Community Overview

Introduction and Mission

The world of AI model development is evolving at a rapid pace. Today, thousands of AI models can be deployed in a multinational corporation or organization, both developed in-house and supplied by a wide array of vendors. This can lead to important discoveries and insights, but it also presents a new issue - what framework is in place to integrate all these instances, and to correlate and present their collective findings?

We propose the next stage of AI evolution - the Enterprise Neurosystem. A unified AI intelligence framework connected to every area of a business or organization, gathering and integrating millions of real-time and historical data points for actionable reference. This framework will autonomously analyze and manage day-to-day operations while providing observations/guidance to human management.

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 are quite similar 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. It will enable a deeper and more complete level of AI awareness and perception - building on the vision of a large-scale collective framework, and turning it into a single federated AI instance.

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, to maximize the use of popular models and minimize unintended effects like bias.

The consortium is most importantly, a community.

The Challenge / The Opportunity

The majority of AI applications are built upon a few dozen core frameworks that are repurposed for a variety of analytics requirements across a dizzying array of verticals and functions in a multinational corporation, governmental body or in service of a global initiative.

The growing list of AI-assisted corporate functions includes, but is not limited to, supply chains and logistics, autonomous drones (infrastructure inspection), IT operations and security, financial systems and forecasts, manufacturing lines, natural language-based customern 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. In more technical realms the functions include 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 many more.

Regardless of the vertical use case, these AI 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.

Differences in vendor AI development philosophies have also produced multiple variants of application architectures. As a result, customers can field AI applications with wildly different degrees of complexity. A lack of application-level cohesion can also contribute to reduced bias assessment capabilities. Based on the broad experience of data scientists, bias is an unfortunate and very 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 quite different, and (b) system-wide insights continue to be incomplete, due a larger integration challenge. Non-integrative analysis and self-reflection also leads often to inaccurate assessments - just as they do across humanity.

If there was a way to fundamentally address these issues, and truly unify and cross-correlate all of these elements, we could achieve transformative infrastructure efficiencies. Over time, a far more accurate and powerful analytics engine would be the end result.

Solution

We propose a unified analytics framework that spans all aspects of corporate and organizational operations. Similar to 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 business or organization – and autonomously regulate and optimize many day-to-day functions. This capability will increase efficiencies in all areas of IT and operations, and will provide an operations interface of unparalleled depth and clarity.

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 an enterprise or organization, tying together areas of operation via corporate networks, manufacturing and logistics systems, mobile networks, the web, IoT sensors, satellite systems and IT software systems and databases.

Data integration will be fine-tuned by users to reframe the deployment and target additional areas of analysis. They will teach the system to understand new areas of focus and further enable the system to be self-directed. Eventually, this AI analytics system will run autonomously by maintaining a communication and analytics web across all aspects of the business or organization.

In its end state, an overarching intelligence will draw in and cross-correlate the data from the many areas of operations, autonomously providing management for lower-level functions – while systemically adjusting analytics and operations in real time. For larger strategic challenges, human guidance will be brought in to deliver optimal decision-making capability.

For businesses, the Enterprise Neurosystem framework will result in improved service delivery, deeper customer insights, supply chain and delivery efficiencies, data center operations and environmental savings, security optimization, and more. The resulting cost savings and increase in revenue will be significant. For governmental and other scientific applications, the federated intelligence and insight will be invaluable.

To support the creation and deployment of this architecture, high-impact proof of concept projects are being mounted to build the community's momentum. The following projects will lay the foundation to arrive at this unified end state.

Project Activity

We have established the following 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 Proof of Concept:**
  An X-ray research AI framework that supports Stanford LCLS objectives, and helps validate the baseline Neurosystem architecture.
- **Bee Population / AI Acoustics Proof of Concept:**
  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 Proof of Concept:**
  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.

Project Genesis

Four years ago, some of the project's founders 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.

It became clear 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.

It was clear that a similar evolutionary framework could emulate this neural topology within the IT domain and enable a deeper more complete level of AI awareness and perception. A large-scale collective framework could be built, turning countless elements into a single federated AI instance. This federation could open up the larger possibilities of AI in any enterprise, regardless of its focus or business vertical. This framework, when employed, could provide the leadership of an organization with a window into many aspects of its business functions or information in real-time by identifying trends, predicting challenges, and delivering guidance on potential outcomes and solutions.

Its architecture would include high-value AI models dedicated to specific functions, to drive interest and adoption. Concurrently a top-tier interpretive and reporting intelligence would be built – 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 management team.

The framework would be assembled and nurtured by a research community that would welcome corporations, government agencies, academics and the private sector in an open forum. It is important that an architecture of this scale and critical importance will 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.

Climate Application 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.

Founding Members

This concept has attracted 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!

Community Members and their Parent Organizations

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 our Proof of Concept Projects follow below.

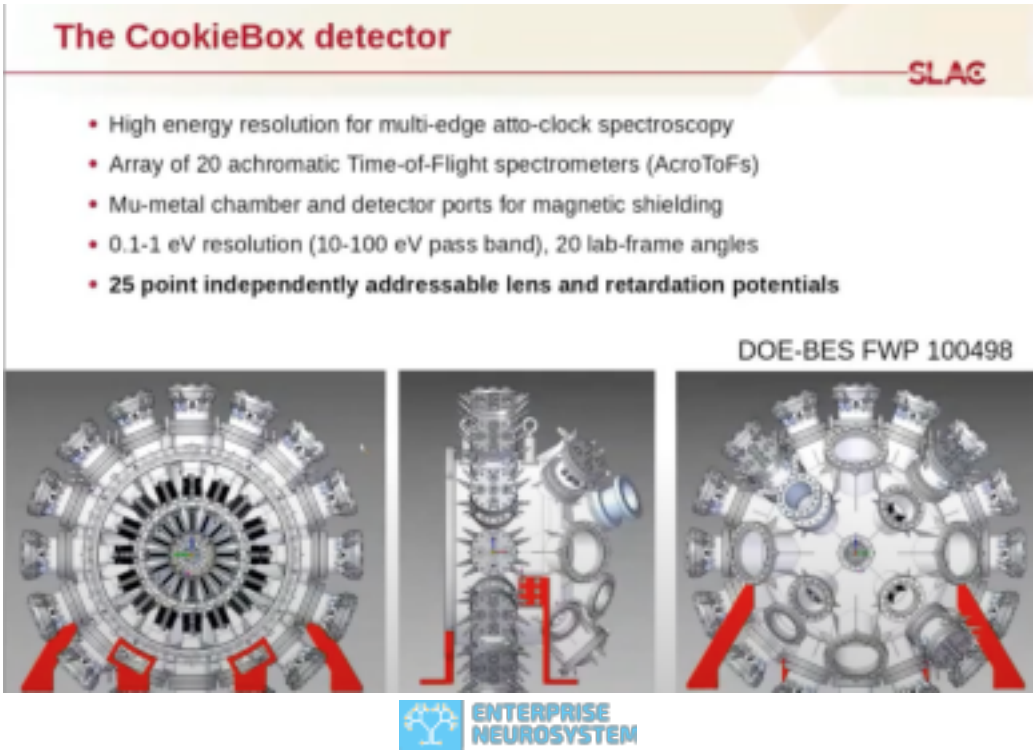
Proof of Concept Projects

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 – Chair: David Wood, Software Engineering, IBM Research