Community Overview

Overview

Climate change is a rapidly becoming a threat to human existence. And concurrently, AI model development is evolving at a rapid pace. Thousands of AI models could be deployed across multinational boundaries to help gather and cross-correlate climate data in real time. This can lead to new discoveries and more accurate solutions, but also points to a larger issue - what global-scale framework could be used to integrate all these AI instances, to look for deeper patterns in the data and relay these findings?

We propose the next stage of AI evolution - the Enterprise Neurosystem. A single AI instance connected to every area of climate monitoring, with millions of real-time and historical data points for reference. This framework will autonomously analyze and quantify this data, while providing guidance on current climate conditions and course correction.

Just like the human body, in which many functions are fully autonomic (heartbeat, oxygen exchange, energy assimilation, etc.) and yet others require higher-order / conscious analysis (relationship management, movement and guidance systems, etc.), an IT data center and network environment possess an interesting similarity to this biological model. Humanity, it seems, has been subconsciously using its own architecture as the basis for building these connective systems.

The Enterprise Neurosystem is the final stage of this evolution. The combined effect of human and machine deterministic capabilities is beyond the power of either one functioning alone - and therefore highly complementary, given their respective strengths.

An open-source research consortium is required to achieve this objective - one that spans academia, industry, and government. A community that can incorporate a wide array of viewpoints and experiences in the development of this technology, which will be shared in a transparent and accessible manner.

Project Genesis

Four years ago, we were challenged by a multinational telecommunications firm to examine the role of artificial intelligence in mobile networks. The charter at that time was to explore the role of AI/ML in these architectures at both a granular level and a much larger organizational scale, and then provide deeper operational insight.

During the research phase, it became clear that many enterprises are creating AI models for specific analytics functions. But in many cases, these models run independently from other AI instances. Large-scale correlation of findings, particularly in real-time scenarios, was also missing.

We eventually came to realize that mobile networks and related data centers possess an interesting parallel to the human neurosystem. Our neurology is a connective biological framework that links different sensor and analytics subsystems, and cooperatively helps each one to function autonomously. It also correlates different inputs from each of these subsystems, and centralizes this analysis to determine when corrective attention is required.

We decided a similar evolutionary framework would emulate this neural topology within the IT domain, and in turn, enable a deeper and more complete level of analysis and perception - to build on the vision of a large-scale collective framework, and turn it into a single federated AI instance. This system would connect AI nodes and data sets across different areas of operation in a unified manner. And this framework, when employed, would provide a window into many aspects of climate change by identifying environmental trends and emerging challenges, and delivering guidance on potential outcomes and solutions.

This architecture includes high-value AI models dedicated to specific functions, to drive interest and adoption. Concurrently, we propose to build a top-tier interpretive and reporting intelligence – one that takes in all forms of real-time and historical data, both structured and unstructured, and can respond with autonomous remedial action or course-correction recommendations to a climate management team.

The framework would be assembled and nurtured by a research community that would welcome developing nations, government agencies, academics and the private sector in an open forum. We recognized that an architecture of this scale and critical importance should be open source for a variety of reasons: model design and code base transparency, ongoing community input - especially relating to algorithmic bias - and to provide freely available AI solutions and infrastructure capabilities to the world.

Challenge and Opportunity

The majority of AI applications are built upon a few dozen core frameworks that are repurposed for various analytics requirements. This can span a dizzying array of functions in a climate initiative, not to mention a government or multinational corporation.

The growing list of required climate IT system functions include, but are not limited to, network/messaging operations and security, IoT device and satellite data integration, AI model provenance and related digital asset management, dynamic IT resource allocation, MLOps frameworks, data usage/policy engines, and more.

And in the corporate and government domains, this can encompass supply chains and logistics, autonomous drones (infrastructure inspection), IT operations and security, financial systems and forecasts, manufacturing lines, natural language-based customer interaction/management, data center and network equipment maintenance, human resources, legal functions, insurance and risk analysis, global and regional regulatory frameworks, taxation analysis, and facilities security/employee verification.

Regardless of the vertical use case, these activities are typically siloed by their function-specific applications, with little or no integration to other systems. This limits the potential for collective inference and ultimately, the larger value of AI. Visibility is limited to a small subset of the total data, out of reach of wide frame data analysis and insight.

In addition, differences in vendor AI development philosophies have produced multiple variants of application architectures. As a result, customers will field AI applications with different degrees of complexity. A lack of application-level cohesion can also contribute to reduced bias assessment capabilities. Given the broad experience of data scientists, bias is both unfortunate and a real challenge.

This multi-architecture and vendor situation reveals fundamental AI operational issues, as a) the various vendors and the underpinnings of their architectures are vastly different, and b) system-wide insights continue to be incomplete, due a larger integration challenge.

And non-integrative analysis and self-reflection leads to inaccurate assessments - just as they do across humanity. If there were a way to fundamentally address these issues, and truly unify and cross-correlate all of these elements, we could achieve transformative infrastructure efficiencies. And over time, a more accurate and powerful analytics engine would be the end result.

Solution

We propose a unified analytics framework that spans all aspects of climate operations. Much like the neurology of the human body, it will be a series of interconnected AI/ML models, tailored to permeate and assess every aspect of the environment – and autonomously adjust to analyze new climate events.

The initial program will start with a small series of open source AI projects designed to tackle fundamental issues, each with a clear path to value. These will be linked by a common communication and AI/ML framework, which will be developed concurrently. The structure and relationships of these respective efforts will be determined by the members of these projects. Over time, this connective framework and its associated projects will extend awareness across the globe, tying together areas of climate analysis via the web, mobile networks, IoT sensors and satellite systems, and historical databases.

In its end state, an overarching intelligence would draw in and cross-correlate the data, and autonomously provide management for lower-level functions - systemically adjusting analytics operations in real time. For larger strategic challenges, human guidance would be required to deliver the optimal assessment.

To initiate the creation and deployment of this architecture, high-impactful projects and PoCs will be mounted to build the project's momentum. These initial small steps are each designed to increase climate assessment capabilities and architectural proof points, and in turn, provide practical value to citizen scientists and developing nations.

The following projects will lay the foundation to arrive at the unified end state. By initially targeting a tight set of high-value deliverables, and sequentially building this framework according to open source principles and best practices, the end result will be a significant advance in climate analysis, and AI platform capabilities in general.

Project Activity

We have established the objectives of the overall initiative and created individual workstreams. We have also identified and assigned development personnel, which include Data Scientists, Application Developers, Systems Architects and IT Ops personnel.

Workstreams include:

- Central Analytics / Intelligence: The overarching AI intelligence and reporting instance at the top tier of data correlation and analysis. A self-identifying digital asset and AI model catalog is the first phase of development.
- Autonomous Middleware and Messaging: The self-identifying connective fabric of this framework, with related security capabilities. A PoC is planned in partnership with Stanford SLAC.
- Data Governance and Operational Guiding Principles: The underlying operational principles of this system, with a focus on humanitarian/ethical objectives and outcomes.
- Stanford SLAC LCLS PoC: An X-ray research AI framework that supports Stanford LCLS objectives, and helps validate the baseline Neurosystem architecture.
- Bee Population / AI Acoustics PoC: IBM Research has provided an AI software donation for Acoustics analysis, that the community is now using to analyze hive health and understand bee population declines. This is using low cost technology to assist developing nations.
- Stanford SLAC/HTM PoC: Given the biological parallels, Stanford SLAC has initiated a study of HTM for AI anomaly detection. The community is interested in the general intelligence promise of this approach.

Founding Members

This effort will first involve a team of core participants to achieve the initial codebase and participate in forum discussions. The following list includes the firms associated with our current and original founding members:

America Movil

Dell

Equinix

Ericsson AI

Harvard Analytics

IBM Research

Intel

Kove

Meta

Microsoft Azure

Penguin Computing

PerceptiLabs

Red Hat

Reliance Jio

Seagate

Stanford SLAC

UC Berkeley Data-X

Verizon Wireless

Yahoo!

Conclusion and Future State

An overarching singular framework is the final phase of evolution in terms of AI and enterprise IT. And given the widespread distribution of climate data, and real time feeds via satellites and various IoT devices that lack true collective integration, we feel this is a research opportunity that has yet to be addressed. We feel a collective AI infrastructure and central analysis approach will also be of great interest across various scientific communities and enterprise verticals.

Yet, there is a future use case that we plan to develop as well. This climate system would begin by extending across every geography, and reach into the most remote regions of the world. From the earliest days of this community's inception, we foresaw the use of this infrastructure for a larger egalitarian purpose, as this neurosystem eventually evolves into a *connective, intelligent fabric between humans and their ecosystem*. It would capture sensor and other data on a planetary scale to correlate, load balance and autonomously offset the effects of overpopulation and pollution, and lead to more efficient natural resource/species preservation.

Development Commitment

As this is a volunteer organization, the development commitments can be considered variables based on each member firm's resource availability. However, resourcing should include an executive sponsor and a dedicated technical resource or resources (architect, developer, or both). Again, these are part-time activities, given the volunteer aspect of community engagement. As well, industry leaders with resource constraints can assign technical observers to contribute their guidance and best practices as part of the regular meetings.

Community involvement can span a variety of activities. This includes any of the following: community management, marketing support, general participation via sharing current challenges and historical knowledge of architectures, operational and technical project management, outbound communications, and of course, code development and documentation. A wide variety of talents and interest areas are welcome.

In return, participants can expect a number of benefits – an open exchange of ideas on the latest AI research, new areas of enterprise development, leveraging and incorporating solutions built on an open-source framework, and discovering commonality with fellow community members and their areas of endeavor.

Please contact Bill Wright (bwright@redhat.com) and Ganesh Harinath (ganesh.harinath@fiduciaai.com) for any questions regarding community membership.

Details on PoCs and the Focus Groups follow below.

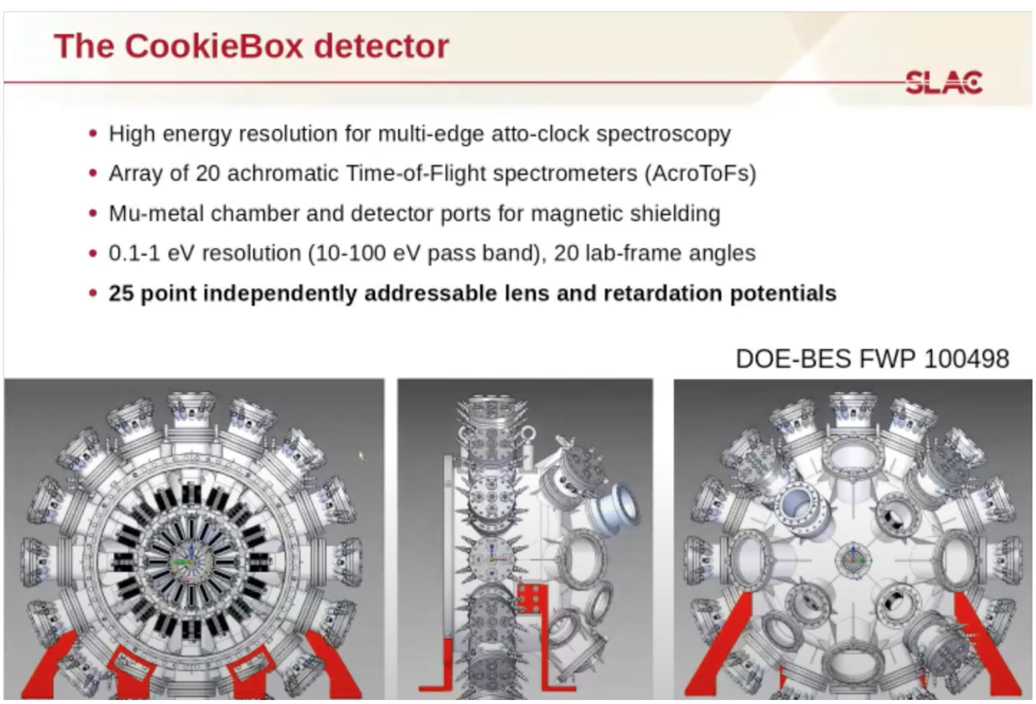
Workstreams

Stanford SLAC Neurosystem PoC

Stanford SLAC LCLS (Stanford Linear Accelerator Laboratory - Linac Coherent Light Source) has invited our community's participation in support of a Proof of Concept for the LCLS Cookiebox detector.

It will use FFT algorithms to quickly process image data at the edge of the detector, and then forward the data to more powerful NN capability in a core data center. Images currently arrive at a rate of 120 images a second per spectrometer, and this will increase to one million images per second (20M/sec total rate) over the next two years.

AI/ML capability will need to be created to quickly sift through this flood of data in real time, and search for anomalies that indicate new discoveries in physics. This is a concise example of a Neurosystem topology - real-time data processing in multiple nodes, connected to a central cross-correlation intelligence. This can serve as the first step to a broader-based intelligent system, as this example is using analytics in edge nodes connected to a core data center for deeper neural network discovery and analysis. The same architecture pattern can be replicated in a large scale climate monitoring network – and a wide variety of industries, including telecommunications, manufacturing, financial services and health care.



Acoustics, AI and Global Bee Populations

One of the gravest ecological challenges of our time is the potential loss of bees, and the shock that would have to earth's ecosystem and global food supply. Bees enable functioning ecosystems and support human habitation. Pollination from bees is responsible for over 35% of global food crops - staples like coffee, cocoa, tomatoes, almonds, apples and blueberries depend on bees for production.

Unfortunately, bee populations are in decline - for example, beekeepers across the United States lost over 45.5% of their colonies in 2021. This is significantly higher than the 20% loss considered normal, and only proactive bee colony management by humans has prevented a catastrophe to date.

A number of factors may be the cause, including pesticides, pollution, climate change, intensive farming and the invasive Varroa mite. But researchers suspect there may be other factors at play that haven't been identified yet.

Enter acoustics and AI. Honeybees generate a variety of sounds that serve as a form of communication inside the hive. And it's not only the frequencies that determine the meaning of these sounds, but there is an underlying structure to these signals. Research has shown that the health and wellbeing of a hive is correlated to the acoustics of that colony - and identifying and tracking these signal changes could help navigate the colony's issues, improve the well-being of the bees in real time, uncover new modalities of communication, and help understand all the related factors.

With science and technology, we can hopefully uncover evidence of what is killing off the bees, and offer solutions and tools that can help with the preservation of this crucial species. We also will focus on providing free open source software and datasets to others that want to engage in collecting data on bee health. In addition to developing nations, arming the citizen scientist and hobbyist beekeeper is also crucial, as this is a challenge that requires as many participants as possible.

IBM Research has arranged for a generous code base donation to the Enterprise

Neurosystem community - Acoustics AI for Enterprise. Given the community charter, this project is an ideal use case to highlight the benefits of AI analysis of these waveforms. And this application is just one of thousands that would be part of an eventual global-scale neurology, helping to mitigate the effects of humanity on our habitat.

Secure AI Integration Fabric

AI application networking and security presents major challenges due to the risk of leaks of massive data sets, which data owners could be responsible for. This is magnified in multinational efforts, since data from multiple countries needs to be used, but it may not even be legal to have a single "data lake" holding the combined information. Recent open source innovations in the use of Layer 7 Application networking techniques allow us to address these security concerns in AI model connectivity, and have been carefully examined by some of the largest financial institutions. By tying applications together in a temporary network, and disengaging all connectivity when the job or requirement is complete, security is greatly enhanced due to compartmentalization of data and time-locked access.

We have initiated a new proof of concept with the Stanford SLAC AI lab, and will be conducting a baseline proof of concept to test the functionality of this new technology. We will also apply an autonomous policy engine to this integration fabric, and we are exploring the use of a related platform developed by community member Dinesh Verma at IBM Research.

This fabric can then be used to securely tie AI models together in a large-scale multinational framework, with the strongest networking security currently attainable. This open source development could allow developing nations to securely share innovation while preserving sovereignty and economic competitiveness by dynamically adapting running AI models in environmental and digital environments with intermittent connectivity.

Hierarchical Temporal Memory Analysis

The cross-correlation engine is one of the final stages of development of the Neurosystem framework. While development is currently underway on the foundation layer (self-identifying digital asset catalog), the community wants to explore potential AI approaches to upper tier cross-correlation and anomaly detection. The community has partnered with Operate First to do a study of HTM (biologically-based software architecture), in partnership with the Stanford SLAC AI lab effort related to fusion energy, to test its capabilities for use in higher order multi-sensor correlations for anomaly event forecasting.

Community Leadership

Governing Board:

Chair: Bill Wright, Head of AI/ML and Intelligent Edge, Red Hat
Vice Chair: John Overton, CEO of Kove
Technical Committee Lead: Dinesh Verma, CTO Edge at IBM Research
Government Representative: Ryan Coffee, Sr Scientist, Stanford National Accelerator Lab
Financial Services Representative: Vishnu Hari, PM at Meta AI
IT Vendor Representative: Ganesh Harinath, CEO of Fiducia AI
Telco Industry Representative: Tong Zhang, Principal R&D Engineer at Intel, Network AI

Working Groups:

Central Intelligence Working Group – Chair: Dinesh Verma, CTO Edge, IBM Research
Physical Image Analytics – Chair: Ryan Coffee, Sr Scientist, Stanford National Accelerator Lab
Intelligent Connectivity Working Group – Chair: Sanjay Aiyagari, Principal Architect for Telco, Edge and AI, Red Hat
Telco Working Group – Chair: Ravi Sinha, Director of Technology and Solutions, Reliance Jio
AI-based Signal Processing Working Group – Chair: David Wood, Software Engineering, IBM Research