NGCRC Project Proposal
Intelligent Autonomous Systems Based on Data Analytics and Machine Learning

August 28, 2017

Prepared for
The MS Sector Investment Program

Prepared by
Bharat Bhargava
CERIAS, Purdue University
Table of Contents

1 Executive Summary ........................................................................................................... 3
  1.1 Abstract.......................................................................................................................... 4
  1.2 Graphical Illustration .................................................................................................... 4
2 Description of Project ........................................................................................................ 6
  2.1 Statement of Problem .................................................................................................... 6
  2.2 State of Current Technology and Competition ............................................................ 7
  2.3 Proposed Solution and Challenges ............................................................................... 8
    2.3.1 Cognitive Autonomy ............................................................................................... 9
    2.3.2 Knowledge Discovery ............................................................................................. 11
    2.3.3 Reflexivity of the system ....................................................................................... 14
    2.3.4 Trust in Autonomous Systems .............................................................................. 16
  2.4 Distinctive Attributes, Advantages, and Discriminators .............................................. 17
  2.5 Tangible Assets to be Created by Project ..................................................................... 17
    2.5.1 Software ................................................................................................................. 18
    2.5.1 Documentation ....................................................................................................... 18
  2.6 Technical Merit and Differentiation ............................................................................. 18
3 Project Milestones ............................................................................................................. 19
  3.1 Statement of Work ........................................................................................................ 19
    3.1.1 Cognitive Autonomy and Knowledge Discovery ..................................................... 19
    3.1.2 Reflexivity of the system ....................................................................................... 20
    3.1.3 Trust in Autonomous Systems .............................................................................. 21
    3.1.4 Integration with NGCRC and NGC IRAD projects ................................................. 21
  3.2 Milestones and Accomplishments ................................................................................ 22
4 Project Budget Estimate .................................................................................................. 23

List of Tables

Table 1: Executive Summary .................................................................................................. 3
Table 2: Machine learning techniques for outlier/anomaly detection .................................. 14
Table 3: Milestones and Accomplishments ......................................................................... 23
Table 4: Project Budget Estimate ......................................................................................... 23
## 1 Executive Summary

<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Intelligent Autonomous Systems based on Data Analytics and Machine Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Bharat Bhargava</td>
</tr>
<tr>
<td><strong>Project Lead</strong></td>
<td>Bharat Bhargava</td>
</tr>
<tr>
<td><strong>University</strong></td>
<td>Purdue University</td>
</tr>
<tr>
<td><strong>Requested Funding Amount</strong></td>
<td>$199,999</td>
</tr>
<tr>
<td><strong>Period of Performance</strong></td>
<td>September 1, 2017 - August 31, 2018</td>
</tr>
<tr>
<td><strong>Is this an existing Investment Project?</strong></td>
<td>No</td>
</tr>
<tr>
<td><strong>TRL Level of Project</strong></td>
<td>5</td>
</tr>
<tr>
<td><strong>Key Words</strong></td>
<td>Autonomous system, data provenance, reinforcement learning, cognitive autonomy, data analytics, machine learning analytics, ontological reasoning, blockchain, trust</td>
</tr>
<tr>
<td><strong>Key Partners &amp; Vendors</strong></td>
<td>Context-based Adaptable Defense Against Collaborative Attacks in SOA; End-to-End Security Policy Auditing and Enforcement in Service-Oriented Architecture; Monitoring-Based System for E2E Security Auditing and Enforcement in Trusted and Untrusted SOA; Privacy-Preserving Data Dissemination and Adaptable Service Compositions in Trusted &amp; Untrusted Cloud</td>
</tr>
</tbody>
</table>

| **NGC projects you have collaborated with in the past** |

**Table 1: Executive Summary**
1.1 Abstract

Intelligent Autonomous Systems (IAS) are highly cognitive, reflective, multitask-able, and effective in knowledge discovery [1]. Examples of IAS include software that is capable of automatic reconfiguration, autonomous vehicles, network of sensors with reconfigurable sensory platforms, and an unmanned aerial vehicle respecting privacy by deciding to turn off its camera when pointing inside a private residence. Research is needed to build systems that can monitor their environment and interactions, learn their capability, and adapt to meet the mission objectives with limited or no human intervention. The systems should be fail-safe and should allow for graceful degradations while continuing to meet the mission objectives.

This project will advance the science of autonomy in smart systems through enhancement in real-time control, auto-configurability, monitoring, adaptability, trust. We plan to contribute to autonomy in smart systems and research in NGC IRAD (smart autonomy, Multi-intelligence (MINT) Enterprise Analytics, and Rapid Autonomy prototype among others). The main objective is to realize the vision as presented by Thomas Vice of NGC based on his talk at Purdue in 2016 and efforts of Donald Steiner based on the following approaches.

(1) Employ machine learning techniques on sensor and provenance data to learn and understand the underlying patterns of interaction, conduct forensics to detect anomalies, and provide assistance in decision making by on-the-fly semantic and probabilistic reasoning.

(2) Apply advanced data analytics techniques to incomplete and hidden raw system data (provenance data, error logs, etc..) to discover new knowledge that contributes to the success of the IAS mission.

(3) Enhance the autonomous system’s self-awareness, self-protection, self-healing, and self-optimization by learning from the knowledge discovered through data analytics.

(4) Utilize blockchain technology for storing provenance data for providing monitoring, trust, and verification, using the NGC-WaxedPrune system envisioned by Donald Steiner, Jason Kobes, and Leon Li, and demonstrated at TechFest in 2015.

1.2 Graphical Illustration

We propose a novel approach that performs on-the-fly analytics on data streams gathered from sensors/monitors of autonomous systems to discover valuable knowledge, learn from the system’s interactions with the runtime environment and adapt its actions in a way to maximize its benefits over time for enhanced self-awareness and auto-configuration capability, and track the provenance of the data gathered/generated by the system to provide increased trust in the actions of the system. By integrating components for streaming data analytics, cognitive computing with deep reinforcement learning and knowledge discovery through unsupervised/supervised learning on streamed data, the proposed model aims to provide a unified architecture for smart autonomy, applicable to various systems that NGC is developing. The overall architecture of the proposed model is demonstrated in Figure 1.
General characteristics of the proposed solution are as follows:

- Data obtained through the sensors/monitors of the autonomous system are fed into data stream processor, which contains modules for pre-processing of the data to prepare it for analytics to derive valuable knowledge. The dimensionality of the data is reduced and data is sampled to allow for real-time processing.

- The pre-processed data is fed into the data analytics module (knowledge discovery engine), which applies unsupervised machine learning algorithms to detect deviations from the normal behavior of the system. The gathered data is used to build a model of the system’s environment and actions by storing it in a knowledge discovery module, which is consulted repeatedly through the lifetime of the system, acting like the memory of a human-being to decide which actions to perform under different contexts.

- The provenance of the data gathered by the sensors/monitors of the system is logged in an immutable private ledger based on the blockchain technology. This provides verifiability of the data which is used in the knowledge discovery process. It helps in building and measuring the level of trust of an IAS.

- The data pre-processed by the data stream processor and the provenance data are fed into the cognitive computing engine, forming the observations for reinforcement learning in the system, so that the system gains self-awareness over time through a reward-based process. The reward can be based on the type of the system; for a UAV, it could be based on the quality of image processing, while for a missile defense system it could be accuracy and time needed to mitigate an attack. The reinforcement learning process utilizes deep neural networks to build a model of the big data gathered, rather than utilize a trial-and-error learning approach. This enables the system to gain increased self-awareness in time, and gain auto-configuration/self-healing abilities. The system acts upon its environment based on the outcomes of the reinforcement learning and knowledge discovery processes, keeping it in an action-value loop as long as it functions.
2 Description of Project

2.1 Statement of Problem

Systems with smart autonomy should be capable of exhibiting high-level understanding of the system beyond their primary actions and their limitations and capacity. They should predict possible errors, initiate backup plans, and adapt accordingly. They should be able to multitask: collaborating with their human counterparts, communicating, and executing actions in parallel. A smart system is also required to monitor its interactions with the environment, find problems, optimize, reconfigure, and fix those problems autonomously, while improving its operations overtime. A comprehensive IAS should be rich in discovered knowledge on which it can reason with that knowledge at various levels of abstraction using several quantitative and qualitative models: semantic, probabilistic, ontological, symbolic, and commonsense. Hence, an IAS is contingent on its cognizance of its operational boundaries, operating environment, and interactions with clients and other services. An IAS should demonstrate reflexivity implying that it continuously adjusts its behavior and adapts to new unpredictable situations. It should have reasoning where it can introspect about its own reasoning limitations and capacity.

![Fig. 2 Conceptualization of Comprehensive Intelligent Autonomous Systems (IAS)](image)

These characteristics lead to the following research problems and directions: (a) how to enhance the cognizance of IAS using novel cognitive processing approaches that enable the system to be aware of the underlying operating and client context where the data is being generated, (b) how to conduct distributed processing of streaming data on-the-fly (and in parallel) in order to apply advanced analytics techniques and machine learning models for knowledge discovery, (c) investigating new analytics techniques for finding underlying patterns and anomalies, thus increasing the value of the gathered data, (d) how to facilitate learning from data to improve the adaptability of the IAS, (e) how to innovatively apply blockchain technology in order to provide trust and verifiability to IAS, (f) how to contribute to representation and reasoning approaches based on both qualitative and quantitative models—probabilistic, ontological, semantic, and commonsense—to discover new knowledge, and finally, (g) how to advance science of learning algorithms to enable autonomy in self-optimization, self-healing, self-awareness, and self-protection, and to reason about making decisions under uncertainties.
2.2 State of Current Technology and Competition

Wes Bush, CEO of Northrop Grumman presented in Kansas State University [2] several insights that relate to our proposed research. According to him an autonomous system should be able to act without the lapses of human judgment or execution inadequacies and provide the same level of concern as a human to a particular task. This is defined as cognitive autonomy [2]. A concept generation system for cognitive robotic entities is implemented by Algorithm of Machine Concept Elicitation (AMCE) [13]. AMCE enables autonomous concept generation based on collective intention of attributes and attributes elicited from formal and informal definitions in dictionaries. In [14], a bio-inspired autonomous robot with spiking neural network (SNN) is built with a capability of implementing the same SNN with five variations through conditional learning techniques: classical conditioning (CC) and operant conditioning with reinforcement or punishment and positive or negative conditioning. A wideband autonomous cognitive radio (WACR) has been designed and implemented for anti-jamming in [15]. The system has the collected data on spectrum acquisition as well as the location of the sweeping jammer. This information and reinforcement learning is used to learn the perfect communication mode to avoid the jammer. Here, the system is self-aware about the current context. We will investigate learning models and analytics to attain cognitive autonomy in IAS. To conduct data analytics on-the-fly and change the analytics techniques automatically, an instrumented sandbox and machine learning classification for mobiles is implemented in [16]. The analysis is conducted, adjusted, and realigned based on the information of mobile applications submitted by the subscribers. There are well-known knowledge discovery mechanisms that can be applied on raw data to discover patterns. In [17], the authors outline scalable optimization algorithms and architectures encompassing advanced versions of analytics techniques such as principle component analysis (PCA), dictionary learning (DL), and compressive sampling (CS). We will be employing advanced data analytics techniques to discover patterns and anomalies from raw data.

Thomas E. Vice, corporate vice president of NGC, gave a talk at Purdue University about the future of autonomous systems [54]. He outlined the projects on autonomous systems and how Trusted Cognitive Autonomous Systems will be the future. Our project complements the vision of NGC. Through discovered knowledge, an IAS can continuously learn, reason, predict, and adapt to the future events. A lightweight framework for deep reinforcement learning is presented in [18]. The learning algorithm uses the asynchronous gradient descent for optimization of deep neural networks. In this paper [19], the authors introduce an agent that maximizes the reward function by continuous reinforcement learning with an unsupervised auxiliary task. Reinforcement learning is one of the major machine learning methods that is used primarily on automated cyber physical systems such as autonomous vehicles [20-22] and unmanned aerial vehicles (UAVs) [23-25]. Defender-and-attacker game, a game theoretic approach, is employed in general learning models of security as well. When the attacker information is very limited and attacker persistently makes her moves (in the game) to affect the system, the defender needs to constantly adapt to the attackers’ novel strategies. So the defender constantly reinforces her beliefs based on the attacker moves and creates a robust defense strategy for future attacks [26]. We will use reinforcement learning algorithms to enhance automated decision making and dynamic reconfiguration capabilities to increase the reflexivity of the system.

Data provenance is used in forensics and security for providing robust support for the underlying systems, sometimes autonomous, through valuable meta-information about the system.
and its interactions [7]. Data provenance has been modeled for and used in autonomous systems in service-oriented architecture [3] [4] [12] and autonomous information systems [5] [6]. Further investigation is needed to model the use of provenance in enabling autonomy. The Database-Aware Provenance (DAP) architecture [8] provides a workflow that detects the addition of any new autonomous unit of work for fielding any service request and tracks its activities to extract the relevant operational semantics. Provenance data is also used to enhance trust and security in autonomous systems. Trust in information flow can be maintained and verified by provenance data [9], where trust of autonomous entities can be quantified by data provenance and internal values of the data items. Piercing perimeter defenses in autonomous systems can be resolved by provenance-aware applications and architectures [10]. To enable autonomy, systems must be able to reason about and represent provenance data at multiple levels of abstraction. Quantitative and qualitative reasoning can enable semantic knowledge discovery and predictable events. Semantic ontologies are widely used in autonomous cyber-physical systems (CPS) [27]. Ontology-like reasoning over several intelligence representations of new entities can enable the autonomous system to reason about unexpected entities present in their environment [28] [29]. A recent study [11] shows that trust and immutability are provided through provenance on blockchain technology, where smart contracts can be created. This increases trust, provides consensus, and reduces the need for third party intervention: creating a decentralized autonomous setting. Provchain—a blockchain-based data provenance architecture is proposed in [30] to provide enhanced availability and privacy in cloud environments. Blockchain provides integrity to provenance data through its immutable property [31]. Our research will utilize data provenance with blockchain technology for modeling autonomy in smart systems.

2.3 Proposed Solution and Challenges

We propose a comprehensive approach to enable autonomy in smart systems by enhancing the following fundamental properties of IAS: cognitive—mindfulness of the current state of the system (self-awareness), reflexivity—ability of the system to monitor and respond to known and unknown scenarios, and adjust accordingly with limited or no human intervention (self-optimization and –healing), knowledge discovery—ability to find new underlying patterns and anomalies in system interactions through advanced data analytics techniques, predictive—learn and reason from the discovered knowledge, anticipate possible future events, and recalibrate corresponding actions, and finally trust—ability to provide verification and consensus for the clients as well as for the system (self-protection).

The quality and trustworthiness of data in an IAS is of prime importance for achieving the abovementioned goals. We will utilize the following data storage/sharing technologies and data sources when modeling the system and conducting experiments.

**NGC-WaxedPrune prototype system:** Data are stored in the Active Bundle [39] [40] [41], which is a self-protected structure that contains encrypted data items, access control policies, and a policy enforcement engine. It assists in privacy preserving data dissemination. The design of this system received the first rank (voted by corporate partners) at the 2015 annual symposium competition of the Purdue CERIAS center. This system can be used to deal with all data generated and monitored in IAS and its interactions with outside entities.

**Provenance data:** In the Active Bundle scheme, provenance metadata is generated, attached to an Active Bundle and sent to a central monitor each time a service accesses data. Provenance metadata contains information on when data was accessed, where, by whom, as well as several
execution environment parameters, such as OS version, Java version, libraries, CPU model at data recipient's side. Using provenance as a basis for decision making largely depends upon the trustworthiness of provenance [36]. We can deploy Active Bundle as used in WaxedPrune and blockchain storage for provenance data [33] in order to provide trust and integrity to IAS.

**Monitoring Data:** Log files are one of the most numerous data collection methods to record activities, user-and-system generated errors, notifications, transactions, interaction with third parties, etc., [31]. Employing advanced data analytics techniques can provide us with rich knowledge of patterns and anomalies. We intend to use the log files of the WaxedPrune system. Analytics on numerical data from sensors/monitors of autonomous systems can be used to verify the convergence of reinforcement algorithms [34]. We will use publically available data to test the proof of concept in terms of accuracy and convergence of machine learning techniques, reinforcement algorithms, and reasoning models for IAS.

The individual components of the proposed smart autonomy model are described in the subsections below.

### 2.3.1 Cognitive Autonomy:

An IAS in a distributed environment should be aware of its three major system, software, and interaction layers: (1) its own state of the system and software as well as operational parameters, (2) state of its neighboring systems, and (3) client or third party services and their interactions with the system.

---

![Cognitive Computing Process for Autonomous Systems](Image)

---

We propose a novel approach that uses Artificial Intelligence (AI) techniques to monitor and learn the state of autonomous systems to automatically adapt to meet mission objectives with no human intervention. The main idea of the proposed research is actively monitoring the system to provide those results as inputs to decision-making machine learning algorithms that determine the new configuration of the system based on the resulting outputs. This research will focus on the
analysis of two types of data: (1) performance parameters, such as response time, CPU usage, memory usage, etc., and sensor data peculiar to the system and (2) data access patterns stored as data provenance in blockchain for misbehavior detection. By integrating system performance and either benign or malicious behavior data in making decisions from past experience the proposed model aims to provide a unified and comprehensive architecture for self-healing intelligent autonomous systems.

Deep reinforcement learning [18] will be utilized as the primary machine learning technique for cognitive computing in the system to achieve adaptability to different environments, learn from previous vulnerabilities and maximize the security. As stated by Mnih et al. [39], reinforcement learning provides a way to model human behavior in terms of optimizing control of an environment of the agent, through an action-value feedback loop. Reinforcement learning is a difficult task due to the complexity of representing an environment with high-dimensional sensory data. Nevertheless, recent advancements in deep learning allow for building more abstract representations of data from sensors through utilizing multiple levels of nodes, which can be used as the model to optimize the action-value function in the reinforcement learning process. Deep reinforcement learning has recently been successfully applied for tasks like playing Atari games [18].

The deep neural network (DNN) component of the cognitive computing engine will be used to approximate the optimal action-value function for the reinforcement learning model. Deep neural networks also solve the problems of adversarial search and Markov decision processes. The Markov property is nothing but the probability of the current event ($E_i$) depending on the probability of the previous event ($E_{i-1}$). With DNNs, we can store and build more memory in the previous state. Through this increased memory, we can build effective Higher-order Markov models, which recollect more data history, enhancing more predictive capability of the system. We can represent the Markov decision process as follows: in the $n^{th}$ Markov model,

$$\Pr(E_i \mid E_{i-1}, E_{i-2}, ..., E_1) = \Pr(E_i \mid E_{i-1}, ..., E_{i-n})$$

We will employ customized higher-order Markov Decision Processes (MDP) to create novel reinforcement algorithms. For example, consider a smart system executing functionalities in a cloud environment. In the Markov model, there are states before (past, present, and future states) but currently the future states of the system are not only affected by the past state but also affected by the current actions of the client services and the system. There will be a reward function for the autonomous system, and in the transaction, the system must maximize the rewards. Given time ($t$), actions ($A_t$), rewards ($R_t$), and states ($S_t$), a reinforcement learning model is represented below,

$$S_t \xrightarrow{A_t} S_{t+1} \xrightarrow{A_{t+1}} S_{t+2} \xrightarrow{A_{t+2}} S_{t+3} \ldots$$

Each state is combined with the actions and maximized reward function, so the system learns which actions to perform to gain more rewards and which actions to reduce the loss. The cognitive computing engine in the proposed research takes as input the data preprocessed by the data stream processor as well as provenance data, which represent the state/observations of the autonomous system for reinforcement learning. The task of the engine is to enable the system to make the best decision for the next action given the context of interaction, the current states of the various system
parameters and the knowledge discovered through performing on-the-fly analytics on the streamed data. The overall goal is to select actions in a way to maximize the cumulative QoS parameters that include security and trust, performance, real-time response, and degradation. We will deploy NGCRC funded research on active monitoring tools for measurements of the performance and behavior of services, ideas from MTD [50] for switching replicas and will incorporate new tools for both supervised and unsupervised learning to allow dynamic reconfiguration under various unknown environments, context, and situations.

2.3.2 Knowledge Discovery: The knowledge discovery component of an IAS employs methodologies from pattern recognition, machine learning, and statistics to extract knowledge from raw, and sometimes unknown data. Knowledge discovery is an important element in supporting cognitive autonomy since new knowledge discovered can trigger changes to the smart system to adapt to the new parameters, thus enabling autonomy. Discovered knowledge constitutes the representation of unknown data, its form, and its degree of certainty. The generic process of knowledge discovery is shown in Figure 4 below.

![Knowledge Discovery Process Diagram](image)

Fig. 4 Knowledge Discovery in Autonomous Systems

Knowledge discovery on large data, in particular streaming data, needs efficient data processing. Distributed data processing on streaming data becomes a necessity for faster classification and storage of data [52]. We will introduce a parallel processing of data items that can classify and categorize the streaming data considerably fast. The classification and clustering techniques must be capable of on-the-fly processing of data streams: distributed data processing can accommodate simultaneous processing of sequential/parallel data streams: the key idea behind the parallel processing is to host distributed data processing units (DDPU) that can (a) read (R) to load the data, (b) Analyze (A) to process and classify the data, and (c) toggle (T) to shift to/from read or analyze. For example,

<table>
<thead>
<tr>
<th>DDPU 1</th>
<th>R item 1</th>
<th>T→</th>
<th>R item 2</th>
<th>T→</th>
<th>A item 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDPU 2</td>
<td>A item 1</td>
<td>T→</td>
<td>R item 3</td>
<td>T→</td>
<td>A item 4</td>
</tr>
<tr>
<td>DDPU 3</td>
<td>R item 4</td>
<td>T→</td>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Cycle 1 Cycle 2 Cycle 3
The representation above shows a fundamental distributed data processing technique—\(RAT\)—to process data on-the-fly, which is scalable to process Big Data streams. Depending on the priority and availability of the data items each processing unit prioritizes the RAT operation for each data item. In this way, instead of relying on static rules and heuristics to determine prioritization of data processing, we can compute the value of the data on-the-fly based on data’s quantitative/qualitative system metrics such as sensitivity, dependence, and importance of the data, and process the data items accordingly. This distributed processing of data streams will contribute to the Distributed Data Processing IRAD of NGC.

The processed data can contain both categorized (easy to label) such as data origin, time of creating, and modification, etc., and uncategorized data such as error logs (text). Hence, we will employ both customized—combination of multi-level decision trees and Bayesian probabilistic methods—classification and regression algorithms [53], and advanced clustering techniques to achieve high dimensionality and to label the data, and prepare it for analysis. We will be using Bayesian statistics to estimate the reliability of the autonomous system and quantify the unknown due to lack of data (missing data). Bayes’ theorem states that, given two data items \(D_1\) and \(D_2\),

\[
\Pr(D_1 \mid D_2) = \frac{\Pr(D_2 \mid D_1) \cdot \Pr(D_1)}{\Pr(D_2)}
\]

The reliability of the autonomous system can be measured using Bayesian statistical methods with conditional probability and prior distribution of the autonomous system’s states. New knowledge can be discovered through reliability analysis of autonomous system, which will contribute to the self-awareness of the system, enabling smart autonomy. Our Bayesian statistics approach will contribute to the Reliability Analysis Data System (RADS) IRAD of NGC.

Quantitative and qualitative reasoning can enable semantic knowledge discovery and predictable events. Semantic ontologies are widely used in autonomous cyber-physical systems. We will apply ontologies to generate semantic reasoning over the provenance data. For example, semantic ontology reasoning will be used to extract attributes of provider-client interaction such as: platform, data requested, update, and access. Applying semantic reasoning models to the log files of provenance data will help the system discover new knowledge about the client. This will be stored and used to make decision and contribute to autonomy.

Of particular interest to the knowledge discovery process in the proposed system are the following methods that we will investigate and integrate into the knowledge discovery engine:

1. **Association Rule Mining**: Association rule mining discovers patterns of the form “if \(X\) then \(Y\)”, where \(X\) and \(Y\) are item sets. This allows us to find frequent patterns of co-occurrence in large datasets. Typical algorithms for association rule mining include the Apriori algorithm, sampling algorithm, frequent pattern tree and partition algorithm. For IAS, we will utilize the mentioned association rule mining algorithms to discover system events that co-occur frequently under normal and anomalous circumstances (e.g. CPU and memory usage spiking up together). This will allow the system to have increased awareness of what environment and system conditions to expect when a certain event occurs and adapt itself accordingly.

2. **Clustering**: Clustering allows us to partition data without having a training sample, which is useful in situations where the system has just started functioning and we need to discover groups of events/data similar to each other in terms of certain parameters, representing different states of the system. We will employ \(k\)-means clustering, a typical algorithm to cluster multi-dimensional data \(D\) consisting of \(m\) records \(r_1\ldots r_m\) into \(k\) clusters \(C_i\) with
centroids \( m_i \) using the squared error criterion for optimization, such that each record is assigned to the cluster with the minimum distance to the centroid of that cluster. The error is measured as:

\[
\sum_{i=1}^{k} \sum_{r_j \in C_i} \text{Distance}(r_j, m_i)^2
\]

Here the most effective distance function can depend on the nature of the data, therefore we will experiment with multiple distance functions. Finding clusters of IAS data along various dimensions will allow for detection of anomalies when incoming data does not belong to any of the previously built clusters. This is also useful for discovering cases like zero-day attacks, which have no known attack signature through detecting deviations from the normal behavior of the system.

3. **Sequential/Temporal Pattern Mining:** Sequential/temporal pattern mining discovers patterns in a dataset that occur frequently in a particular sequence. The gold standard for time series analysis is Hidden Markov Models (HMM), therefore we will utilize HMM to build a representation of IAS behavior through observation of the system states and state transitions over time.

Based on HMM, the system can be in one of the \( N \) possible states \( \{S_1, S_2, \ldots, S_N\} \), and undergoes a transition from one state to another at particular times. The state transition probabilities of the system depends on the immediate past, i.e.

\[
P(q_t = S_j \mid q_{t-1} = S_i, q_{t-2} = S_k \ldots) = P(q_t = S_j \mid q_{t-1} = S_i)
\]

Additionally, the observations (data gathered through sensors/monitors) are a probabilistic function of each state, i.e.

\[
P(o_t = v_k \mid q_t = S_j)
\]

where \( o_t \) is the data observed at time \( t \) and \( v_k \) is a distinct observation in the set of possible observations for the system. Using HMM, we will build a probabilistic model of the system from a sequential set of observations/data, which best explains the behavior of the system in terms of transitioning between different states and the data resulting from the transitions. For example, a low CPU usage observation can be associated with a malfunctioning module state with high probability, while an extremely high CPU usage observation can be associated with a system under attack state. Based on the knowledge discovered over time with HMM, the IAS will be able to predict current and next states more accurately, and take adaptability actions accordingly. Critical node analysis in higher order Markov models can lead to identifying critical steps in complex attack strategies of adversaries, reducing resource usage for target analysis. Once the pattern is discovered, the systems can reinforce its understanding and adapt to the new set up.

In addition to the abovementioned techniques, various models for detection of outliers in different types of data have been devised by the machine learning community. While supervised and unsupervised learning models have been applied with success to a variety of domains, robust models for detecting anomalies and failures in IAS operation are still lacking. The main shortcoming of supervised anomaly detection models are that they require a large amount of training data and can only provide accurate results on anomalies that were previously observed in the system. This makes such models unable to capture threats/anomalies that are completely new, which is essential in an environment of ever-growing security vulnerabilities and attacks. A
significant advantage of unsupervised models is that the training data required is gathered from the behavior of services operating under normal conditions (possibly in an isolated environment/private cloud); i.e. no attack data is required to train these models. We will consider the advantages and disadvantages of existing models as listed in Table 2 and focus on the development of techniques that are both accurate and have low runtime overhead, possibly using an ensemble of models from the literature.

<table>
<thead>
<tr>
<th>Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-means Clustering</td>
<td>Low complexity</td>
<td>Sensitive to noisy data</td>
</tr>
<tr>
<td>EM Meta Algorithm</td>
<td>Adaptable to different distributions</td>
<td>Converges slowly in some cases</td>
</tr>
<tr>
<td>One-Class Support Vector Machine (SVM)</td>
<td>Can handle very high-dimensional data, usually has high accuracy</td>
<td>High memory and CPU, needs positive and negative examples</td>
</tr>
<tr>
<td>Unsupervised Neural Network</td>
<td>Has learning capability</td>
<td>Long processing for big networks</td>
</tr>
<tr>
<td>Self-Organizing Map</td>
<td>High dimensionality reduction</td>
<td>Time consuming</td>
</tr>
<tr>
<td>Hidden Markov Models (HMM)</td>
<td>Representative of the time-based relations and states of services</td>
<td>Have scalability issues</td>
</tr>
</tbody>
</table>

Table 2: Machine learning techniques for outlier/anomaly detection

2.3.3 Reflexivity of the system: The goals of IAS in the proposed approach are (1) replacing anomalous/underperforming modules with reliable versions or adapting to a new mechanism to avoid anomalies, (2) reconfiguring system parameters to respond to anomalous system behavior, (3) swiftly self-adapting to changes in context, (4) enforcing proactive and reactive response policies to achieve performance and security goals, and (5) achieving continuous availability even under attacks and failures.

Providing adaptability in order to achieve increased autonomy in IAS relies on two main elements:

1. Being cognitive and determining action: Monitoring of systems is of utmost importance in achieving high self-awareness, as systems in environments with highly dynamic contexts may exhibit frequent changes in many QoS parameters. We measure the assurance level, (integrity/accuracy/trust) of the system from the performance parameters such as response time, throughput, packet loss, delays, consistency, acceptance test success, etc. Compliance with all the requirements of IAS is hard to achieve in such dynamic environments, making monitoring a must for accurate decision-making. The tasks involved in effective monitoring and analysis of the obtained data include the following: (a) identification of QoS metrics, such as response time, CPU usage, memory usage, etc., to determine the performance and behavior of IAS; (b) development of models for identifying deviations from performance (e.g., achieving the total response time below a specific threshold) and security goals (e.g., having trust levels above a certain threshold).

2. Autonomous system reconfiguration based on changes in context: Changes in the context of IAS can affect system behavior, requiring autonomous reconfiguration. While changes in user context can result in updated priorities such as trading accuracy for lower response time in an emergency, changes in system context can result in failures requiring the restart of a component of the IAS. Dynamic reconfiguration of system modules based on the updated constraints and contexts enables success of mission objectives.
Adaptability allows dynamic configuration of software and execution to meet the changing demands of autonomous systems for performance, reliability, security, and resource utilization. Adaptable systems provide graceful degradation and can respond to the timing, duration, type, extent, severity of failures and attacks. Adaptation must satisfy the consistency and integrity constraints. The granularity of formally defined classes of algorithms will determine the overhead and benefits of adaptation. Experiments in adaptability allow systems to identify conditions for satisfying the Quality of Service (QoS) requirements of mission objectives and provide guidelines for reconfiguring algorithms, protocols, sites and associated servers, communication software and routers, and others components. We have explored many ideas about how to create new replicas and determine when to execute the replacement of nodes. One of them is based on graceful degradation. The main idea is having primary and alternate modules and using an acceptance test to validate their operation. Initially a primary module is used and constantly tested. In case of failure there are two alternatives: (1) weaken the acceptance test or (2) replace the primary module with the alternate/replica that can pass the acceptance test. Figure 5 illustrates the concept. In the case that an alternate module replaces the primary module of the IAS not able to pass an acceptance test, the composition of a process in the IAS can change as shown in the lower part of the figure (Note that here the system has two module alternatives for a process A, which invokes a process B with three module alternatives, that further invokes process M having three module alternatives chosen based on the acceptance test process in the upper part of the figure).

Fig. 5 Dynamic Adaptation based on Recovery Block Scheme

Adaptable autonomous systems should be able change their system configuration to guarantee mission critical operations at the cost of sacrificing performance. Because some services may continue their effort to compromise these systems there exists a need for more adaptable solutions
to protect systems. Our proposed Moving Target Defense (MTD)-type [50] is a defensive strategy that aims to reduce the need to continuously fight against attacks by decreasing the gain-loss balance perception of attackers. The framework narrows the exposure window of a node (module) to such attacks, which increases the cost of attacks on a system and lowers the likelihood of success and the perceived benefit of compromising it. The achieved reduction in the vulnerability window makes this strategy optimum for adaptable autonomous systems.

The proposed framework introduces reflexivity and adaptability to systems. Reflexivity has two main components: (1) continuing operation and (2) adapting to counter anomalies. The MTD-style approach takes into consideration these components since it transforms systems to be able to adapt and self-heal when ongoing anomalies are detected, which guarantees operation continuity. The initial target of the framework is to prevent successful compromises by establishing short lifespans for nodes/services to reduce the probability of attackers’ taking over control. In case an attack occurs within the lifespan of a node, the proactive monitoring system triggers a reincarnation of the node.

2.3.4 Trust in Autonomous Systems: Self-protection (automatic identification and protection from security threats) and self-healing (automatic fault discovery and correction) are important properties of an IAS [43]. We propose an approach for data provenance with blockchain-based mechanisms to build trustworthiness of the data and ensure identities of network participants. Integrity of data will be guaranteed by blockchain technology. Data can be used for threat detection. Optimized access for transaction validation procedure allows to reduce number of blocks in the blockchain. There is one Merkle tree per Active Bundle and it gets updated with the hash of the data each time a transaction occurs, i.e. either data is read from Active Bundle or data inside Active Bundle gets updated by an authorized service. Provenance record contains information on what data type has been accessed / updated, by whom (by which service), when and who sent the Active Bundle to the service.

Challenges of blockchain technology deployment
1. Performance: Blockchain is replicated to all the network participants and this imposes a performance overhead. This was discussed with Peter Meloy of NGC-UK.
2. Access Control (Read): In case of access revocation or subject’s role change, access to data must be revoked immediately within an information system when authorization is no longer valid. However, revoked access to data on a blockchain can be bypassed in the following ways: (1a) by replaying old blocks against an empty blockchain and stopping before the revocation block is appended; (1b) An attacker holding a copy of a blockchain could use a modified client to just ignore the revocation block. Even if read access to local blockchain requires an off-chain token handshake with a centralized authority for authorization; then that token would continue to work forever in the future. The requirement to revoke previously granted access can be bypassed by rolling the local clock back and restoring unauthorized access to blockchain data.

We discussed and learnt many blockchain ideas from Steve Seaberg (NGC). We plan to collaborate with Peter Meloy, Steve Seaberg, and Vladimiro Sassone (University of Southampton, UK) to work on blockchain – based methodology for provenance data storage and verification.
2.4 Distinctive Attributes, Advantages, and Discriminators

- **Cognitive autonomy with advanced data analytics**: Cognitive processing will be employed using customized activity pattern identification (multilevel decision trees and hierarchical Bayesian inference) models to understand the parameters of the environment such as processing platform, entities (both systems and humans), and their features to automatically detect and adjust the provenance metrics such as importance factor. If the data is accessed by the client in an unsecured or unverified environment the Bayesian inference to assess the probability of security risks is utilized. Once the risk is determined, the system can store additional provenance data points to monitor closely. If an anomaly is detected with on-the-fly provenance data analysis, then the system would take appropriate actions.

- **Automated reconfiguration**: Existing approaches for autonomy in IAS lack robust mechanisms to monitor compliance of systems with security and performance policies under changing contexts, and to ensure uninterrupted operation in case of failures. The proposed work will demonstrate that it is possible to enforce security and performance requirements of IAS even in the presence of anomalous behavior/attacks and failure of system modules. The self-healing will be accomplished through automated reconfiguration, migration, and restoration of modules.

- **Reflexive systems**: We will design and implement the reflexive machine learning model to make decisions and trigger corresponding actions for adaptations, while proactively monitoring (being cognitive) the system. We abstract the system runtime from autonomous system to formally reason about its correct behavior. This abstraction allows the framework to enable MTD-style capabilities to all types of systems regardless of its architecture or communication model (i.e. asynchronous and synchronous) on all kinds of platforms). The modules of cognitive monitoring, trust, and automated migration/reconfiguration can be easily integrated into NGC enterprise analytics flow. The modular architecture and use of standard software in the monitoring framework allows for easy plugin to IRAD software. The automation work will allow identification of NGC clients’ requirements for building capabilities in prototypes for Air Force Research Lab (AFRL). We plan to work closely with NG on BAA proposals.

- **Blockchain-based provenance for trust**: We will use blockchain technology to store provenance data and utilize customized direct acyclic graph (cDAG) data structure for blockchain implementation. A DAG has no direct cycles and it is a finite directed graph and it aids autonomy by reducing the interventions from external sources and making it easier to create robust blockchain protocols. Merkle tree optimizations will be implemented by assigning quantitative measures to the provenance data points to decide the significance of each data point (or set of data points) so that they are sufficient for data analysis and making an informed decision. Only these significant data points will be stored in blockchain. These two techniques will increase the efficiency of implementation.

2.5 Tangible Assets to be Created by Project
2.5.1. Software

- **Stream Data Analytics Engine:** This module will accept data streams of various formats and types and perform analytics on the data on-the-fly through parallel processing. It will include a novel stream processing architecture aided by Apache Kafka, data sampling, and dimensionality reduction components to speed up the processing and reduce the noise in the data.

- **Knowledge Discovery Engine:** This module will be implemented as a suite of advanced data analytics techniques and machine learning algorithms to discover useful patterns in the data gathered by the autonomous system. The outputs of the algorithms will help in the detection of anomalies and normal system behavior, contributing to the building of a cognitive model of the system that could also be consulted by other IAS.

- **Cognitive Computing Engine:** This module will utilize data from the stream processor and the knowledge discovery modules to build a cognitive model of the system. It will primarily be based on a state-of-the-art deep reinforcement learning algorithm. This will allow the system to learn about its current state, its current context, and systems that are interacting with it.

- **Data Provenance Ledger:** This will be implemented as a private ledger keeping immutable records of provenance data, based on the blockchain technology.

2.5.1 Documentation

We will provide four types of documentations that would help NGC researchers. They include:

1. **Source code:** Code for the software will be well self-documented for possible extensions/modifications by future developers.

2. **Deployment and user manuals:** All software components created in the project will be clearly documented with deployment guides and user guides on how to use each component separately as well as how to use the whole prototype.

3. **Reports:** We will provide mid-term and final reports that describe algorithm implementations, and the experimental results that characterize the performance of the presented solutions. These results will include both system performance and security evaluation of the system.

4. **Demonstrations:** To be made at NGC meetings and to NGC researchers.

We will provide high-quality documents adhering to the standards used at NGC.

2.6 Technical Merit and Differentiation

The proposed approach offers many advantages over existing solutions for autonomy, learning, and adaptation in IAS. The main benefits of the proposed approach are:

(a) The solution is generic and targets multiple layers (cognitive, reflexive, knowledge discovery, and predictive: all are done with automation i.e. with limited or no human help) of the NG software stack, as opposed to traditional techniques for manual mitigation, reconfiguring, and decision making.
(b) The solution will be built upon award winning research at Purdue on adaptability, V2V, UAS, NGC-Waxedprune system infrastructure, with the potential of funding from AFRL and NSF. 

(c) The solution is based on industry-standard technologies such as blockchain and distributed cloud environment, providing seamless integration into existing systems. For example, our cognitive computing model is utilized as the base technology of the framework for autonomous vehicles and UAVs.

(d) The proposed reflexive framework facilitates proactive mitigation of anomalies and failures through active monitoring of the performance and behavior of systems and can incorporate new tools for resiliency and antifragility under various failures, security threats, and insider attacks. The solution enables formal reasoning about system self-awareness, self-optimization, and self-reconfiguration contributing to the science of autonomy.

(e) Blockchain-based provenance storage will be supported for providing trust and immutability of data in autonomous systems. Data provenance tracking enables learning from data leaks, repairing them, mitigating the loss and modifying the system so any future leak can be avoided.

(f) Continuous monitoring, self-restoration and self-healing of smart system operations and data allows highly automated and cognitive system by learning from anomalies and failures and self-reconfiguring the underlying system accordingly to increase smart autonomy.

3 Project Milestones

3.1 Statement of Work

3.1.1 Cognitive Autonomy and Knowledge Discovery

A cognitive computing engine for autonomous systems that utilizes diverse machine learning techniques will be developed. This engine will create models for normal behavior and anomaly detection in autonomous systems. To this end, several phases [8] need to be completed as below.

- **Observation selection**: The single data entity that represents the state of the autonomous system. Each observation consists of several features (a particular type of information). We will focus our study in data from the performance evaluation of the system (e.g. response time, CPU usage, memory usage, etc.) and access patterns of users stored in blockchain.

- **Dataset**: A collection of observations, each containing values for each of the features. The dataset must be representative to guarantee generalization.

- **Feature generation and selection**: Any creation of new features based on original or derived datasets. After having a comprehensive set of features we will select the subset that best fits our interest for system behavior modeling.

- **Method selection**: We will explore both supervised and unsupervised methods. With supervised methods a trained engine with labeled data will allow to classify an observation as either benign or malicious. On the other hand, unsupervised methods will allow clustering observations as benign or malicious without previous training.

- **Outlier/anomaly detection**: The cognitive computing engine will trigger the reconfiguration of the system if an anomaly is detected.

The following experiment is planned for proof of concept and further tune our research ideas and approaches.
Experiment 1: Anomaly detection to trigger system reconfiguration

- **Input:** A dataset of time series provenance data, network sensor data, and system and software performance data. A significant amount of data generated about the system state is discrete in nature.
- **Output Parameters:** Model that detects all anomalous patterns after analyzing the input data. Pattern discovery provides information about potential malfunctions, security loopholes, insider attacks, and other failure events in autonomous systems.
- **Experimental Setup:**
  - Select a representative dataset of the system state.
  - Explore a variety of unsupervised techniques and supervised techniques (deep learning) to determine the model that performs the best in the anomaly detection process.

3.1.2 Reflexivity of the system

We will develop approaches for automated observation and detection of latent anomalies that will trigger the reconfiguration of the system to guarantee system performance. Specific tasks for automated monitoring include:

- Defining metrics to quantify effectiveness of system components (services, data, networks etc.)
- Defining costs of software-based reconfiguration/monitoring/healing of system components.
- Developing models and mechanisms for optimized automated monitoring and reconfiguration of system architectures to achieve maximum possible reflexivity with minimum operational cost.

The reflexivity a system can be measured in its capacity of changing without interrupting the services running on it. We plan to work on the following adaptability and restoration tasks:

- Developing techniques for adaptable reconfiguration of IAS, which utilize performance and security data gathered by monitors (e.g. response time, CPU and memory usage, number of authentication failures, response status) to create system configurations that better meet quality of service (QoS) and security requirements. These can assure effective missions through assessment & control of the cyber situation in mission context. This effort allows agile operations and the capability to escape harm by dynamically reshaping cyber systems as conditions/goals change. It enhances automated reflexivity.
- Designing techniques for reconfiguring system parameters with a graceful degradation approach to replace individual modules not meeting acceptance tests with more reliable alternate versions at the expense of possibly lower performance.
- Developing risk and performance estimation models and optimization algorithms that will be integrated into the reconfiguration process to achieve optimal performance in system reconfiguration with careful consideration of costs and benefits of adaptability.

The following experiment is planned for proof of concept and further tune our research ideas and approaches.

Experiment 2: Measuring level of autonomy in reflexive capabilities of the system

- **Input:** Machine learning models obtained through experiment 1 vs. the graceful degradation method with different acceptance tests (under different conditions CPU usage, memory usage, etc.)
• **Output Parameters:** We will be measuring convergence time—the time it takes for an algorithm to complete based on input data, reaction time—the time it takes for the system to make decisions and corresponding actions based on the ML algorithms, and computation and communication overhead of the approaches specified in the input.

• **Experimental Setup:**
  - Select a representative dataset of the system state to train the cognitive computing engine
  - Set of acceptance tests under different environmental conditions (dynamic context)
  - Measure the performances of a variety of unsupervised techniques and supervised techniques (deep learning) using the indicated output parameters to compare the different models specified in the input.

### 3.1.3 Trust in Autonomous Systems

We will develop approaches for automatic detection of misbehavior to increase the automatic response of the system, as it will react not only to performance degradation but also to undue use of the resources by users. To that end, a blockchain will be used to keep the history data access patterns of users as provenance data. The specific tasks to accomplish this goal are:

- Develop smart contracts (code running in the blockchain) that will process and keep the autonomous system access patterns of users in the blockchain.
- Develop a friendly interface that will allow the access to the access patterns when it is required.
- Develop the interface to supply the data access patterns to the knowledge discovery module of the cognitive computing engine for data analytics.

The following experiment is planned for proof of concept and further tune our research ideas and approaches.

**Experiment 3: Measuring cost/benefit of trust in autonomous system**

- **Input:** Access patterns of each user/service of the autonomous system.
- **Output Parameters:** We will be measuring processing time—the time it takes for the smart contract running in the blockchain to store the data, accessing time—the time it takes for the system to access the data in the blockchain, anomaly detection effectiveness and the computation and communication overhead.
- **Experimental Setup:**
  - Select the format of users’ access pattern passed to the smart contract (service running in the blockchain) for processing and storage.
  - Create the interface for communication between our autonomous system and the blockchain
  - Measure the effectiveness of the anomaly detection mechanism (models of experiment 1) over access pattern data.
  - Measure the impact in performance (communication and computational overhead).

### 3.1.4 Integration with NGCRC and NGC IRAD projects

We have collaborated with Dr. Donald Steiner and Jason Kobes on Waxedprune. A prototype was demonstrated at NGC TechExpo in June 2016. We are communicating with NG researchers and
plan to contribute to their efforts in the following IRADs:

- Smart Autonomy (with Donald Steiner, Will Chambers, and Miguel Ochoa)
- Rapid Autonomy prototype
- Multi-intelligence (MINT) Enterprise Analytics (with Brock Bose).
- Reliability Analysis Data System (RADS)

We have discussed research questions and plan to collaborate with Peter Meloy in NG-UK, And Steve Seaberg. We will coordinate with Jason Clark/Joshua Bernstein to target BAA in Air Force Research Laboratory, Rome, NY and coordinate with Dr. Donald Steiner for a proposal to NSF for funding via regular and transfer technology programs.

3.2 **Milestones and Accomplishments**

The following table shows the list of tasks to be accomplished during the project period, broken down in a quarterly basis. We plan to hold weekly meetings in Fall 2017 and Spring 2018 with NG researchers to accomplish the development of demos for Tech Expo 2018.

<table>
<thead>
<tr>
<th>Task</th>
<th>Q1 (Sep - Nov)</th>
<th>Q2 (Dec - Feb)</th>
<th>Q3 (Mar - May)</th>
<th>Q4 (Jun - Aug)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setup of the autonomous system integrated with the NGC WaxedPrune Project</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementation of blockchain network</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integration of the autonomous system with the blockchain network</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Automatic derivation of data (system performance and data provenance)</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Setup of data stream processor</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Development of data analytics models based on the collected data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Development of adaptable models for graceful degradation</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Development of deep reinforcement learning model</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Experiments to test the effectiveness of the solution and tuning of parameters of data analytics models</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Prototype demonstration at NGC TechFest 2018 (if approved)  X
Integration of developed autonomous framework with smart autonomy IRAD at NG  X  X  X

Table 3: Milestones and Accomplishments

4 Project Budget Estimate

The project will involve one faculty, two Ph.D. students (one working on Ph.D. dissertations on intelligent autonomous systems and second one who will facilitate experiments and prototypes and demos for 2018 Tech Expo). Budget will consist of salary for the faculty and salary for Ph.D. students. The total budget including fringe benefits, tuition fees, and Purdue University overhead will be $199,999.

<table>
<thead>
<tr>
<th>Category</th>
<th>Items</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labors Hours</td>
<td>Total Hours: 3120</td>
<td>117,725.00</td>
</tr>
<tr>
<td>Materials and Travel</td>
<td></td>
<td>4,687.00</td>
</tr>
<tr>
<td>ODCs</td>
<td></td>
<td>10,260.00</td>
</tr>
<tr>
<td>Other Indirect costs @ 55% MTDC</td>
<td></td>
<td>67,327.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>$199,999</strong></td>
</tr>
</tbody>
</table>

Table 4: Project Budget Estimate

5 References:


https://engineering.purdue.edu/AAE/aboutus/lectures/rolls_royce/2016_Tom_Vice