2022

**Reduce Friction**. Outstanding essay on software engineering friction and development team productivity by C J Silverio: it explains the concept of "friction" (and gives great definitions of "process", "ceremony" and "formality" in the process) as it applies to software engineering, lays out the challenges involved in getting organizations to commit to reducing it and then provides actionable advice on how to get consensus and where to invest your efforts in order to make things better.

# 10:25 pm / software-engineering, management

#### July 26, 2022

<u>viewport-preview</u> (via) I built a tiny tool which lets you preview a URL in a bunch of different common browser viewport widths, using iframes.

# 12 am / css, iframes, mobile, projects, testing

**Cosmopolitan: Compiling Python**. Cosmopolitan is Justine Tunney's "build-once run-anywhere C library"—part of the αστμαlly pδτταblε εχεςμταblε effort, which produces wildly clever binary executable files that work on multiple different platforms, and is the secret sauce behind redbean. I hadn't realized this was happening but there's an active project to get Python to work as this format, producing a new way of running Python applications as standalone executables, only these ones have the potential to run unmodified on Windows, Linux and macOS.

# 8:43 pm / python, redbean, cosmopolitan, justine-tunney

## **July 27, 2022**

<u>SQLite Internals: Pages & B-trees</u> (via) Ben Johnson provides a delightfully clear introduction to SQLite internals, describing the binary format used to store rows on disk and how SQLite uses 4KB pages for both row storage and for the b-trees used to look up records.

# 2:57 pm / algorithms, databases, sqlite, ben-johnson

Fastest way to turn HTML into text in Python (via) A light benchmark of the new-to-me selectolax Python library shows it performing extremely well for tasks such as extracting just the text from an HTML string, after first manipulating the DOM. selectolax is a Python binding over the Modest and Lexbor HTML parsing engines, which are written in no-outside-dependency C.

# 5:55 pm / html, python

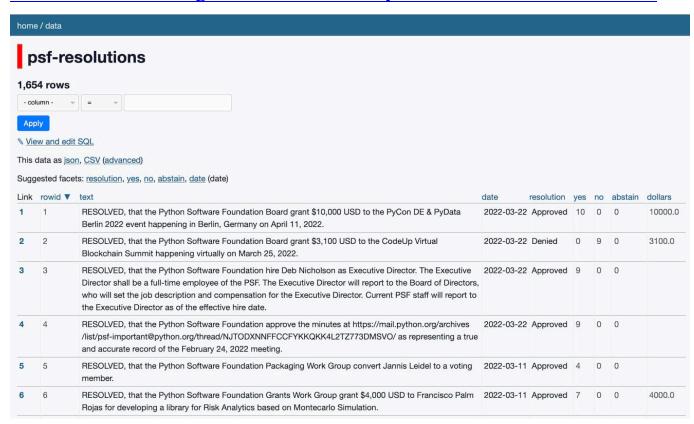
#### July 29, 2022

<u>Packaging Python Projects with pyproject.toml</u>. I decided to finally figure out how packaging with pyproject.toml works—all of my existing projects use setup.py. The official tutorial from the Python Packaging Authority (PyPA) had everything I needed.

# 11:18 pm / packaging, python

#### July 30, 2022

#### Weeknotes: Joining the board of the Python Software Foundation



A few weeks ago I was elected to the board of directors for the Python Software Foundation.

[... 2,081 words]

7:02 pm / python, sqlite, datasette, weeknotes, sqlite-utils, datasette-lite, psf

Introducing sqlite-lines—a SQLite extension for reading files line-by-line (via) Alex Garcia wrote a brilliant C module for SQLIte which adds functions (and a table-valued function) for efficiently reading newline-delimited text into SQLite. When combined with SQLite's built-in JSON features this means you can read a huge newline-delimited JSON file into SQLite in a streaming fashion so it doesn't exhaust memory for a large file. Alex also compiled the extension to WebAssembly, and his post here is an Observable notebook post that lets you exercise the code directly.

#7:18 pm / json, sqlite, observable, webassembly, alex-garcia

**GPSJam** (via) John Wiseman's "Daily maps of GPS interference" —a beautiful interactive globe (powered by Mapbox GL) which you can use to see points of heaviest GPS interference over a 24 hour period, using data collected from commercial airline radios by ADS-B Exchange. "From what I can tell the most common reason for aircraft GPS systems to have

degraded accuracy is jamming by military systems. At least, the vast majority of aircraft that I see with bad GPS accuracy are flying near conflict zones where GPS jamming is known to occur."

#7:51 pm / gis, gps, mapping, john-wiseman

GAS-ICS-Sync (via) Google Calendar can subscribe to ICS calendar feeds... but polls for updates less than once every 24 hours (as far as I can tell) greatly limiting their usefulness. Derek Antrican wrote a script using Google App Script which fixes this by polling calendar URLs more often and writing them to your calendar via the write API.

# 11:47 pm / google-calendar, icalendar

page 1 / 2 next »

#### 2022 » July

	M	T	W	T	F	S	S
					1	2	3
	4	5	6	7	8	9	10
	11	12	13	14	15	16	17
	18	19	20	21	22	23	24
	25	26	27	28	29	30	31

Colophon 2018 2019 2020