Radar Trends to Watch - January 2022

Developments in AI, IoT, Programming, and More
January 4, 2022

Learn faster. **Dig deeper**. **See farther**.

All things considered, December was a short month: I'm writing this on December 20, with a week and a half left until January. But it hasn't been inactive. We've seen a lot of exciting developments, including the (beta) release of APIs to GPT-3; new language models from Google, one of which is significantly smaller and more efficient than most large language models; and new tools for documenting the biases of natural language datasets.

We've also had bad news on the security front. Log4J, a logging library that's used in a lot of enterprise software, has multiple critical vulnerabilities that are being exploited. While the developers are working hard to find and release patches, these events underscore a big problem with open source software. The developers are a small group of dedicated, but underfunded, vol-

unteers. What processes can be put in place to ensure that open source software is maintained? (Please don't say DAOs. That just siphons funding away to others who don't contribute to maintenance.)

Artificial Intelligence and Machine Learning

- Coqui started working on open source tools for multilingual speech-to-text conversion. Pete Warden shows how to get started. James Cham argues that speech is a better route to augmented reality than vision and goggles.

- APIs to GPT-3 are now in beta, so GPT-3 can be called directly from programs. The APIs are all REST-based, with Python bindings; bindings for other languages are provided by the community. Fine-tuning GPT-3 with your own data is also now supported (in beta) by OpenAI.

- Google has created a new language model called Retro that has performance equivalent to models 25 times its size (comparing it specifically to a new

model named Gopher with 280 billion parameters). Retro incorporates a large database of sentences that it can consult to make its results more accurate.

- HuggingFace has developed the Data Measurements Tool to create documentation for natural language datasets. There's a python API in addition to a no-code interface. The interface provides descriptive statistics (size, average record length, etc.), distributional statistics (e.g., word counts), and comparative statistics (information about topics, biases, and associations).

- AWS SageMaker Canvas claims to allow businesspeople to develop Machine Learning applications to solve business problems with no programming experience. While it clearly makes programming easier, it provides no help on issues like data inbalance, bias, and validity.

- Timnit Gebru has founded the Distributed AI Research Institute (DAIR). This group will do AI research that is not influenced by corporate or military aims, but rooted in communities, and prioritizes people in the groups that are currently constantly being harmed.

- Data-centric AI is the next step in the development of AI: improved tools for data collection, augmentation, labeling, quality evaluation, governance, and more.

- AI as an assistive technology for psychotherapy: The AI doesn't interact with the patient, but analyzes the conversation between the patient and the therapist to determine which parts of the conversation are effective.

- Unlike other social media sites, Twitch is actually doing something about sockpuppet accounts. Suspicious User Detection uses AI to detect people who have created new accounts after being banned. Their posts are subjected to additional moderation.

- Google has developed a Multimodal AI model: a single model that can handle still image, video, and audio classification. This is a big step forward over models that can only handle a single kind of input.

- A deep learning framework can detect phishing messages with 99% accuracy.

- Learning robots for everyday tasks are beginning to move into the mainstream. In Google X's office, they have robots for wiping tables, sorting trash, and performing other "useful tasks," and they are starting to deploy these robots across the rest of the company. These robots aren't specialized, like Roomba; they are robots that are capable of learning how to do different things.

Security and Privacy

- The NSO Group may be in serious legal trouble, but it is only the tip of the iceberg. Is the "hacking for hire" industry too big to fail? Plenty of governments are willing customers.

- A critical zero-day in Log4J, a logging library widely used in enterprise software, has IT departments scrambling to patch and update their systems. Although it is widely used, like many older open source projects, Log4J maintenance is poorly funded and relies largely on volunteers.

- RLBox is a tool Mozilla developed to run libraries in their own fine-grained sandbox to protect against vulnerabilities in third-party libraries. Modules are compiled to WebAssembly, and then to native machine code, which effectively places boundaries on memory access and control flow.

- Thieves place AirTags in concealed places on a desirable car, then use it to track the car until it's in a place where they can conveniently steal it. As the month has gone on, we've seen more and more reports of thefts like this.

- Employees working from home are using "mouse movers" to defeat intrusive bossware that records whether or not they are at their computers.

Virtual and Augmented Reality

- Sexual harrassment in the Metaverse: Harrassment doesn't have to be physical. Nor does racism.

- Based on job listings, Google appears to be working on consumer-oriented augmented reality glasses (along with Facebook, Microsoft, Apple, and many others).

Connected Hardware (aka IoT)

- Subscribing to your car: Toyota users will have to pay an $80 annual subscription fee to use their car's "remote start" feature. The immediate cause is the "sunsetting" of 3G services, but this promises to be a much bigger trend across all kinds of devices.

- A camera the size of a grain of salt can produce images comparable in quality to traditional lenses. This could lead to advances in minimally invasive surgery.

- Having mended relations with Samsung, Google's WearOS for wearable devices is challenging Apple's watchOS for market share. We rarely mention market share, but this could be an important shift.

- Self-replication in Xenobots, living programmable robots made from frog cells. This sounds like a science fiction nightmare, but fortunately, they die quickly. At least right now.

Operations

- Mess with DNS is a tool written by Julia Evans (@b0rk) for experimenting with DNS. It gives you a real subdomain for which you can create and query records, and shows you everything that DNS is doing behind the scenes.

Programming

- Zed is a real-time collaborative editor for Rust, based on conflict-free replicated datatypes (CRDTs). CRDTs are likely to be an important tool for a new generation of collaborative software.

- HTTP/3, an update to (replacement for) HTTP/2 is here, and it's impressively fast. Google and Facebook are using it, nginx supports it, along with "modern" browsers.

- To go with Copilot, GitHub has an improved code search that is now available as a technology preview.

- Support for Rust in the Linux kernel is still experimental, but making good progress.

- Current container systems are designed for scaling, but don't address performance issues well. Apptainer is a specialized container system for high-performance computing, stressing highly-coupled parallelism and inter-process communications.

- Unifying observability, business analytics, and data infrastructure is a big opportunity for developers, operations, and business users. Observability gives operations staff the kind of insight into systems that is badly needed in other areas of business management.

- Bank Python is an (informal) dialect of Python that appears to be widely used at banks. There's no access to the filesystem, but there is access to a database of financial objects, which includes all the Python programs themselves.

Law and Legislation

- Facebook is being sued for $150B for its role in contributing to the Rohingya genocide in Myanmar. The amount is probably meaningless, but the strategy of using a civil lawsuit to force social media companies to be accountable for their actions is interesting.

- A popular family safety app, Life360, is selling highly accurate location data about its customers into the very shady data market.

- More state and national governments are considering legislation for AI "accountability" and requiring auditing of algorithms for bias. New York City, in particular, is requiring annual audits of AI models used by the city, and imposing fines for undisclosed use.

- China is restricting the import of encryption technology with key lengths greater than 256 bits. This restriction could have significant effect on supply chains, the use of cryptocurrency, and privacy. And it could be a model for other states that want to limit cryptography.

Quantum

Quantum chess isn't chess played by quantum computers; it's chess where the pieces move according to quantum rules. Superposition, entanglement, wavefunction collapse (measurement): all these come into play. Makes 3D chess look like a children's game.