SYM-AM-23-095

Excerpt from the Proceedings of the Twentieth Annual Acquisition Research Symposium

Acquisition Research:
Creating Synergy for Informed Change

May 10–11, 2023

Published: April 30, 2023

Approved for public release; distribution is unlimited.

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

The research presented in this report was supported by the Acquisition Research Program at the Naval Postgraduate School.

Management and Business Knowledge Representation for Decision Making: Applying Artificial Intelligence, Machine Learning, Data Science, and Advanced Quantitative Decision Analytics for Making Better-Informed Decisions

Dr. Johnathan Mun— is a Professor of Research at the Naval Postgraduate School in Monterey, CA, and is a specialist in advanced decision analytics, quantitative risk modeling, Monte Carlo simulation, strategic flexibility real options, predictive modeling, and portfolio optimization. He has authored over 32 books, holds 22 patents and patents pending, and has written over 200 technical reports, white papers, analytical notes, and academic journal articles. His prior positions include being Vice President of Analytics at Oracle/Crystal Ball and a senior manager at KPMG Consulting. He has developed multiple software applications that run Monte Carlo risk simulations, predictive modeling, portfolio selection and optimization, strategic real options, analysis of alternatives, total ownership cost, and others. Mun has performed over 100 consulting projects in private industry as well as with the Department of Defense, in areas including decision analytics, return on investment analysis, acquisitions analysis, analysis of alternatives, enterprise risk management, program and project selection and optimization, cost and schedule risk modeling, and others. His latest technical reports with the Naval Research Program and Naval Acquisitions Research Program deal with the return on investment on military education and research, as well as total cost ownership modeling and return on investment on DARPA's 5G protected waveform open-source initiative. He holds a PhD in finance and economics from Lehigh University with specializations in advanced risk analytics, econometric modeling, financial analytics, decision analysis, and strategic real options. He also holds an MBA in management and a BS in physics and biology, as well as multiple other designations, such as the Certified in Quantitative Risk Management (CQRM), Certified in Financial Risk Management (FRM), and Chartered Financial Consultant (CFC), among others. [jcmun@nps.edu]

Abstract

How were the decisions made in the past, and what were the drivers, strategies, or rationale? The old adage holds true on how organizations should learn from the past to help make better decisions in the future. This current first-phase research looks at how the Department of Defense (DoD) can inculcate institutional and corporate memory. Specifically, the research tests and develops recommendations about how a transparent Decisions Options Register (DOR) integrated intelligent database system can be developed, where the DOR helps capture all historical decisions (assumptions, data inputs, constraints, limitations, competing objectives, and decision rules) for programs within the DoD. Information in this DOR will be compatible with meta-semantic searches and data science analytical engines. The DOR is used for modeling future decision options to enable making decisions under uncertainty while leaning on past best practices and allowing senior leadership to make defensible and practical decisions. The current first phase of research uses stylized data and examples to illustrate the recommended methodologies.

This research implements industry best-in-class decision analytics using advanced quantitative modeling methods (stochastic simulation, portfolio optimization) coupled with Artificial Intelligence (AI) and Machine Learning (ML) algorithms (data scraping, text mining, sentiment analysis) and Enterprise Risk Management (ERM) procedures. The DOR will be partially based on ERM methods of using risk registers, where different risk elements are subdivided into different GOPAD groups, or Goals (military capability, cost savings, novel technology, future weapons capability, public safety, government priorities, command preference, etc.), Organization (Air Force, Army, Navy, Marines), Programs (acquisition, commercial-off-the-shelf, joint-industry, hybrid, etc.), Activity (inventory, replacement, new development, research and development, and so forth), and Domain (air, sea, cyber, etc.) categories.

Multiple competing stakeholders (e.g., the Office of the Secretary of Defense, Office of the Chief of Naval Operations, the U.S. Congress, and the civilian population) have their specific objectives (e.g., capability, efficiency, cost-effectiveness, competitiveness, and lethality, as well as alternatives and trade-offs), constraints (e.g., time, budget, schedule, and manpower), and mission-based domain requirements (e.g., balancing the needs of digital transformation in cybersecurity, cyber-counterterrorism, anti-submarine warfare, anti-aircraft warfare, or missile defense).

This research takes a multidisciplinary approach where methods from advanced analytics, artificial intelligence, computer science, decision analytics, defense acquisitions, economics, engineering and physics, finance, options theory, project and program management, simulation with stochastic modeling, applied mathematics, and statistics are applied. The ultimate goals are to provide decision-makers actionable intelligence and visibility into future decision options or flexible real options, complete with the assumptions that led to certain comparable decisions.

The recommended approaches include the use of supervised and unsupervised AI/ML sentiment text analysis, AI/ML natural language text processing, and AI/ML logistic classification and support vector machine (SVM) algorithms, coupled with more traditional advanced analytics and data science methods such as Monte Carlo simulation, stochastic portfolio optimization and project selection, capital budgeting using financial and economic metrics, and lexicographic rank approaches like PROMETHEE and ELECTRE.

Example case applications, code snippets, and mock-up DORs are presented, complete with stylized data to illustrate their capabilities. The current research outcome will provide a stepping stone to the next phase's multiyear research, where prototypes can be built and actual data can be run through the prescribed analytical engines.

Introduction

The purpose of this proposed research is to generate a transparent Decisions Options Register (DOR) integrated intelligent database system that helps to capture all historical decisions going forward, including their assumptions, data inputs, constraints, limitations, competing objectives, and decision rules for the Department of Defense (DoD). Information in this DOR will be compatible with meta-semantic searches and data science analytical engines. The DOR is used for modeling future decision options to implement and enable making decisions under uncertainty while leaning on past best practices and allowing senior leadership to make defensible and practical decisions.

The DOR is based on Enterprise Risk Management (ERM) practices in private industry, which typically lists risks and lessons learned from past, current, and proposed future projects. The creation of a documentation database of decision history is critical. There is no learning curve if there is no curve, and you cannot have a curve without any data or information. With the recommended DOR and associated methodologies in this current research, we can compute probabilities of the success and failures of a new program by looking at its characteristics and using historical data as a reference to predict the outcomes. Of course, there will be a need to operationalize and define success versus failure. Just because a program is under budget, on time, requires little rework, and hits all the required specifications and technology release levels, does it mean it is successful? What other metrics might we use to determine definite success or definite failure, and what about all the other levels in between? We need to identify available data as well as the gaps to get us a solid DOR. What are some statistically significant predictors of success and failures as we have operationally defined them? The other issue is risk mitigation and strategic flexibility.

This research will showcase industry best-in-class decision analytics and ERM procedures. The DOR will be partially based on ERM methods of using risk registers, where different risk elements are subdivided into different GOPAD groups, or *Goals* (military capability, cost savings, novel technology, future weapons capability, public safety, government priorities, command preference, etc.), *Organization* (Air Force, Army, Navy, Marines), *Programs* (acquisition, commercial-off-the-shelf, joint-industry, hybrid, etc.), *Activity* (inventory, replacement, new development, research and development, and so forth), and *Domain* (air, sea, cyber, etc.) categories.

Multiple competing stakeholders (e.g., the Office of the Secretary of Defense, Office of the Chief of Naval Operations, the U.S. Congress, and the civilian population) have their specific objectives (e.g., capability, efficiency, cost-effectiveness, competitiveness, and lethality, as well as alternatives and trade-offs), constraints (e.g., time, budget, schedule, and manpower), and mission-based domain requirements (e.g., balancing the needs of digital transformation in cybersecurity, cyber-counterterrorism, anti-submarine warfare, anti-aircraft warfare, or missile defense). These elements are critical when new decisions are to be considered. A DOR database that preserves institutional knowledge and memory will assist in such endeavors and instill trust in the decisions.

This research will take on a multidisciplinary approach where we will be applying methods from advanced analytics, artificial intelligence, computer science, decision analytics, defense acquisitions, economics, engineering and physics, finance, options theory, project and program management, simulation with stochastic modeling, applied mathematics, and statistics. The ultimate goals are to provide decision-makers actionable intelligence and visibility into future decision options or flexible real options, complete with the assumptions that led to certain comparable decisions.

Research Current State-of-the-Art

In a legal dispute, courts use precedents when deciding the outcomes of cases. The use of precedence has been in practice for over 200 years, often to appeal or overturn previous judgments. However, precedent-based decision-making is something that industries and governments have not yet fully embraced. Organizations, including the DoD, tend to have a short memory due to the fluctuations and outflows of human capital and the loss of institutional knowledge when employees leave or are reassigned elsewhere. The current research is intended to include an examination of how related research into the state of the art of precedent-based decision-making is performed today, what might be considered state of the art, and what its current limitations are.

Research Approach

The research applies multiple novel approaches to enhance its success in generating a powerful and searchable DOR database. The recommendations will include key parameters, assumptions, input data, saved models and computations, decisions made, leadership inputs and overrides, constraints and limitations, end goals, and other pertinent information, which can then be mined using *Sentiment Analysis* with *Machine Learning*, coupled with *Scraping Algorithms* and *Text Mining* with *Custom Lexicographic* sets. Users of the system will be able to apply precedent-based insights into their current and future programs. In addition, whenever possible, predictive values will be complemented by actual values captured over time. This allows postmortem analysis of previous programs and provides for lessons learned along the way. Capturing the history of key decisions will help senior leadership make more credible and defensible decisions, which may eventually lead to legal and regulatory changes for the DoD.

The proposed methodologies will allow the collection of data that can be applied in a variety of areas, including, but not limited to, Integrated Risk Management® approaches where *stochastic analyses* like *Monte Carlo simulations, stochastic portfolio optimization*, and advanced *data analytical* approaches, *artificial intelligence*, and *data science* methods can be run. Over time, lookback analyses can be applied to update the DOR, making it more closely aligned with the needs of the DoD. The system should be able to collect different types of economic data (total lifecycle cost, total ownership cost, acquisition cost, cost deferred, and schedule and risk costs), logistics data (e.g., inherent availability, effective availability, mission reliability, operational dependability, mean downtime, mean maintenance time, logistics delay time, achieved availability, operational availability, mission availability, fielded capabilities, and Likert levels of creative and novel technology, as well as other metrics), qualitative subject matter expert estimates (strategic value, value to society, command priorities, legal and regulatory impact scores, etc.), and market comparables to operationalize various elements of DoD benefit. At appropriate time intervals, backfitting analyses such as *nonlinear discriminant analysis, neural networks, distributional fitting, limited dependent variables, path-dependent partial least squares,* and others can be applied to tease out the *critical success factors* that lead to the success or failure of certain decisions within a program or acquisition.

Research Application

The current research is important because it will create a significant difference in the DoD's decision-making process. The DoD is continually looking for better theoretically justifiable and quantitatively rigorous analytical methods for decision analysis, capital budgeting, and portfolio optimization. The specific interest lies in how to identify and quantify the value of each program to the military and optimally select the correct mix of programs, systems, and capabilities that maximizes some military value (strategic, operational, or economic) while subject to budgetary, cost, schedule, and risk constraints. This research applies private-sector and industry best practices coupled with advanced analytical methods and models to help create these methodologies to do so. However, the uniqueness of the DoD requires that additional work be done to determine the concept of value to the military while considering competing stakeholders' needs. The DoD requires defensible and quantitatively robust concepts of military value in its return on investment for making optimal funding decisions such as where, how much, and how long to invest. These decision options (strategic sequential compound real options, optimal timing options, growth options, and other options to expand, contract, and abandon) are critical when performing an analysis of alternatives and balancing cost-benefit trade-offs in a non-economic DoD environment. The DOR will provide historically preserved insights into the various alternate futures assumed, the alternatives modeled, and why certain decisions were made.

Artificial Intelligence and Data Science

Artificial Neural Network (NN) is a data-driven, distribution-free nonparametric family of methods that can be used for nonlinear pattern recognition, prediction modeling, and forecasting. NN is often used to refer to a combinatorial network circuit of biological neurons. The modern usage of the term often also refers to "artificial neural networks," comprising artificial neurons, or nodes, recreated within a software environment. Such artificial networks attempt to mimic the neurons or neuronal nodes in the human brain in terms of the way humans think, identify patterns, and, in our situation, identify patterns for forecasting time-series data. NN methods can be used in well-behaved time series as well as chaotic physical systems. When used in Big Data (BD) and in conjunction with Machine Learning (ML) approaches, it can be considered as a cross-over to a semi-supervised Artificial Intelligence (AI) system. NN is still considered semi-supervised, as neural networks require a multilayered training process as part of the activation function. For instance, the neural node weights and interactive convolution can be run autonomously once the activation is triggered in the system. In multilayered neuronal nodes, the results from the first node layer will become the inputs into subsequent layers of nodes.

This paper proposes the addition of an internal optimization process to be iteratively run to continually train the nodes to minimize a series of error measurements, such as the standardized sums of squares of errors while balancing and constraining the Akaike Information Criterion, Bayes Criterion, and Hannan-Quinn Criterion. In addition, the proposal here is to add a Combinatorial Fuzzy Logic methodology to the mix to generate the best possible forecast. The term *fuzzy logic* is derived from fuzzy set theory to deal with reasoning that is approximate rather than accurate. As opposed to crisp logic, where binary sets have binary logic, fuzzy logic variables may have a truth value that ranges between 0 and 1 and is not constrained to the two truth values of classic propositional logic. This fuzzy weighting schema is used together with a combinatorial method to yield time-series forecast results.

Augur (2016) provides a good summary of the history of data science. According to his research, the term "data science" first appeared as early as 1974, when Peter Naur published his article entitled "Concise Survey of Computer Methods" and defined it as "the science of dealing with data, once they have been established, while the relation of the data to what they represent is delegated to other fields and sciences." The term took a while to catch on, having not fully integrated the vernacular until 2010. The term "data scientist" is often attributed to Jeff Hammerbacher and D. J. Patil, of Facebook and LinkedIn, in 2008. Between 2011 and 2012, "data scientist" job listings increased by 15,000%, with an emphasis on working with Big Data. By 2016, data science started to become entrenched in the fields of Artificial Intelligence, specifically in the subfields of Machine Learning and Deep Learning.

Literature Review

Artificial intelligence (AI) is a broad term that refers to a variety of technologies. It's a catch-all term for a group of inorganic computer science technologies that are used to simulate intelligence. The word AI is often associated with the hazy notion of machine learning, which is a subset of AI in which a computer system is trained to recognize and categorize external real-world data. It is "The ability of machines to perform tasks that normally require human intelligence—for example, recognizing patterns, learning from experience, drawing conclusions, making predictions, or taking action—whether digitally or as the smart software behind autonomous physical systems," according to the DoD's (2019) AI strategy. The DoD (2019) is particularly interested in these expanded automation capabilities since prospective future near-peer enemies such as Russia and China are investing extensively in this field for military purposes. Given the vast AI field of study, this study focuses on the AI processes deemed most ideal for procurement, such as Machine Learning (ML), Natural Language Processing (NLP), and Robotic Process Automation (RPA), as illustrated in Figure 1 (modified from Sievo [2019]). The image depicts AI as a combination of AI sciences such as machine learning and natural language processing, and while RPA benefits from AI applications, it is not a simulation of human intelligence, but rather a mimic of skills.

The science of AI was established in 1956 to determine whether inorganic robots could execute human-level intelligence capabilities (Denning, 2019). It went through various hype cycles, mostly as a result of sensationalizing what it could do, with numerous disappointments (Figure 2). Significant interest in AI resurfaced about the same time as Big Data computer capacity became more widely available to researchers and businesses, allowing them to apply the science to a variety of practical applications (Haenlein & Kaplan, 2019). Manufacturing robots, smart assistants, proactive healthcare management, illness mapping, automated financial investing, virtual travel booking agents, social media monitoring, conversational marketing bots, NLP tools, and contract management are all examples of commercially feasible AI applications (Daley, 2019).

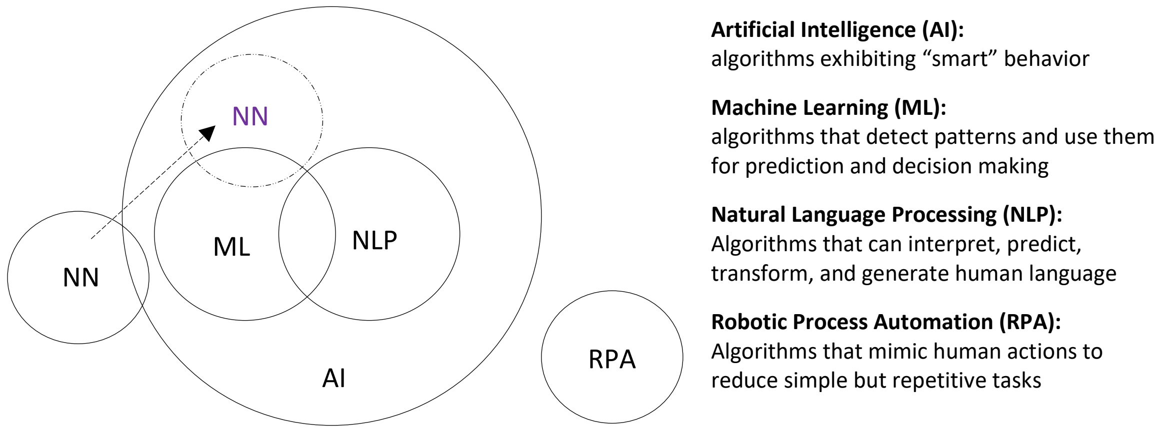


Figure 1. Types of Artificial Intelligence

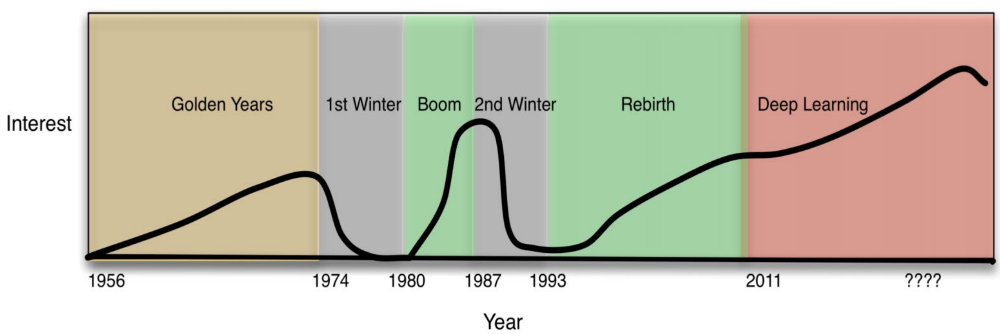


Figure 2. The Timeline of Interest in AI During Different Phases of Its Development (Denning, 2019)

A Brief History of Data Science

Several good explanations of the history of data science can be found in Press (2013) and Augur (2016). The timeline of data science development is summarized here. We can see how mathematical statistics has evolved into applied statistics, data science, artificial intelligence, and machine learning.

1962: John Tukey wrote "The Future of Data Analysis," and as a mathematical statistician, he considered his critical expertise as one able to analyze data.

1974: Peter Naur published the "Concise Survey of Computer Methods," where he coined the term data science. He defined it as "the science of dealing with data, once they have been established, while the relation of the data to what they represent is delegated to other fields and sciences." This term took a while to catch on.

1977: The International Association for Statistical Computing (IASC) was founded. Its main goal was to "link traditional statistical methodology, modern computer technology, and the knowledge of domain experts to convert data into information and knowledge."

1994: The early forms of modern marketing began to appear, with the main emphasis on Database Marketing.

1996: The term Data Science appeared for the first time at the International Federation of Classification Societies in Japan. The inaugural topic was entitled "Data Science, Classification, and Related Methods."

1997: Jeff Wu gave an inaugural lecture titled simply "Statistics = Data Science?"

2001: William Cleveland published "Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics." He put forward the notion that data science was an independent discipline and named six areas in which he believed data scientists should be educated: multidisciplinary investigations, models and methods for data, computing with data, pedagogy, tool evaluation, and theory.

2008: The term "data scientist" is often attributed to Jeff Hammerbacher and DJ Patil of Facebook and LinkedIn.

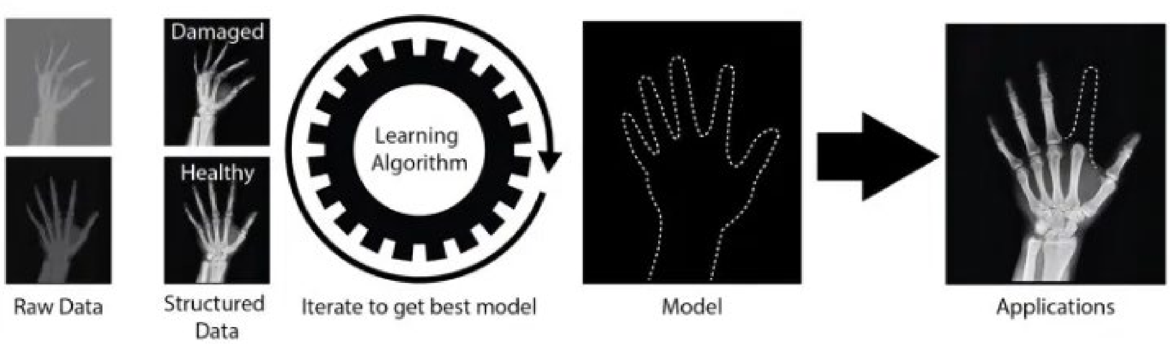
2010: The term "data science" has fully infiltrated the vernacular. Between 2011 and 2012, "data scientist" job listings increased by 15,000%.

2016: Data science started to be entrenched in Machine Learning and Deep Learning.

Machine Learning

Intelligence is the ability to process a specific sort of data, allowing a processor to solve significant problems (Gardner, 1993). Beyond the traditional idea of a person's intelligence quotient (IQ), which can often simply evaluate how well someone performs on an IQ test rather than their natural talents, psychologists have postulated multiple categories of intelligence. Howard Gardner (1993) proposed a theory of multiple intelligence, which suggests that traditional psychometric views of intelligence are too narrow. Intelligence should be expanded to include more categories in which certain processors, in this case, people, are better at making sense of different stimuli than others. Visual-spatial, linguistic-verbal, interpersonal, intrapersonal, logical-mathematical, musical, body-kinesthetic, and naturalistic intelligence are some of the categories of intelligence (Gardner, 1993). A counter-argument would be that these categories essentially represent learned and disciplined habits that people adopt over their lives as a result of their personality and surroundings. Regardless, both definitions of intelligence (traditional and many) are relevant to the stages involved in creating an artificial intelligence machine.

A computer is capable of doing computations and returning a response based on the data provided. It can be programmed and configured to repeat particular stages or algorithms and even change its conclusions based on previously calculated results using error-correcting techniques. The underlying principle of machine learning is a combination of these two phases. A computer system is fed data that is structured in such a way that the algorithm can identify it, deduce patterns from it, and make assumptions about any unstructured data that is presented later (Greenfield, 2019). In an x-ray learning method, Figure 3 explains how this works.



The image shows the steps an AI algorithm goes through to make a recommendation to a physician on where a missing body part should be. It takes in structured data and develops its understanding of what "right" looks like. When given unstructured data, it compares the image against previously trained models and identifies the abnormality with a recommendation on where to apply a fix, such as a prosthetic.

Figure 3. AI Training Algorithm
(Greenfield, 2019)

Supervised Learning

An algorithm is taught the patterns using past data and then detects them automatically in new data. Supervision comes in the form of correct answers that humans provide to train the algorithm to seek out patterns in data. This is commonly used within procurement areas such as spend classification (Sievo, 2019).

Unsupervised Learning

The algorithm is programmed to identify and potentially detect patterns in new data. Without any human supervision, the algorithm is not expected to surface specific correct answers; instead, it looks for logical patterns within raw data. This is rarely used within critical procurement functions (Sievo, 2019).

Reinforcement Learning

The algorithm helps to make decisions on how to act in certain situations, and the behavior is rewarded or penalized depending on the consequences. This is largely theoretical in the procurement context (Sievo, 2019).

Deep Learning

Deep learning is an advanced class of machine learning inspired by the human brain where artificial neural networks progressively improve their ability to perform a task. This is an emerging opportunity in procurement functions (Sievo, 2019).

Natural Language Processing

Anyone who has used devices that appear to be able to understand and act on written or spoken words, such as translation apps or personal assistants like Amazon's Alexa, is already familiar with NLP-enabled AI. NLP is a set of algorithms for interpreting, transforming, and generating human language in a way that people can understand (Sammalkorpi & Teppala, 2019). Speech soundwaves are converted into computer code that the algorithms understand. The code then translates that meaning into a human-readable, precise response that can be applied to normal human cognition. This is performed using semantic parsing, which maps a passage's language to categorize each word and, using machine learning, creates associations to represent not just the definition of a term but the meaning within a specific context (Raghaven & Mooney, 2013). Figure 4 depicts this categorization and analysis procedure in the context of a DoD procurement contract.

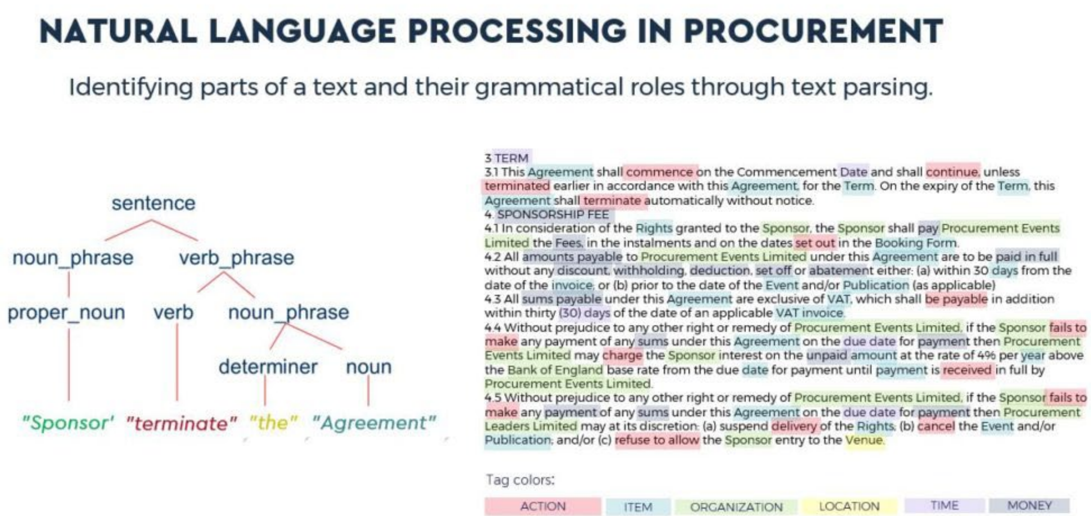


Figure 4. Semantic Parsing in Procurement
(Sievo, 2019)

Robotic Process Automation

RPA is not AI, as previously stated; rather, it is an existing process that has been augmented by AI. RPA is defined as "the use of technology by employees in a firm to set up computer software or a robot to capture and interpret current applications for processing transactions, altering data, triggering reactions, and communicating with other digital systems" (Institute for Robotic Process Automation & Artificial Intelligence, 2019). When used correctly, robotic automation offers numerous benefits because it is not constrained by human limitations such as weariness, morale, discipline, or survival requirements. Robots, unlike their human creators, have no ambitions. Working harder will not get you more money or get you promoted, and being permanently turned off will have no effect because robotic automation just duplicates the practical parts of the human intellect, not the underlying nature of mankind (Zarkadakis, 2019). (Note, however, that machine learning relies on an incentive system to make judgments about positive or negative reactions.) A future AI-enabled RPA option is for a machine to learn how to control the source of positive reinforcement fully independent of the rules required to achieve its aim. Things that survive evolve to stay alive because of positive reinforcement from their surroundings and the fact that they continue to act in a way that is regarded as survivable. This should be taken into account in any future AI efforts, especially in the case of why a human must always be present when final judgments are made. Regardless of whether or not AI systems have a perfect track record, they should not be entirely trusted.

Proposed Artificial Intelligence, Machine Learning, and Advanced Quantitative Methodologies

Decision Options Database

As discussed, the purpose of this research is to create a transparent Decisions Options Register (DOR) integrated intelligent database system that helps to capture all historical decisions going forward, including their assumptions, data inputs, constraints, limitations, competing objectives, and decision rules for the DoD. Information in this DOR will be compatible with meta-semantic searches and data science analytical engines. The DOR is used for modeling future decision options to implement and make decisions under uncertainty while leaning on past best practices and allows senior leadership to make defensible and practical decisions.

AI/ML Data Reduction and Classification and Logistic Predictive Modeling

The dataset comprises textual information as well as whether the project was successful (completed) or failed (the program was rejected or canceled). Using the quantitative variables, AI/ML classification routines can be applied to determine the probability that a potential or future program will also be successful or fail.

The classification routine we will use applies in the situation where the dependent variable contains data that are limited in scope and range, such as binary responses (0 or 1 for failures/successes), truncated, ordered, or censored data. For instance, given a set of independent variables (e.g., age, income, education level of credit card or mortgage loan holders), we can model the probability of defaulting on mortgage payments using maximum likelihood estimation (MLE). The response or dependent variable Y is binary. That is, it can have only two possible outcomes that we denote as 1 and 0 (e.g., Y may represent the presence/absence of a certain condition, defaulted/not defaulted on previous loans, success/failure of some device, answer yes/no on a survey, etc.) and we also have a vector of independent variable regressors X, which are assumed to influence the outcome Y. A typical ordinary least squares regression approach is invalid because the regression errors are heteroskedastic and non-normal, and the resulting estimated probability estimates will return nonsensical values of above 1 or below 0. MLE analysis handles these problems using an iterative optimization routine to maximize a log-likelihood function when the dependent variables are limited.

A Logit or Logistic regression is used for predicting the probability of occurrence of an event by fitting data to a logistic curve. It is a generalized linear model used for binomial regression, and like many forms of regression analysis, it makes use of several predictor variables that may be either numerical or categorical. MLE applied in a binary multivariate logistic analysis is used to model the dependent variable to determine the expected probability of success of belonging to a certain group. The estimated coefficients for the Logit model are the logarithmic odds ratios, and they cannot be interpreted directly as probabilities. A quick computation is first required, and the approach is simple.

Specifically, the Logit model is specified as Estimated $Y = LN[P_i/(1 - P_i)]$ or, conversely, $P_i = EXP(Estimated\ Y)/(1 + EXP(Estimated\ Y))$, and the coefficients β_i are the log odds ratios. So, taking the antilog or $EXP(\beta_i)$, we obtain the odds ratio of $P_i/(1 - P_i)$. This means that with an increase in a unit of β_i, the log odds ratio increases by this amount. Finally, the rate of change in the probability $dP/dX = \beta_i P_i (1 - P_i)$. To estimate the probability of success of belonging to a certain group (e.g., predicting if a program will develop issues and eventually fail given a certain combination of lifecycle cost, ROI, FTE requirements, length of time, strategic value, etc.), we simply compute the $Estimated\ Y$ value using the MLE coefficients and convert it into the inverse antilog of the odds ratio as discussed previously. Next, we can take the statistically significant variables and apply them to a Gaussian Support Vector Machine (SVM) to classify the programs into high probabilities of approval or rejection categories.

First Step

Model Inputs:
VAR1
VAR2; VAR3; VAR4; VAR5; VAR6; VAR7; VAR8; VAR9
Status (D)
Monthly FTE, Complexity Level, Strategic Value, Value to Command, Length in Months, Program Cost, Overrun Ratio, Annual Cost Savings

Generalized Linear Model (Logit with Binary Outcomes)

	Coefficient	Std. Error	Wald Test	P-value	Exp(B)	Lower	Upper
Intercept	-1.634198	0.754434	4.692098	0.030302	0.195109	0.000000	0.000000
VAR1	0.028625	0.020496	1.950585	0.162524	1.029039	0.988520	1.071218
VAR2	0.076812	0.144371	0.283071	0.594695	1.079839	0.813711	1.433004
VAR3	-0.262500	0.040630	41.7411	**0.000000**	0.769127	0.710254	0.832879
VAR4	-0.096195	0.027419	12.3083	**0.000451**	0.908287	0.860764	0.958434
VAR5	0.000823	0.012687	0.004210	0.948266	1.000824	0.976243	1.026022
VAR6	0.074324	0.039911	3.467833	**0.062573**	1.077155	0.996106	1.164799
VAR7	0.564136	0.134590	17.5689	**0.000028**	1.757929	1.350325	2.288569
VAR8	0.049994	0.101851	0.240943	0.623526	1.051265	0.861027	1.283535

Log-Likelihood	-199.9830
Restricted Log-Likelihood	-285.4773
McFadden’s R-Squared	0.299479
Cox and Snell’s R-Squared	0.289636
Nagelkerke’s R-Squared	0.425440
Raw Akaike Info. Criterion	417.9659
Raw Bayes Criterion	455.8974

Log-Likelihood	-199.9830
Restricted Log-Likelihood	-285.4773
Chi-Square	170.9886
Degrees of Freedom	8
P-value	0.000000

Second Step

Model Inputs:
VAR1
VAR4; VAR5; VAR7; VAR8
Status (D)
Strategic Value, Value to Command, Program Cost, Overrun Ratio
Generalized Linear Model (Logit with Binary Outcomes)

	Coefficient	Std. Error	Wald Test	P-value	Exp(B)	Lower	Upper
Intercept	-0.781188	0.305330	6.545958	0.010512	0.457862	0.000000	0.000000
VAR1	-0.239706	0.033215	52.0818	**0.000000**	0.786859	0.737266	0.839788
VAR2	-0.074519	0.023632	9.942889	**0.001615**	0.928190	0.886178	0.972194
VAR3	0.082202	0.022767	13.0359	**0.000306**	1.085675	1.038294	1.135218
VAR4	0.588673	0.108123	29.6424	**0.000000**	1.801597	1.457549	2.226855

Log-Likelihood	-201.7171
Restricted Log-Likelihood	-285.4773
McFadden's R-Squared	0.293404
Cox and Snell's R-Squared	0.284691
Nagelkerke R-Squared	**0.418177**
Raw Akaike Info. Criterion	413.4342
Raw Bayes Criterion	434.5072

Log-Likelihood	-201.7171
Restricted Log-Likelihood	-285.4773
Chi-Square	167.5204
Degrees of Freedom	4
P-value	0.000000

Third Step

Model Inputs:
Status (D)
Strategic Value, Value to Command, Program Cost, Overrun Ratio
Sigma, Lambda, Omega, Calibration Level: 1.00, 1.00, 0.40, 1.00

AI Machine Learning: Classification with Gaussian SVM (Supervised)

Relax: 8.218332

Accuracy	68.20%	67.40%	68.20%	**69.40%**	67.80%	67.80%	66.60%	65.00%	63.40%	62.20%
Omega	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00

Forecast	Group
1.118101	1.00
0.971805	0.00
. . .	. .
. . .	. .
0.971828	1.00

Stochastic Simulation and Probabilistic Analysis

Another recommended approach is to perform stochastic distributional fitting; that is, how do the collected historical data fit known probability distributions? These fitted distributions can be used as the variable's input parameters (e.g., a Fréchet or Weibull distribution with shape and scale parameters of 0.5 and 1.2). Figure 5 illustrates an example where historical program costs were fitted to determine its distributional properties. With the fitted distribution, these can be used as inputs into a Monte Carlo simulation model to forecast and predict a new program's chances of success.

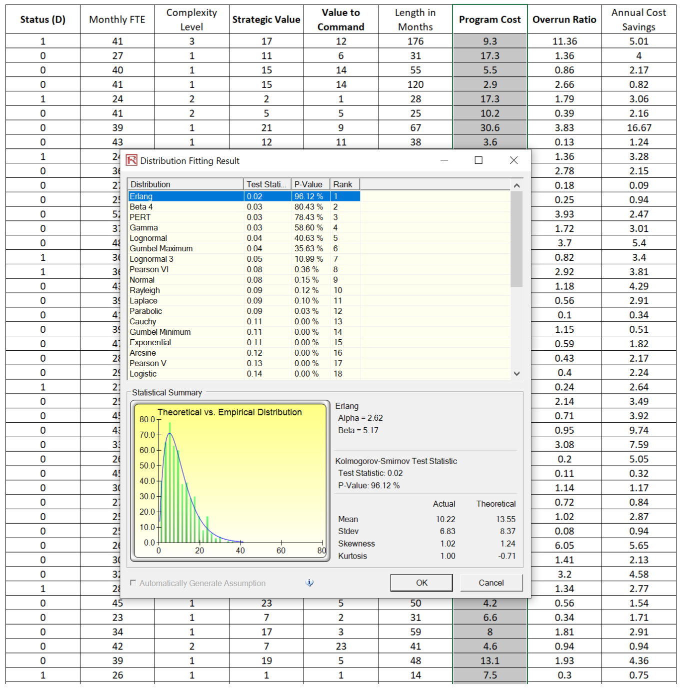


Figure 5. Distributional Fitting for Probabilistic Analysis and Stochastic Simulation

Predicted	Actual	Correct 72.4%	Datapoints
positive	negative	0	a few months before the release of star wars episode 1 , the phantom menace , 20th century fox decides to release another space film
negative	negative	1	bad movies described as " a swift descent into sinful pleasure , decay , and debauchery " are hard to watch . bad 2000 ' s movies that r
positive	negative	0	bruce willis needs to stay away from straightforward action pictures . mercury rising adds to a growing list (including such stinkers as th
negative	negative	1	capsule : godawful " comedy " that ' s amazingly shabby and cut - rate , and rather bereft of laughs . i was having a bad week in my life
negative	negative	1	battlefield earth is the worst film of 2000 , and i guarantee you that nothing else this year will even come close . in fact , i ' ll be surprised
negative	negative	1	jet li busted onto the american action movie scene , when he stole the show in 1998 ' s lethal weapon 4 , with his wicked looks , his nast
negative	negative	1	one night , during a torrential downpour that flooded the streets , we went to see -- what else -- hard rain . " so , are we all going to die
negative	negative	1	it was with a huge lack of something to do that i decided to watch this on good old upn on sunday afternoon , when the only good things
negative	negative	1	vampires starts out almost in the style of a spaghetti western with an attack on a small homestead in new mexico . the house has a nest
positive	negative	0	when considering david fincher ' s latest film , " the game " , four words come to mind . " don ' t believe the hype . " this michael douglas
positive	negative	0	weighed down by tired plot lines and spielberg ' s reliance on formulas , _saving private ryan_ is a mediocre film which nods in the directi
positive	negative	0	if you ' re going to make a two - hour hollywood in - joke , why bother releasing it to the general public ? if you ' re going to create a film
positive	negative	0	it used to be that not just anyone could become a vampire . usually , you had to be an aristocrat - a count such as dracula or karnstein .
positive	negative	0	hav plenty , as we are told in the beginning and reminded during the film , is a true story . life itself is a series of true stories , but most a
positive	negative	0	it seems that i ' ve stopped enjoying movies that should be fun to watch . take payback , for example , a movie that most people seem to
negative	negative	1	note : some may consider portions of the following text to be spoilers . be forewarned . " quick , robin ! the anti - shark repellant ! " - ada
negative	negative	1	everything about this ninth trek movie seems on the cheap , from the roger corman - grade special effects to its highly derivative and ugl
positive	negative	0	this is one of the worst big - screen film experiences i ' ve had for a while . with this film , plus ` showgirls ' and ` basic instinct ' , paul ver
negative	negative	1	phew , what a mess ! for his fifth collaboration with director rich - ard donner (lethal weapon i - iii , maverick) , mel gibson plays a moto
negative	negative	1	in french , the phrase " film noir " literally means " black film . " webster defines it as " a type of crime film featuring cynical malevolent ch
positive	negative	0	plot : lara croft is british , rich and kicks a lot of ass . she also likes to raid tombs but when the illuminata discover that all nine planets ar
negative	negative	1	it happens every year -- the days get longer , the weather gets warmer and the studios start releasing their big - budget blockbusters . t
negative	negative	1	i heard actor skeet ulrich discussing this film in a couple of interviews , and in both instances , he felt the strange compulsion to compare
negative	negative	1	it is with some sad irony that i screened fright night part 2 on the day that one of it ' s stars , roddy mcdowall passed away at the age of
positive	negative	0	sometimes a stellar cast can compensate for a lot of things , and " pushing tin " certainly features some name stars who are going place
negative	negative	1	the most depressing thing about the depressingly pedestrian james bond film " the world is not enough " is its final frame : white letters o
negative	negative	1	when walt disney pictures announced a live - action feature based on the ' 60s cartoon series of " mr . magoo , " special interests groups
positive	negative	0	david spade has a snide , sarcastic sense of humor that works perfectly on the tv sitcom just shoot me . it also served as a good showc
positive	negative	0	9 : its pathetic attempt at " improving " on a shakespeare classic . 8 : its just another piece of teen fluff . 7 : kids in high school are not th
negative	negative	1	sometimes i wonder just what the censors are thinking . take this film , " naked killer " , among it ' s ingredients are heavy doses of violer
positive	negative	0	after seeing blaze and driving miss daisy , i was ready for some mindless fun -- oh , maybe something like tango & cash . maybe not ! m
negative	negative	1	in " the 13th warrior , " arab poet ahmed ibn fahdlan (antonio banderas) finds himself kicked out of baghdad for feeling up the king ' s ol
positive	negative	0	there isn ' t much good about this movie . not much i can say about the acting , directing , or writing that would make you consider seeing
positive	negative	0	" varsity blues " is the best film of 1999 thus far . unfortunately , it is also the first film i have seen from 1999 . it is another one of those s
negative	negative	1	for a film touted as exploring relationships and black sexuality , trois is surprisingly tame . despite it ' s lurid subject matter and it ' s pass
positive	negative	0	" mission to mars " is one of those annoying movies where , in the middle of the movie , you get the sneaking suspicion that the reason tl
negative	negative	1	my friend here in film school just made a two minute - long film for one of his classes that includes a staged anal rape scene , done by tw
positive	negative	0	remember back in the mid 1990s when crime and macabre movies were all the rage ? " pulp fiction " and " fargo " both managed to get
positive	negative	0	_dirty_work_ has a premise of deliciously mean - spirited potential . mitch weaver (norm macdonald) and his lifelong best friend sam m
positive	negative	0	america ' s favorite homicidal plaything takes a wicked wife in " bride of chucky , " and their unholy matrimony is something old , nothing

Figure 6. Text Scraping and Sentiment Analysis Dataset and AI Classification Results

Lessons Learned in the Prototype Application

The prototype was very insightful in that it provided a myriad of lessons learned. For instance, the issue of hypernyms and hyponyms can be developed to create a hierarchical structure of a custom dictionary, whereby using text scraping methodologies, we can complement the learning algorithm with our custom lexicon. Sayings, proverbs, adages, and other types of word structures will also need to be considered, as will concise wordings or mixed negatives (e.g., “no good” is a negative connotation as opposed to a positive “good” implication despite the fact that the word exists in the context). The impact scores of certain words and their frequencies can also be used to generate word clouds and help create visuals of the most frequent and impactful comments from past programs.

Neural Network Pattern Recognition Prediction Methods

Using the Box-Jenkins method of forward-looking predictive steps, we have

$$\hat{x}_{t+1} = f(x_t, x_{t-1}, \dots, x_{t-n})$$

$$\hat{x}_{t+2} = f(x_{t+1}, x_t, \dots, x_{t-n+1})$$

$$\dots$$

$$\hat{x}_{t+k} = f(x_{t+k-1}, x_{t+k-2}, \dots, x_{t-n+k-1})$$

Where x_t is the observation of x at time t. This means that if we use a k step ahead predictive model, we have

$$x_{t+1} = f_1(x_t, x_{t-1}, \dots, x_{t-n})$$

$$x_{t+2} = f_2(x_t, x_{t-1}, \dots, x_{t-n})$$

$$\dots$$

$$x_{t+k} = f_k(x_t, x_{t-1}, \dots, x_{t-n})$$

Here, we see that f_i are computed in the neural network paradigm.

Activation Transfer Functions for Neural Networks

Logistic sigmoidal function: $f(x) = (1 + e^{-x})^{-1}$

Hyperbolic tangent function: $f(x) = (e^x - e^{-x})(e^x + e^{-x})^{-1}$

Sine and cosine function: $f(x) = sin(x)\ or\ f(x) = cos(x)$

Linear function: $f(x) = x$

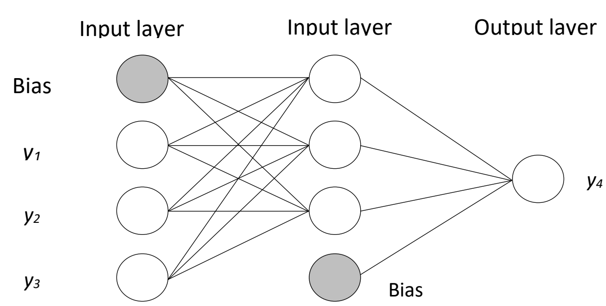


Figure 7. A Multiple Layered Perceptron Neural Network

The neural mapping assumes that y_4 is the dependent variable, whereas y_1, y_2, y_3 and a constant term is the set of independent variables. The neural network has an input layer, a hidden layer, and an output layer. There are three inputs in the input layer, a neuron for the biases, four neurons in the hidden layer, and one neuron in the output layer.

Error Measurements and Error Correction for Parameter Calibration

Total Variables (Dependent and Independent): v

Mean Absolute Deviation: $MAD = \frac{\sum|e_t|}{n}$

Root Mean Squared Error: $RMSE = \sqrt{\frac{\sum(e_t)^2}{n}}$

Sums of Squared Errors: $SSE = \sum(e_t)^2$

Maximum Log-Likelihood: $MLL = \frac{n}{2} ln(2\pi) - \frac{n}{2} ln \frac{SSE}{2} - SSE\left[\frac{n}{2SSE}\right]$

Akaike Information Criterion: $AIC = \frac{-2MLL}{n} + \frac{2k}{n}$

Bayes Information Criterion (BIC): $BC = AIC + \frac{2(v+2)(k+3)}{n-k-3}$

Pesaran-Timmermann Test: $PT = \frac{p(xf)-p\prime}{\sqrt{v-w}}$ where $v = \frac{p\prime(1-p\prime)}{n}$, $p' = f^+x^+ + (1-f^+)(1-x^+)$, and where $w = \frac{(2f^+-1)^2x^+(1-x^+)}{n} + \frac{(2x^+-1)^2 f^+(1-f^+)}{n} + \frac{(4x^+f^+)(1-x^+)(1-f^+)}{n^2}$, $x^+$is the proportion of positives on the data and $f^+$is the proportion of positives on the forecast

Hannan-Quinn Information Loss Criterion: $HC = \frac{-2MLE}{n} + \frac{2kln(ln(n))}{n}$

Training Algorithms

For model 1:

Variables used in Training set	Predicted value
y_1, y_2, y_3	$\hat{y}_4$
y_2, y_3, y_4	$\hat{y}_5$
…	…
$y_{400}, y_{401}, y_{402}$	$\hat{y}_{403}$

Using the coefficients obtained from the training set, we do the following on the testing set:

Variables used in Testing set	Predicted value
$y_{401}, y_{402}, y_{403}$	$\hat{y}_{404}$
$y_{402}, y_{403}, y_{404}$	$\hat{y}_{405}$
…	…
$y_{420}, y_{421}, y_{422}$	$\hat{y}_{423}$

When forecasting, we use

Independent Variables	Predicted value
$y_{401}, y_{402}, y_{403}$	$\hat{y}_{404}$
$y_{402}, y_{403}, \hat{y}_{404}$	$\hat{y}_{405}$

For model 2:

Variables used in Training set	Predicted value
$y_1, y_2, y_3, y_4, y_5, y_6$	$\hat{y}_7$
$y_2, y_3, y_4, y_5, y_6, y_7$	$\hat{y}_8$
…	…
$y_{300}, y_{301}, y_{302}, y_{303}, y_{304}, y_{305}$	$\hat{y}_{306}$

Using the coefficients obtained from the training set, we do the following on the testing set:

Variables used in Testing set	Predicted value
y_{301}, y_{302}, y_{303}, y_{304}, y_{305}, y_{306}	$\hat{y}_{307}$
y_{302}, y_{303}, y_{304}, y_{305}, y_{306}, y_{307}	$\hat{y}_{308}$
…	…
y_{330}, y_{331}, y_{332}, y_{333}, y_{334}, y_{335}	$\hat{y}_{336}$

When forecasting, we use

Independent Variables	Predicted value
y_{301}, y_{302}, y_{303}, y_{304}, y_{305}, y_{306}	$\hat{y}_{307}$
y_{302}, y_{303}, y_{304}, y_{305}, y_{306}, $\hat{y}_{307}$	$\hat{y}_{308}$

Figure 8 shows a neural network for time series forecasting. The functions ϕ_0 and ϕ_h are called activation functions, and the logistic function and linear function are usually chosen. One of the input nodes is sometimes called the data bias node.

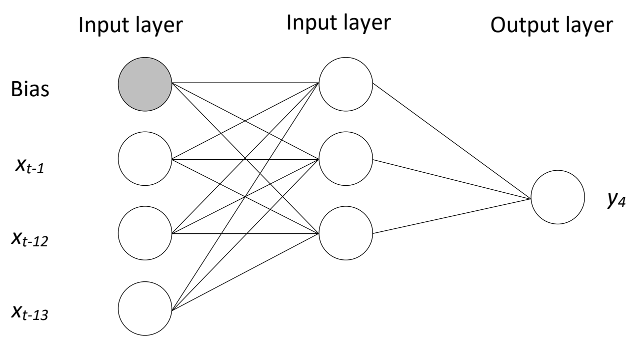


Figure 8. Example of Multiple Layered Perceptrons

The functional form to be modeled looks like this:

$$\hat{x}_t = \phi_0\left(w_{co} + \sum_h w_{ho}\phi_h\left(w_{ch} + \sum_h w_{ih}x_{t-j_i}\right)\right)$$

w_{ch} is the weight for the connections between the constant input and the hidden neurons, w_{co} is the weight of the direct connection between the constant input and output, and w_{ih} and w_{ho} are the weights for the other connections between the inputs and hidden neurons.

The model is built in the following steps:

- Obtain periodic time-series data.
- Calculate log-returns $log\frac{x_t}{x_{t-1}}$.
- Transform the log returns to the interval [0, 1]. The reason is that we use logistic functions in both the hidden layer and the output layer. The output of the logistic function lies between [0, 1].
- Decide which part of the data are as a training set, i.e., to train the neural network to obtain the weights w_{ih}, w_{ch}, w_{ho}, w_{co} in the figure, and which part of the data is used as the testing set. That is, after we train the neural network, we use those weights for forecasting and comparing with the data in the testing set.

- Decide on the inputs: x_{t-1}, x_{t-2}, x_{t-3} are used to predict x_t. Then, x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6} are used to predict x_t.
- Build the model to calculate the output.
- Use the data table to calculate the output by neural network for the data in the training set. Calculate the sum of squares of errors. Minimize this term using an internal optimization routine. We obtain the trained neural network, and the weights can be used for forecasting.
- Use the data table to calculate the output by neural network for the data in the testing set. Compare with true value.

Decision Analytics Methodology

A Combined Lexicographic Average Rank Approach for Evaluating Uncertain Multi-Indicator Matrices with Risk Metrics[1]

In many situations, projects are characterized by several criteria or attributes that can be assessed from multiple perspectives (financial, economic, etc.). Each criterion is quantified via performance values (PV), which can either be numerical or categorical. This information is typically structured in a multi-indicator matrix **Q**. A typical problem faced by a decision-maker is to define an aggregate quality (AQ) able to synthesize the global characteristics of each project and then derive the rankings from the best to the worst base-case ranking (Mun et al., 2016).

Ranking techniques can be classified as parametric and nonparametric. A parametric technique requires information about decision-maker preferences (e.g., criterion weights). According to Dorini, Kapelan, and Azapagic (2011), some examples of parametric techniques include the ELECTRE methods (Roy, 1968) and Preference Ranking Organization Methods for Enrichment Evaluations (PROMETHEE; Brans & Vincke, 1985). Nonparametric techniques, such as Partial Order Ranking (Bruggemann et al., 1999) and Copeland Scores (Al-Sharrah, 2010), do not require information from the decision-maker. In general, all of these techniques can produce a ranking of the alternatives from the best to the worst.

Therefore, given a matrix **Q**, the selected procedure generates a ranking defined as the base-case rank (BCR). As a result of this assessment, for each alternative, a specific rank R_i that considers the multiple perspectives defined by the decision-maker is obtained. The set of R_i corresponds to the global evaluation under the first synthetic attribute, defined and named as *base ranking* and capable of characterizing the alternatives in the base case.

However, in real-life situations, each performance value could be affected by uncertain factors. Several approaches have been presented for analyzing how the uncertainty in the performance values (the input) affects the ranking of the objects (the output; Rocco & Tarantola, 2014; Corrente et al., 2014; Hyde et al., 2004; Hyde & Maier, 2006; Yu et al., 2012). The approaches, based on Monte Carlo simulation, consider each uncertain factor as a random variable with known probability density functions. As a result, the AQ of each alternative and, therefore, the ranking also become random variables with approximated probability distributions. In such situations, the decision-maker could perform probability distribution evaluations. For example, the decision-maker could be interested in determining not only the worst rank of a specific alternative, but also its probability and volatility (risk evaluation).

---

[1] Some of the material discussed in this section is based on previous work by the author, Dr. Johnathan Mun, and his team, Dr. Elvis Hernández-Perdomo and Dr. Claudio M. Rocco. Their work has been published as a chapter, "A Combined Lexicographic Average Rank Approach for Evaluating Uncertain Multi-Indicator Matrices with Risk Metrics," in *Partial Order Concepts in Applied Sciences*, M. Fattore and R. Brüggemann (eds.), Springer International Publishing (2017).

In the standard approach, the probability of an alternative being ranked as in the BCR is selected as the synthetic attribute *probability* able to characterize the alternatives under uncertainty.

The stochastic nature of the AQ of each alternative could be further assessed to reflect the risk evaluation induced by uncertainty. In this case, it is required to compare several random variables synthesized through their percentiles and statistical moments. Several approaches have been proposed to this end, such as a simple comparison of the expected value, the expected utility (Von Neumann & Morgenstern, 1947), the use of low-order moments (Markowitz, 1952), risk measures (Jorion, 2007; Mansini et al., 2007; Rockafellar & Uryasev, 2000), the Partitioned Multiobjective Risk Method (PMRM; Asbeck & Haimes, 1984; Haimes 2009), and the stochastic dominance theory (Levy, 2006), among others.

To consider the risk evaluation induced by uncertainty, each alternative is represented by the third synthetic attribute: *compliance*. This new attribute is based on a simultaneous assessment of several risk measures and some moments of each AQ distribution (Mun et al., 2016).

At this point, each alternative is assessed from three different angles:

1. Multiple decision-making perspectives that include several aspects such as economic, financial, technical, and social (*base ranking*)
2. Uncertainty propagation on performance values (*probability*)
3. A risk evaluation based on the generated probability distribution (*compliance*)

These perspectives are then used for defining a new multi-indicator matrix $\mathbf{Q}_1$ correlated to projects and synthesized using a ranking technique. However, in some situations, decision-makers need to select projects following their most preferred criteria successively. For this reason, an aggregation ranking technique that allows compensation is useless.

Therefore, the final assessment is derived using a combined approach based on a *nonparametric aggregation rule* (using the concept of average rank) for attributes 1 and 2; a simple procedure for score assignment for attribute 3; and a *lexicographic rule*. In addition, a preliminary analysis of the alternatives is performed using a Hasse diagram (Bruggemann et al., 1999). To the best of the researcher's knowledge, this type of combined assessment has not been reported in the literature.

Average Rank Approach

Let P define a set of n objects (e.g., alternatives) to be analyzed and let the descriptors q_1, q_2..., q_m define m different attributes or criteria selected to assess the objects in P (e.g., cost, availability, environmental impact). Attributes must be defined to reflect, for example, that a low value indicates low rankings (best positions), while a high value indicates high ranking (worst positions; Restrepo et al., 2008). However, for a given problem or case study, this convention could be reversed.

If only one descriptor is used to rank the objects, then it is possible to define a total order in P. In general, given $x, y \in P$, if $q_i(x) \leq q_i(y)\ \forall i$, then x and y are said to be comparable. However, if two descriptors are used simultaneously, the following could happen: $q_1(x) \leq q_1(y)$ and $q_2(x) > q_2(y)$. In such a case, x and y are said to be incomparable (denoted by $x \| y$). If several objects are mutually incomparable, set P is called a partially ordered set or *poset*. Note that since comparisons are made for each criterion, no normalization is required.

The objects in a poset can be represented by a directed acyclic graph whose vertices are the objects $\in P$, and there is an edge between two objects only if they are comparable and one covers the other, that is, when no other element is in between the two. Such a chart is termed a Hasse diagram (Bruggemann et al., 1995).

A Hasse diagram is, then, a nonparametric ranking technique and can perform ranking decisions from the available information without using any aggregation criterion. However, while it cannot always provide a total order of objects, it does provide an interesting overall picture of the relationships among objects.

A useful approach to producing a ranking is based on the concept of the average rank of each object in the set of linear extensions of a poset (De Loof et al., 2011). Since the algorithms suggested for calculating such average ranks are exponential (De Loof et al., 2011), special approximations have been developed, such as the Local Partial Order Model (LPOM; Bruggemann et al., 2004), the extended LPOM (LPOMext; Bruggemann & Carlsen, 2011), or the approximation suggested by De Loof et al. (2011).

From the Hasse diagram, several sets can be derived (Bruggemann & Carlsen, 2011). If $x \in P$,

1. $U(x)$, the set of objects incomparable with x: $U(x) := \{y \in P : x || y\}$
2. $O(x)$, the *down* section: $O(x) := \{y \in P : y \leq x\}$
3. $S(x)$, the successor section: $S(x) := O(x) - \{x\}$
4. $F(x)$, the *up*: $F(x) := \{y \in P : x \leq y\}$

Then, the following average rank indexes are defined:

a) $LPOM(x) = (|S(x)| + 1) \times (n + 1) \div (n + 1 - |U(x)|)$

b) $LPOMext(x) = |O(x)| + \sum_{y \in U(x)} \frac{p_y^<}{p_y^< + p_y^>}$

where *n* is the number of objects,

$|V|$ defines the cardinality of the set *V*,

$p_y^< = |O(x) \cap U(y)|$, $p_y^> = |F(x) \cap U(y)|$, and $y \in U(x)$

Lexicographic Approach

A lexicographic technique enables decision-makers to develop choice rules in which they select more items based on their most important criteria. When two objects have the same influence on the most preferred criteria, decision-makers prefer the one with the biggest impact on the second most preferred criteria, and so on, according to Saban and Sethuraman (2014). This lexicographic form simulates situations in which decision-makers have a strong preference for one criterion over another or are in charge of non-compensatory aggregation (Yaman et al., 2011; Pulido et al., 2014).

Finally, decision-makers can model their strong preferences for the criteria chosen since, after additional investigation of the situation, they are neither indifferent nor uncertain about their preferences for the criteria considered. In other words, they will always favor one criterion over another, regardless of criterion weights.

Risk Metrics and Compliance

Risk metrics are statistical indicators or measurements that enable decision-makers to assess the dispersion (volatility) of specific events or outcomes. As a result, a random variable can be evaluated using statistical moments (e.g., mean, variance, skewness, kurtosis), or risk metrics, such as Value at Risk (VaR) and Conditional VaR, can be used to investigate extreme values (Bodie et al., 2009; Fabozzi, 2010; Matos, 2007; Mun, 2015).

Risk metrics are used to analyze the volatility or stability of a set of options or a portfolio of alternatives in decision problems, such as financial risk management (Chong, 2004), portfolio risk management (Bodie et al., 2009), enterprise risk management (Scarlat et al., 2012), and a variety of other areas (Fabozzi, 2010).

A compliance strategy, or the establishment of a set of rules to guide decision-makers, is used to evaluate how risky an object is and its interaction with other objects (Hopkins, 2011). For assessing compliance, several methodologies have been presented. Barrett and Donald (2003), for example, propose a stochastic dominance analysis to compare probability distributions before establishing a hierarchy; Boucher, Danielsson, Kouontchou, and Maillet (2014) use risk metrics and forecasting to adjust models based on historical performance; and Zanoli, Gambelli, Solfanelli, and Padel (2014) investigate the effects of risk factors on non-compliance in UK agriculture.

Because it permits evaluating whether an item performs according to decision-makers' preferences and overstated risk measures, the compliance approach is more user-friendly for decision-making. The main concept is to divide the risk spectrum into two categories (Hopkins, 2011). As a result, the higher the compliance with a stated risk metric, the closer the decision-makers' preferences are aligned. Scarlat et al. (2012) and Tarantino (2008) examine similar approaches based on important risk indicators.

PROMETHEE and ELECTRE

Another layer of complexity emerges when decision-makers must integrate potentially conflicting decision criteria (quantitative or qualitative, monetary and nonmonetary) into project management, such as legal (taxes, compliance, social responsibility, etc.), environmental (level of pollution, noise, watershed issues, etc.), and economic (level of economic growth, monetary and nonmonetary). Furthermore, the relative significance (RI) or weights of those criteria may differ. The phrases in BP's (2015) sustainability report that businesses "must earn and keep society's support" and "must take action to assist to conserve the environment for future generations" may imply that certain decision-makers value profit over social responsibility or vice versa. As a result, it is critical to factor those variances into the decision-making process (Mun et al., 2017).

To solve this issue, multicriteria analysis (MCA) has emerged as an effective tool for dealing with multi-dimensional problems and obtaining an Aggregate Quality (AQ) that may be used to support a final decision (Bouyssou et al., 2006; Brito et al., 2010). MCA is a set of strategies, techniques, and tools that aid individuals in solving choice issues (description, grouping, ranking, and selection) by considering multiple objectives or criteria at the same time (Roy, 1996; Ghafghazi et al., 2010; Kaya & Kahraman, 2011; Afsordegan et al., 2016).

The authors propose PROMETHEE (Goumas & Lygerou, 2000; Brans & Mareschal, 2005; Behzadian et al., 2010; Tavana et al., 2013) as an appropriate MCA technique. Outranking the connection S is the basis of PROMETHEE techniques. This notion defines whether "the alternative is at least as good as the alternative b," rather than determining whether the relationship between two alternatives a and b is a strong preference ("a P b"), a weak preference ("a Q b"), or indifference ("a | b"; Brans & Mareschal, 2005).

Because of their theoretical and practical merits, PROMETHEE procedures are appropriate. They can, for example, assign an AQ index to each project that maximizes the available information in terms of decision-makers' preferences for the criteria chosen, as well as the intensity of those preferences among alternatives and the nature of each criterion (Bouyssou et al., 2006). Many energy-related studies have used PROMETHEE methods, including sustainable energy planning (Pohekar & Ramachandran, 2004; Cavallaro, 2005); renewable energy alternatives (Georgopoulou et al., 1997); heating system options (Ghafghazi et al., 2010); and oil and gas pipeline planning (Tavana et al., 2013); and oil and gas pipeline planning (Behzadian et al., 2010).

There are other approaches, such as the ELECTRE methodologies (Bouyssou et al., 2006), the Analytical Hierarchy Process (AHP; Desai et al., 2012; Saaty, 2013), MACBETH (Cliville et al., 2007; Costa et al., 2012), and TOPSIS (Kaya & Kahrama, 2011). These alternative approaches, on the other hand, do not clearly describe the aforementioned benefits, and the AQ is harder to read.

Although some studies have attempted to incorporate real options (RO) into MCA (Cavallaro, 2005; Angelou & Economides, 2008; Tolga & Kahraman, 2008; Zandi & Tavana, 2010; Tolga, 2011, 2012), there is little evidence of an integrated RO-MCA methodology for ranking a portfolio of projects in state-owned energy companies that pursue nonfinancial objectives.

According to the author, while RO values and assesses flexibility and uncertainty for PM, MCA allows for the inclusion of additional factors such as GDP and employment in strategic planning criteria to produce an AQ for picking the best projects.

References

Asbeck, E., & Haimes, Y. Y. (1984). The partitioned multiobjective risk method. *Large Scale Systems, 6*(1), 13–38.

Barrett, G. F., & Donald, S. G. (2003). Consistent tests for stochastic dominance. *Econometrica, 71*(1), 71–104.

Behzadian, M., Kazemzadeh, R. B., Albadvi, A., & Aghdasi, M. (2010). PROMETHEE: A comprehensive literature review on methodologies and applications. *European Journal of Operational Research, 200*(1), 198–215.

Betts, K. D., & Jaep, K. R. (2017). The dawn of fully automated contract drafting: Machine learning breathes new life into decades-old promises. *Duke Law and Technology Review*, *15*(1), 216–233.

Boucher, C. M., Danielsson, J., Kouontchou, P. S., & Maillet, B. B. (2014). Risk models-at-risk. *Journal of Banking and Finance, 44,* 72–92.

Brans, J.-P., & Mareschal, B. (2005). Multicriteria decision aid. The PROMETHEE-GAIA solution in multiple criteria decision analysis: State of the art surveys. *International Series in Operations Research & Management Science, 78*, 163–186.

Brans, J. P., & Vincke, P. H. (1985). A preference ranking organization method: The PROMETHEE method for multiple criteria decision making. *Management Science, 31*(6), 647–656.

Bruggemann, R., Bücherl, C., Pudenz, S., & Steinberg, C. (1999). Application of the concept of partial order on a comparative evaluation of environmental chemicals. *Acta hydrochimica et hydrobiologica, 27,* 170–178.

Cliville, V., Berrah, L., & Mauris, G. (2007). Quantitative expression and aggregation of performance measurements based on the MACBETH multicriteria method. *International Journal of Production Economics, 105*(1), 171–189.

Daley, S. (2019, September 24). *19 examples of artificial intelligence shaking up business as usual*. https://builtin.com/artificial-intelligence/examples-ai-in-industry

Darken, R. (2019, October 21). *Human-machine teaming AI*. Naval Postgraduate School.

Defense Advanced Research Projects Agency. (2019). *AI next campaign*. https://www.darpa.mil/work-with-us/ai-next-campaign

De Loof, K., De Baets, B., & De Meyer, H. (2011). Approximation of average ranks in posets. *MATCH Communications in Mathematical and in Computer Chemistry, 66,* 219–229.

Denning, P. (2019, September). *Harnessing artificial intelligence*. Naval Postgraduate School.

DoD. (2019, February 12). *Summary of the 2018 Department of Defense artificial intelligence strategy: Harnessing AI to advance our security and prosperity*. https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF

Desai, S., Bidanda, B., & Lovell, M. R. (2012). Material and process selection in product design using decision-making technique (AHP). *European Journal of Industrial Engineering, 6*(3), 322–346.

Dorini, G., Kapelan, Z., & Azapagic, A. (2011). Managing uncertainty in multiple-criteria decision-making related to sustainability assessment. *Clean Technologies and Environmental Policy, 13,* 133–139.

Eykholt, K., Evitimov, I., Fernandes, E., Li, B., Rahmati, A., Xia, C., & Song, D. (2018). *Robust physical-world attacks on deep learning visual classification*. Cornell University.

Gardner, H. (1993). *Multiple intelligences*. Basic Books.

Ghafghazi, S., Sowlati, T., Sokhansanj, S., & Melin, S. (2010). A multicriteria approach to evaluate district heating system options. *Applied Energy, 87*(4), 1134–1140.

Golstein, B. (2018, October 10). *A brief taxonomy of AI*. Sharper AI. https://www.sharper.ai/taxonomy-ai/

Greenfield, D. (2019, June 19). *Artificial intelligence in medicine: Applications, implications, and limitations*. Harvard University. http://sitn.hms.harvard.edu/flash/2019/artificial-intelligence-in-medicine-applications-implications-and-limitations/

Goumas, M., & Lygerou, V. (2000). An extension of the PROMETHEE method for decision making in fuzzy environment: Ranking of alternative energy exploitation projects. *European Journal of Operational Research, 123*(3), 606–613.

Gunning, D. (2017, November). *Explainable artificial intelligence*. DARPA. https://www.darpa.mil/attachments/xaiprogramupdate.pdf

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On past, present, and future of AI. *California Management Review*, *61*(4), 5–14.

Haimes, Y. Y. (2009). *Risk modeling, assessment, and management* (3rd ed.). John Wiley & Sons.

Icertis. (2019a). *Icertis customer profile: Mindtree*. https://www.icertis.com/customer/mindtree/

Icertis. (2019b, April 2). *Microsoft streamlined its contract management*. https://www.icertis.com/customers/microsoft-information-exchange-agreements-case-study/

Institute for Robotic Process Automation & Artificial Intelligence. (2019, November 16). *What is robotic process automation?* https://irpaai.com/what-is-robotic-process-automation/

Jorion, P. (2007). *Value at risk: The new benchmark for managing financial risk* (3rd ed.). McGraw-Hill.

Kaya, T., & Kahraman, C. (2011). Multicriteria decision making in energy planning using a modified fuzzy TOPSIS methodology. *Expert Systems with Applications, 38*(6), 6577–6585.

King, A. D. (2019, November 16). *Talk to a transformer*. https://talktotransformer.com/

Knight, W. (2017, April 11). *The dark secret at the heart of AI*. Technology Review. https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/

Lawgeex. (2018, February). *Comparing the performance of artificial intelligence to human lawyers in the review of standard business contracts*. Law.com. https://images.law.com/contrib/content/uploads/documents/397/5408/lawgeex.pdf

Leviathan, Y. (2018, May 8). Google Duplex: An AI system for accomplishing real-world tasks over the phone. *Google AI Blog*. https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance,* 77–91.

Mun, J. & Anderson. M. (2021). Technology trust: System information impact on autonomous systems adoption in high-risk applications. *Defense Acquisition Research Journal, 28*(1), 2–39. https://doi.org/10.22594/10.22594/dau.19-841.28.01

Mun, J. (2015a). *Modeling Risk: Applying Monte Carlo risk simulation, strategic real options, stochastic forecasting, portfolio optimization, data analytics, business intelligence, and decision modeling* (3rd ed.). Thomson-Shore and ROV Press.

Mun, J. (2015b). *Readings in certified quantitative risk management (CQRM)* (3rd ed.). Thomson-Shore and ROV Press.

Mun, J. (2016b). *Real options analysis: Tools and techniques for valuing strategic investments and decisions with integrated risk management and advanced quantitative decision analytics* (3rd ed.). Thomson-Shore and ROV Press.

Mun, J. (2016a). *Real options analysis* (3rd ed.). Thomson-Shore and ROV Press.

Mun, J. (2016c, October). *Empirical cost estimation tool* [Paper presentation]. Naval Acquisition Research Conference, Monterey, CA.

Mun, J. (2019). Empirical cost estimation for U.S. Navy ships. *Universal Journal of Management*, *7*, 152–176.

Mun, J., & Housel, T. (2010). *A primer on applying Monte Carlo simulation, real options analysis, knowledge value-added, forecasting, and portfolio optimization.* Calhoun.

Mun, J., Ford, D., & Housel, T. (2012, October 10). *Naval ship maintenance: An analysis of the Dutch shipbuilding industry using the Knowledge Value Added, Systems Dynamics, and Integrated Risk Management methodologies* (NPS-AM-12-204). https://calhoun.nps.edu/bitstream/handle/10945/33851/NPS-AM-12-204.pdf?sequence=1

Mun, J., George, K., & Ledbetter, E. (2020). *Total ownership with lifecycle cost model under uncertainty for surface ships' electro-optical-infrared-sensors* [Unpublished manuscript].

Mun, J., Hernandez, E., & Rocco, C. (2016). A combined lexicographic average rank approach for evaluating uncertain multi-indicator matrices with risk metrics. In M. Fattore & R. Brüggemann (Eds.), *Partial order concepts in applied sciences*. http://www.springer.com/la/book/9783319454191. eBook ISBN (978-3319454214)

Mun, J., Housel, T., & Wessman, M. D. (2010). PEO-IWS ACB insertion portfolio optimization. *Proceedings of the Seventh Annual Acquisition Research Symposium*, *2* (No. NPS-AM-10-069-VOL-2). Naval Postgraduate School. https://my.nps.edu/documents/105938399/108624025/NPS-AM-10-069.pdf/c71c6830-853a-448b-beac-242bea4cba8b?version=1.0

Nayak, P. (2019, October 25). *Understanding searches better than ever before*. Google. https://blog.google/products/search/search-language-understanding-bert

Oppy, G., & Dowe, D. (2016, February 8). *The Turing test*. In E. N. Zalta (ed.), *Stanford encyclopedia of philosophy.* https://plato.stanford.edu/entries/turing-test/

Parloff, R. (2016, September 28). From 2016: Why deep learning is suddenly changing your life. *Fortune.* https://fortune.com/longform/ai-artificial-intelligence-deep-machine-learning/

Pulido, F. J., Mandow, L., & de la Cruz, J. L. P. (2014). Multiobjective shortest path problems with lexicographic goal-based preferences. *European Journal of Operational Research, 239*(1), 89–101.

Raghaven, S., & Mooney, R. J. (2013). *Online inference-rule learning from natural-language extractions*. The University of Texas.

Restrepo, G., Brüggemann, R., Weckert, M., Gerstmann, S., & Frank, H. (2008). Ranking patterns, an application to refrigerants. *MATCH Communications in Mathematical and in Computer Chemistry, 59,* 555–584.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk, 2,* 21–41.

Roy, B. (1968). Classement et choix en presence de points devue multiples (la methode ELECTRE). *Revue d'Informatique et de recherché opérationelle, 6*(8), 57–75.

Saban, D., & Sethuraman, J. (2014). A note on object allocation under lexicographic preferences. *Journal of Mathematical Economics, 50,* 283–289.

Sammalkorpi, S., & Teppala, J. P. (2019). *AI in procurement*. Sievo Oy.

Scarlat, E., Chirita, N., & Bradea, I. A. (2012). Indicators and metrics used in enterprise risk management (ERM). *Economic Computation and Economic Cybernetics Studies and Research, 46*(4), 5–18.
Shanahan, P. (2018). *DOD cloud strategy*. Department of Defense.
Shaw, M. (2019, October 15). *Why Google is the best search engine (and why businesses should care)*. Tower Marketing. https://www.towermarketing.net/blog/google-best-search-engine/
Sievo. (2019, November 16). *AI in procurement*. https://sievo.com/resources/ai-in-procurement
Tarantino, A. (2008). *Governance, risk, and compliance handbook: Technology, finance, environmental, and international guidance and best practices.* John Wiley & Sons.
Tavana, M., Behzadian, M., Pirdashti, M., & Pirdashti, H. (2013). A PROMETHEE-GDSS for oil and gas pipeline planning in the Caspian Sea basin. *Energy Economics, 36,* 716–728.
Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior.* Princeton University Press.
Yaman, F., Walsh, T. J., Littman, M. L., & Desjardins, M. (2011). Democratic approximation of lexicographic preference models. *Artificial Intelligence, 175*(78), 1290–1307.
Zanoli, R., Gambelli, D., Solfanelli, F., & Padel, S. (2014). Assessing the risk of non-compliance in UK organic agriculture. *British Food Journal, 116*(8), 1369–1382.
Zarkadakis, G. (2019, September 11). *The rise of the conscious machines: How far should we take AI?* Science Focus. https://www.sciencefocus.com/future-technology/the-rise-of-the-conscious-machines-how-far-should-we-take-ai/

Acquisition Research Program
Department of Defense Management
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net