



Third Evolution, Inc.



PROPOSAL ABSTRACT: The Knowledge Molecule

FOR: IARPA-BAA -16-07

By: Third Evolution, Inc.

CAGE #: 7Q781

Mike Baciewicz

602-510-7505

[Mikebaciewicz@thirdevolution.us](mailto:Mikebaciewicz@thirdevolution.us)

[www.thirdevolution.us](http://www.thirdevolution.us)

**INNOVATIVE CONCEPT:** “The Knowledge Molecule”™ by Mike Baciewicz of Third Evolution, Inc.

The Knowledge Molecule is an Artificial Intelligence (AI) engine designed to allow organizations to manage decision making at an optimized level while dealing with the overwhelming exponential growth of available data. This unique application allows for the optimization and automation of data collection, compiling relevant data and automating or recommending first, second and third tier solutions. The Knowledge Molecule is a Multidimensional AI Engine for Correlating Massive, Mixed Parameterized and Non-Parameterized Abstract Data into Quantitative, Correlated and Optimized Best-Case Solution Outcomes; We can provide 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> order extractions which can yield significant inferential predictive outcomes for very complex, time extended data masses otherwise beyond the reach of advanced human cognition. This allows for significant decrease in the time to sort, correlate and analyze data along with improved real-time decision making.

Additionally, the “Building Block” engine can be implemented singularly or combined with itself configured for

- A. Cognitive Computing
- B. Associative Computing
- C. Predictive Computing
- D. All of the above, as a larger molecule.

**DETAILS OF INNOVATIVE CONCEPT:**

Third Evolution provides the combination of Process Engineers, Integrators, Automation and AI Engineers. As a firm, we cannot see, and have not seen, a successful implementation of advanced analytics with the outcome resolution required to be a useful 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> order engine by anyone; and we feel it is because of the dearth of firms with the complete combination of the skills above integrated into a singular mindset for the outcomes required.

Our engines are structurally simple, yet complex in design and integration. The “simple” design equals computationally fast. Additionally, our engines are like Lego Blocks: they can be “snapped” together to build “super-engines” or “super molecules”. The basis for the molecular design is because the solution behind the engines strength can only be achieved if one does not think like a human. Instead, just as one learns other

languages for ideal communication capabilities for transferring ideas to other humans with clarity and specificity (Germanic languages are known for being less capable of emotionally descriptive concepts, for example, as compare to the Romance languages). Similarly, if one embraces the notion that regardless of its level, modern mathematics and Newtonian Physics at some point are detrimental to a significant portion of solutions to advanced problems. What if, for example, the only way to solve a set of problems is tied to the 3<sup>rd</sup> harmonic (or inter-harmonic) of a waveform? We can generally agree that if the previous statement were true, we would “never get there” and that is still within modern mathematics!

We must “stop thinking like a human,” if we are to solve the developing challenges being created with the massive volumes of data from disparate sources arriving at exponential speeds, landing on human desks to be analyzed. The Third Evolution Molecule is designed to “Think” like an advanced machine to capture, sort, correlate and analyze enormous volumes of data and then “present it” to human analysts and decision makers to apply their knowledge, expertise and experience in an optimized fashion. Our team have the skills and tools to overcome these challenges in many of the areas of interest of IARPA listed in BAA-16-07.

**TECHNICAL PATH TO REALIZATION:** As previously indicated, our approach structurally simplifies the path to the solution, but it is not easy. It is also a compound solution with a handful of innovative approaches we developed in the very complex space of real-time (and predictive) environment of 24/7/365 industrial process controls solutions, which we have implemented personally for 30+ years; 17 as a corporate entity.

**REQUIRED CAPABILITY #1:** First, we must address the inputs, or data feeds/sources. For industrial work, where we have 1 project with over 600 unique industrial sites each with hundreds of dissimilar points being monitored (America’s Natural Gas Highway) we have – from a top down look - the layers of corporate networks and all associated security models, which are split into regional networks, which are split into localized plant networks, which are then split into sub-nets and fractional sub-nets which then ultimately traverse down to unique data inputs from the desired singular root point (regardless of its constitution e.g. analog, digital, calculated, inferred, etc.). The only way to be able to confidently assimilate digitally characterized sources is to know that you have done it thousands of times without fail. Further, if the data currently exists only in an

analog format, we have been writing connectors/API's for years. We possess this skill with 100% mission success over 17 years.

**REQUIRED CAPABILITY #2:** Second, we must be able to decompose the process(es) at hand to all roots. Then, from the decomposed root structure of the entire process(es) we must be able to recognize and understand how the re-composition of the process with full optimization can be achieved. And this is only possible through a holistic understanding of the entire super-process from a material-balance perspective. This is a process engineering approach absolutely required for any industrial process that needs to be optimized. It has complete transferability to data-only realms, if one can transition mentally to a relativistic perspective. Hence, our mantra for these types of applications require us to, "Decompose the process, absolutely understand what can be understood, optimize and recompose the process from the root up, automate the process, and then optimize the automation, and then tune the outcomes empirically (against known historically accurate facts), lock down the logic and release for use."

The outcome of this work effort results in (from process engineering accepted best practices) a set of detailed documents (trade secret/IP) that capture the nature of the process in English, symbology and other in-house mechanisms that allow for a complete structural capture of the entire process ready for conversion to our logic engines (trade secret/IP).

**REQUIRED CAPABILITY #3:** Conversion of Step 2 to Logic (ALL Trade Secret/IP)

In this crucial step, we have developed methods to create the logic engines in various structures to allow the incoming data feeds to be processed:

- A. Immediately
- B. 2<sup>nd</sup> order, multiple derivative capability based on value and need
- C. 3<sup>rd</sup> order, multiple derivative capability based on value and need

The Multi-Dimensional AI engines we have are then:

- A. applied to the data,
- B. mixed with circumstantial need(s),

- C. correlated for hierarchical value,
- D. identifying the largest value hinge-points,
- E. presenting the correlated outcomes in hierarchical format tied to the circumstances of value identifying the largest to least contributive data for human review and confirmation.

This allows us to identify potentially newly discovered relevant data on an ongoing basis and allows the system to continue to be iterative and improve over time.

**REQUIRED CAPABILITY #4:** Present the data in whatever manner is required, to wherever it is required in the timeliest manner possible.

Again, due to the SCADA component of our base capabilities, deploying information in the timeliest manner possible based on known value-deployment schedule, the alerting of the appropriate information can be complex but once mapped out with all known recipients and needs analysis, we can present the solution outcomes (be they actionable, informative, approval gates before commencing to another step, etc.) in any manner desired. E.G. Dashboarding, mobile devices, interstitial launching of other applications, etc.

**REQUIRED CAPABILITY #5:** Using our engines (trade secret/IP) combined with client identified outcomes (Best case – get us here whenever you can, Worst case – keep us away from here whenever you can, Conditional cases and any other circumstantial value outcomes) our 2<sup>nd</sup> and 3<sup>rd</sup> order derivatives will gather data for predictive analysis with current-state magnitude of potential likelihood in percent for identified outcomes. This can then be applied to an automated solution or optimized recommended solutions.

#### **MILESTONES FOR PROGRESS ALONG THE PATH**

For the IARPA needs, the following milestones provide insight into where the project is from a percent complete perspective:

- A. Project Description
  - a. Identification of all data sources with protocols and network paths
  - b. Identification of all possible data recipients (hierarchically)
  - c. Identification of all known/desired outcomes (hierarchically)

- B. Complete Detailed Document Exchange (Varied, case by case, but tied to any C3 constituent components)
- C. Complete Detailed Interviews with Subject Matter Experts (client side resources)
- D. Project Process Narrative submittal, review and approval (exactly what we will do, how it will be done in technical details outside of the Trade Secret/IP envelope)
- E. Project Execution
  - a. All API's functional for data ingestion
  - b. All outputs functional to all tested recipients
  - c. All 1<sup>st</sup> order outcomes test cases compared against known historical outcomes (tuning phase for each module type)
  - d. Completed Validation and Verification prior to release for use
  - e. Acceptance and Sign Off

**ESTIMATE OF RESOURCES TO ACHIEVE THE PROPOSED OBJECTIVE:**

The amount of resources needed is tied to the number of modules undertaken at a given time. For 1 module, the team size would be dependent on the specific item(s) of interest to be solved. (e.g. Anticipatory Intelligence for Analysis and forecasting of rare events, OR Analysis for Dynamic means to expose and enhance unprocessed data and interim or alternative analytic products across domains or represent cascade effects of new information on existing models.)

For a minimum team size, we would have at least 4-5 advanced staff at any time working in their respective areas of deliverables. On the client side, the number would be specifically tied to the known subject matter experts for the areas of focus. The larger the area of focus the more clients side staff would likely be engaged.