



# The Enterprise Neurosystem

The connective intelligence of Enterprise AI

Community Vision & Overview

# The Enterprise Neurosystem

## Project Origins

The AI revolution is at an interesting inflection point. Billions in investment, thousands of startups, and models are being deployed in vast numbers across the enterprise. These are highly complex software units of intelligence – but without collective consciousness, or a pathway to greater meaning.

Networks and related datacenters possess an interesting parallel to the human neurosystem. Core functions are managed autonomously (heartbeat, chemical levels, oxygen assimilation, etc.) while others require higher order reasoning (relationship management, course correction, etc.). And while humanity has been subconsciously creating the IT construct in its own image, the final piece – AI/ML – has advanced to a point where the completion of this model is now possible.

There has been a significant wave of bespoke AI/ML applications deployed across the enterprise. The majority of these applications are either created in-house or supplied by vendors, and are purpose-built using a handful of programming frameworks.

This enthusiastic support for AI/ML use cases has inadvertently created a number of complex issues:

- **Model Standards** – There are millions of models now being deployed, but without a common set of standards or best practices. If stronger model standards were implemented, this would answer growing regulatory concerns, and create stability for the developer community.
- **Cross-Correlation** – The vast majority of production models have been deployed as individual applications. There is no central / federated intelligence to make sense of incoming data across the enterprise. If this capability was enabled, a deeper analysis of both real time and historical data would be possible, creating a union of AIOps and predictive Business Intelligence.
- **Model Infrastructure** – The supporting infrastructure for AI/ML should be easily shared, readily transparent and open to community development. It should allow all model formats to integrate in a reliable environment that can be easily managed, explained and repaired.
- **Model Triage** – It is difficult to triage and explain model behavior when data inaccuracy or a general corrective event arises, particularly if they

are created in a closed and proprietary manner. Open and transparent model development, with an eye to standardization, will be required to comprehensively address this area.

- Bias – It is difficult to manage bias effectively if models are independently generated by hundreds if not thousands of different teams, without common oversight. A cross-correlation framework would help address this, but only if the models themselves are easily corrected.
- Synthetic Data Generation – When production data is required for training, it must be anonymized and yet accurately replicated, given regulatory challenges in terms of data usage.

These challenges are daunting, and seem separate and distinct in many respects. But they can certainly be addressed more readily if all elements are integrated within a single transparent architecture. One that could eventually manage the lower order functions of a company, autonomously deploying and integrating AI models across every area of operation.

An evolutionary framework with standardized models will be required, featuring a reliable and transparent neural topology. The primary goal - to enable a deeper and more complete level of AI capability for the intelligent enterprise.

## **The Challenge - and the Opportunity**

The majority of AI/ML applications are built upon a few dozen core frameworks that are repurposed across various business verticals. In general, each one is created with a singular purpose. This is now leading to AI model sprawl - a dizzying array of fragmented functions, particularly in a multinational corporation.

The growing list of AI-assisted functions include, but are not limited to, IT and network operations, visual ID for building security and manufacturing, supply chain and logistics analysis, financial systems and forecasts, NLP-based customer interaction/management and insights, facility and equipment maintenance, human resources, legal functions and contract reviews, insurance and risk analysis, regulatory frameworks, taxation analysis, and so on.

Many of these AI supported functions are separated by individual vendors and their siloed applications. This reduces the business value of AI, given the higher potential of collective inference - visibility is limited to a small subset of the total data available. This also creates firewalls of functionality, and again, reduces the possibility of large-scale insight.

All this is further compounded by differences in vendor AI development philosophies. They each produce multiple variants of AI and ML architectures, and

as a result, customers will field AI applications with wildly differing degrees of complexity, interpretive capability and insight.

Consequently, this fragmentation also means bias becomes an issue. A lack of application-level cohesion leads to reduced or conflicting bias assessment capabilities. Given the broad experience of data scientists, bias is both unfortunate and a real challenge.

This causes fundamental AI operational issues for any enterprise, as a) the various vendors and the underpinnings of their NN architectures are vastly different, and b) unification of interpretive insight in terms of inference, output and bias analysis becomes an integration challenge.

If there were a way to eliminate these issues, and truly unify and cross-correlate all of these elements, then the deepest possible insights would be achieved. Real time and historical data would be fully utilized, and a more linear and efficient business would be the end result.

## **Solution**

We propose a unified analytics framework that spans all aspects of corporate operations. Much like the neurology of the human body, it is a series of interconnected AI/ML functions with a central intelligence, tailored to permeate and assess every aspect of the business, and to autonomously regulate and optimize day to day functions.

It was evident that the final outcome would be to connect the business functions of an enterprise into a single unit of overarching intelligence - the Enterprise Neurosystem. This capability would not only increase efficiencies in every area of network and business operations, but it would enable a corporate operations interface of unparalleled depth and clarity.

It will be a connective fabric that takes existing AI models and integrates them in a unified manner. This in turn preserves existing AI investments, and increases the power of broad-scale analysis. We will also provide individual models that may be missing for various enterprise functions. But the objective is to eventually provide a single AI instance spanning the enterprise, connecting every outlying AI model, drawing together their historical and real-time findings, and conducting analytics and pattern analysis across all company operations.

It will act as a single integrated intelligence, providing the C-suite with a crystal clear window into all aspects of corporate operations. It will autonomously identify

and predict related challenges, and ultimately provide optimal guidance to reach the best solutions.

It will also be built upon commonly available and reliable open source platforms, including model creation and management, storage, middleware and container technology. Again, the transparency of an open source project is a key benefit, when creating a system with these capabilities.

The resulting savings and increase in revenue would be significant. For example, AI-driven predictive maintenance will be one area of focus, and this portion alone is estimated by McKinsey as increasing company cost savings at an equivalent rate of 2% of annual revenues. It also means increased network efficiency and bandwidth, deeper customer insights, increased supply chain and delivery efficiencies, data center operations and environmental savings, facility security optimization and much more.

By targeting a tightly defined set of high value deliverables, and sequentially building this framework in an open community fashion, the end result is a significant advance in F500 business operations.

## **Path To Production**

The initial program will start with a small series of open source projects designed to tackle a few fundamental issues, each with a clear path to value. These will be linked by a common communications and AI/ML framework, which would be developed concurrently. Over time, this effort would extend its awareness across the enterprise, threading in areas of operation via the web, corporate networks, manufacturing and logistics systems, and IT software systems and databases.

The process would then be fine-tuned by corporate users, to reframe the deployment and target additional areas of analysis. They would teach the system to understand these areas of focus and further enable the system to be self-directed. Eventually this AI analytics system would run autonomously by maintaining a communications and analytics web across all aspects of the business. As the Neurosystem will reach across various business units and subsidiaries, it will need to be infrastructure agnostic, and thus this framework will be fundamentally based on common open source platforms and container infrastructure.

There is also a differentiation of purpose for Edge and Core infrastructure environments, and models with lower latency requirements would be positioned accordingly. Load balancing of the model training function would be enacted via connectivity between the Edge and Core, to ensure resource-intensive training will take place in the Core when appropriate. Container management will be a key part

of this requirement, and related integration and communication with the Neurosystem framework will enable autonomous deployments and upgrades.

A single view console would also be created in the early stage of the project, to report the status, output, accuracy and potential drift of all models in real time. And as new data arrives for model training and retraining, both autonomous and human course correction will be enabled at a higher frequency as needed.

In its end state, an overarching intelligence would draw in and cross-correlate the data from the many areas of company operations, and autonomously provide load balancing for lower level functions - adjusting the corporate “system” in real time. In terms of larger operational and strategic issues, human management would be brought in to create a stronger network effect (human/machine), to deliver the best decision capability. This could take the shape of a hologram advisor, sitting with management in human form, and relaying the Enterprise Neurosystem’s findings.

## **Open Source Community Framework**

This will be built within the context of a formal community, one that would welcome enterprise and partner participation in an open R&D forum. We realized an architecture of this dimension and implication should reside in open source for a variety of reasons – factors like code base transparency, ongoing community input, and cost efficiency would serve as the primary drivers.

This community project will also include tightly defined high-value AI functions to drive community interest and user adoption. The first projects would help solve a few of the fundamental issues facing the enterprise community: synthetic data generation, and rapid visual identification scenarios.

As an open source company, Red Hat is delighted to participate in this new community endeavor. Our staff of community professionals and the Red Hat CTO Office have been consulted on this proposal, and basing these development activities on a well-established governance model and open source legal framework. The software will be licensed under Apache 2.0, with DCO code attribution.

## **Founding Partners**

This effort was initiated in August 2020, and involved a team of core customers and partners to start work on the initial architecture and code base. The founding members represent the following companies:

America Movil

Equinix

Ericsson AI

EY AI Engineering

IBM Research / Watson

Intel

Kove

Lambda Labs

Penguin Computing

PerceptiLabs

Red Hat

Seagate

Stanford / SLAC

UC Berkeley / Data-X

Verizon Media

### **Initial Workstreams:**

- 1) Projects 1 and 2: The first two projects will be comprised of achievable AI/ML objectives that will provide functionality that will support the Neurosystem framework.
  - i) Synthetic Twin/Data Framework: Verizon Media
  - ii) IoT/Visual Analytics For Scientific Analysis: Stanford SLAC
- 2) Focus Groups: Concurrent to Phase 1 and 2, two focus groups will be formed. As time goes on, these focus groups will turn to active development, once Phase 1 and 2 are complete.

- i) Middleware and Messaging: The connective aspect of this framework, and the related encryption and security aspects.
- ii) Central Analytics: The overarching intelligence center, that will apply neural network capability at the top tier of data correlation and analysis.

### **Focus Group 1: Middleware and Messaging**

This focus group will meet concurrently during the first phases of development, to plan the connective/communicative fabric between AI models across the enterprise. The architecture will be designed for low latency, high security, and will readily integrate AI models. Middleware best practices will be reviewed and utilized, and open source tooling will be applied wherever possible. The framework will also be modular in nature, and allow the incorporation of existing AI model investments, either internally developed or vendor-sourced. Efficiencies observed in production environments will be adopted, and novel approaches regarding deployment, autonomous integration and encrypted messaging will be explored. Eventually the group will collaborate on a reference architecture, development milestones will be established, and the initial PoC projects will be integrated.

### **Focus Group 2: Central Analytics**

The culmination of the project will result in a central body of perceptive intelligence. This may be represented as a collective of existing neural network frameworks, or an entirely new approach to analytics. This overarching framework will act as a recommendation engine and an autonomous operations manager. It will have a primary charter to enact financial and logistics load balancing across all business units.

Individual representations of this management intelligence may also be distributed in various subsidiaries and geographies. These instances will integrate with the larger scale analytics intelligence - similar to the human prefrontal cortex - that resides in corporate headquarters. This intelligence will be tasked to provide the C-suite with a real time view into all company operations, and deliver recommendations and autonomous responses on a wider scale. Again, the patterns of this initiative are mapped to the human neurosystem, and incorporate the applicable elements of that architecture.



## **Development Commitment and Resources**

The development commitments can be considered variables based on each firm's resource availability. However, resourcing should include a primary project lead or contact for each partner, and dedicated technical resources (architect, developer, or both). These can be part-time or full-time assignments. As well, industry leaders with resource constraints can assign technical observers to contribute guidance and best practices as part of the regular meetings.

Community involvement can span a variety of activities, depending on individual interests. This includes any of the following - active participation in group discussions, sharing current challenges and historical knowledge, whitepaper composition, code development, and offering ideas regarding new solutions. 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 solutions built on an open source framework, speaking at conferences and events on behalf of the community, and discovering commonality with fellow community members and their areas of endeavor.

## **Practical Applications and Future State**

An overarching framework is the final “unified field” phase of evolution in terms of AI and the enterprise. We are now taking this shared vision and launching development as an open community, and inviting a wider group of partners. To build a community upon this proposal is a natural evolution of this concept.

The enterprise would achieve an intelligence and self-awareness that would enable a new age of company operations. It would provide an unparalleled competitive advantage for the early adopters of these frameworks.

Corporations would autonomously share mutually beneficial information between central AI instances, and create vastly more powerful alliances and partnerships. Complimentary products would be quickly aligned and developed for optimal value and reliability. Joint sales opportunities would be recognized and instantly enabled with rapid AI-driven agreements. The list goes on endlessly.

And the Enterprise Neurosystem framework can be applied to any number of use cases. In terms of healthcare, the recent COVID-19 pandemic taught us that the rapid cross-correlation of data can produce vaccines at an exponentially faster rate of speed. And eventually, a hyperdetailed digital twin of the human body – with

individual cells, genetic variables, chemical characteristics, and structural data - could be linked as a template for developing new drugs and restorative therapies.

There are other fascinating future applications. Over time, we foresee the community using this technology for a more egalitarian purpose, as this neurosystem could readily evolve into a connective, intelligent fabric between humans and their environment. It would capture sensor and other data - on a planetary scale - to correlate, load balance and offset the effects of overpopulation and pollution. Ultimately, it could lead to more efficient natural resource consumption and species/ecosphere preservation.

It could also be used beyond the confines of Earth, in terms of detecting extraterrestrial objects with deteriorating orbits, or incoming asteroids. Large frame, wide area AI analysis and tracking would assist in effective identification, and deliver course correction recommendations that could avert future disasters.

And what is most compelling - this will all transpire in the best traditions of open source development. We welcome you to join us.