WHITE PAPER

The Real Facts About Artificial Intelligence

By Thomas Casey, Consulting Director, Analytical Solutions

12.18 / EB10271 / TERADATA CONSULTING/AUTOMATING INTELLIGENCE SERIES/WHITE PAPER

teradata.

Table of Contents

Second in a series of white papers on the topics of Automating Intelligence, this document focuses on how artificial intelligence is a driving force behind analytical automation. Other topics in this series include:

1. AUTOMATING INTELLIGENCE: Industrializing Analytics at Enterprise Scale
2. AUTOMATING INTELLIGENCE: The Real Facts About Artificial Intelligence
3. AUTOMATING INTELLIGENCE: Taking a Multi-Dimensional Approach to Personalization
4. AUTOMATING INTELLIGENCE: Recommended Strategies for Applying Recommendations
5. AUTOMATING INTELLIGENCE: A Pragmatic Approach to Data Management Solutions for Analytics
6. AUTOMATING INTELLIGENCE: Quality Data, the Cornerstone of Effective Analytics
7. AUTOMATING INTELLIGENCE: Developing an Organizational Culture to Maximize your Analytical Initiatives
8. AUTOMATING INTELLIGENCE: The Business-Led Strategy and Framework to Operationalize your Analytical Initiatives

teradata.

Artificial Intelligence...A Driving Force Behind Analytical Automation

Artificial Intelligence, or AI, is a hot topic. There are various articles and stories touting the remarkable opportunities available to companies that leverage AI. Pundits point out that AI may be the most disruptive technological innovation since the steam engine and has the potential to drastically change the way we work, interact, and even live. You have heard from your CMO that you need an AI strategy and you are talking to a myriad of vendors that are telling you why their AI solution is the right one for you. Before that happens, you first need to have a common understanding of what AI is, what it is not, and how it can fit within your overall analytical strategy. Within the Automating Intelligence framework[1] , AI is a key part of the overall analytical strategy but is not an end result by itself. The focus for all analytical initiatives (whether using AI or not) is to automate business processes through analytics—not specifically deploying AI techniques in and of themselves. This document presents a strategy for positioning AI in that context.

The operationalizing of analytics and the processes they drive have always been the implicit goal of why we use analytical techniques in the first place. Going forward, whether it is a sophisticated algorithm that is used to recommend offers or ad hoc analysis to determine what campaigns worked best and why, these all need to be part of improving some business process. In other words, analytical activity does not stop at gaining insight or discovering something interesting or generating a model score. Instead, it should be used to initiate or change some actionable event. The rationale for this is based on three current trends that exist today, each of which continues to evolve (see Figure 1). In this document we will delve into the area of AI by understanding how it has evolved, how it is being leveraged today, and how you should think about it as you build an analytically-driven solution within your own organization.

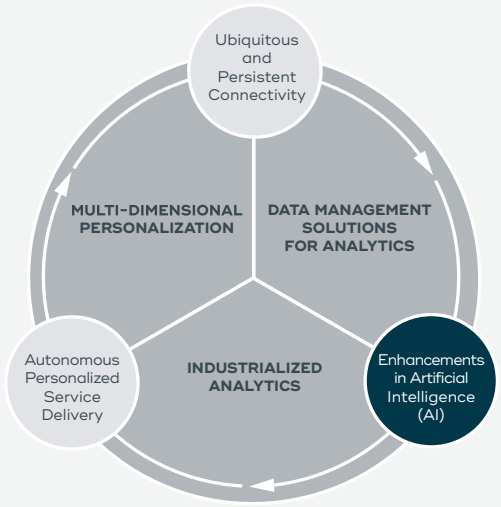


Figure 1. The Drivers Behind Operationalizing Analytics.

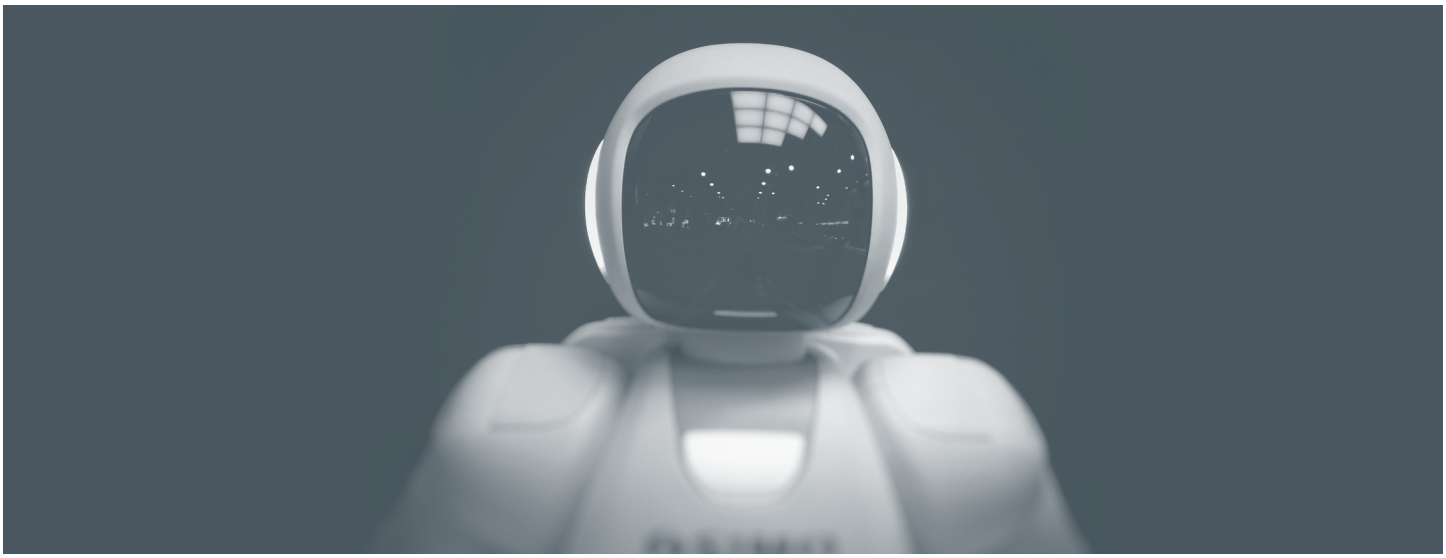
Recent Enhancements in AI

Although many of the current advancements are still overstated, there is a relatively recent convergence of storage technology, processing capability, and revamped analytical methods—leading to a closer realization of AI's initial vision. This evolving capability continues to find new ways to leverage new data types as well as the sheer volume of data being created. Although today's AI still does not replace the need for people in most applications, algorithms can do things that were clumsy or even unthinkable a few years ago. For instance, machines can now recognize handwritten characters about as accurately as a human after seeing just a single example.[2] This is something that was not possible until recently and with increased emphasis on enabling technologies (including aggressive investments into AI technology by countries like China[3]) it is almost assured that these capabilities will continue to evolve. This is why it is considered one of the three major trends driving the Analytics Revolution.

1 Casey, Thomas. Automating Intelligence White Paper Series. Teradata, February 2018, https://www.teradata.com/Resources/Executive-Briefs/Automating-Intelligence.

2 Knight, Will. "This AI Algorithm Learns Simple Tasks as Fast as We Do." MIT Technology Review, December 10, 2015, https://www.technologyreview.com/s/544376/this-ai-algorithm-learns-simple-tasks-as-fast-as-we-do/

3 Xiao, Eva. "The Chinese Government Wants A 100 Billion RMB AI Market by 2018." TechNode. May 27, 2016, https://technode.com/2016/05/27/chinese-goverment-wants-100-billion-level-artificial-intelligence-market-2018/



AI uses a combination of human characteristics and sensory interactions to emulate human thinking and has evolved over several decades. Although AI is a very broad and complex concept with a rich and varied history, many of the advancements have mirrored those of computational power itself. What once was impossible due to computational limitations has suddenly become possible. Current technologies have demonstrated many practical applications for these advancements, but understanding these applications requires some basic understanding of what AI is and how it has evolved.

The Evolution of AI

In the 1950s, several individuals envisioned computers evolving to be able to think as people do, and they termed this field of study, "Artificial Intelligence." Early computers were certainly limited by today's standards and the advancement of any artificial "thinking" was similarly limited. For the most part, practitioners of the AI concept used an ever-increasing set of rules-based algorithms to create an expert system for one action or another. There were certainly gains but learning was rudimentary. After several decades of trying to adapt machines to emulate the thinking processes of a human expert, the limitation still resulted in computers not learning. This led to the next phase of AI, where machines were taught to learn.

As a subset of AI, Machine Learning focuses not on teaching a machine to do something, but rather giving the machine the ability to learn through experience and to improve its "knowledge" over time. Machine Learning showed great promise as we started to see these capabilities manifest themselves in high profile ways such as IBM's Deep Blue™ chess-playing success and later with Watson's Jeopardy™ achievements. In both cases, the machine learned to compete at the highest levels, but the ability to fully realize artificial thinking was still limited to very structured objectives.

For decades, the concept of neural networking existed but was not widely embraced for practical applications. Neural Network modeling attempts to emulate the way the brain works by focusing on a specific outcome and then breaking it into interconnected nodes and assigning weights to the interconnected relationships exhibited between the nodes. The ability to repeatedly do this in order to improve the output results is a powerful analytical method. Early attempts for practical applications, however, were marginalized due to the computational limitations at the time. In recent years, data scientists have leveraged advances in computational power to make these networks enormous and use them to process massive amounts of data. These networks, when applied to the right use case, are very effective. Thus, began the offshoot of AI and Machine Learning, termed "Deep Learning."

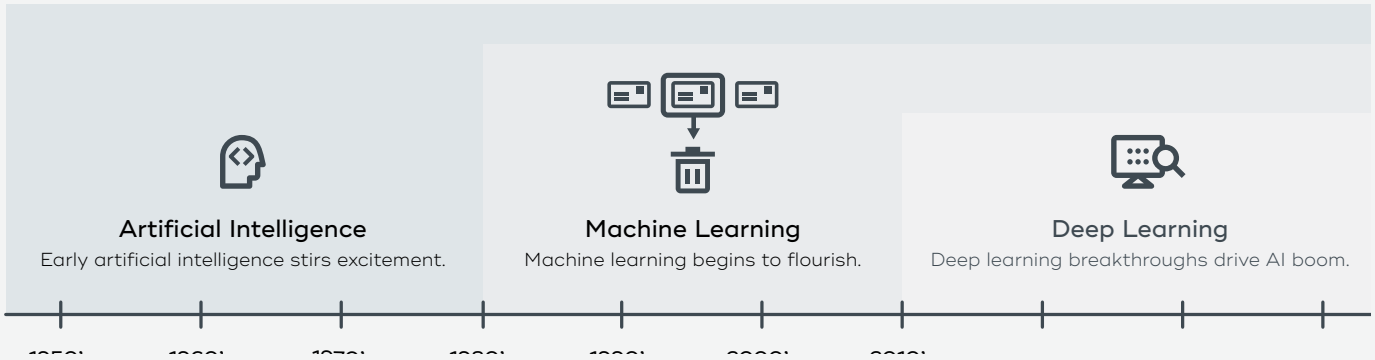


Figure 2. A Timeline for Artificial Intelligence

An example of the AI timeline and the subsets of Machine Learning and Deep Learning is accurately demonstrated in Figure 2.

What this figure shows is that everything we discussed falls into the category of AI. Machine Learning is a subset of AI where algorithms are "trained" to get better autonomously by repeatedly running scenarios. Subsequently, Deep Learning is a subset of Machine Learning that takes the statistical approach of Machine Learning and applied long-standing concepts around Neural Network modeling to drive new and more complex applications for AI. These are not necessarily mutually exclusive or superseding techniques but can support and build upon each other. The real questions for most people are—where are we today with AI, and how much of it is hype versus reality?

How to Classify AI Today

There are many terms used to categorize Artificial Intelligence capabilities. Soft AI vs. Hard AI, Weak AI vs. Strong AI, and Narrow AI vs. Artificial General Intelligence (AGI) are contrasting categories that have been used with various levels of popularity, with the latter probably representing the most commonly used designation today. Regardless of the terms, they all provide a clear distinction within AI capabilities which provides for a clear delineation between when an AI implementation moves from performing a prescribed function towards having the ability to derive outcomes using environmental context (i.e. truly emulate human intelligence).

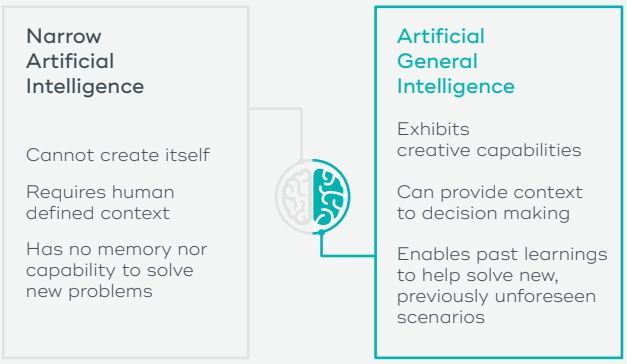


Figure 3. Narrow AI vs. Artificial General Intelligence.

Most pundits generally agree that true AGI is not currently available and there is no real consensus as to how close we are to having this capability. Thus, the scope of the Automating Intelligence framework focuses exclusively on Narrow AI concepts. Despite the terms typically used to describe this discipline (soft, weak, narrow), nearly all AI applications today are clearly in this category and many are very advanced in their capabilities. Whether it is something like facial recognition applications, online language translators, or even self-driving cars, these are all "Narrow AI" applications as they still exhibit the traits identified above. Nevertheless, as evident from these applications' capabilities, "narrow" should not be synonymous with "simple" with respect to Artificial Intelligence deployments.

teradata.

Narrowing Down the Use Cases for "Narrow AI"

Unlike AGI or what some people refer to as "Super AI", Narrow AI is still non-sentient in that it relates to computer models/applications/systems which are not self-aware of what they are designed to do, nor how they can address a problem based on unrelated past experiences. It is the level currently achieved by existing technologies, however, and can still result in some spectacular outcomes. The following is how Narrow AI use cases are categorized within the Automating Intelligence framework.

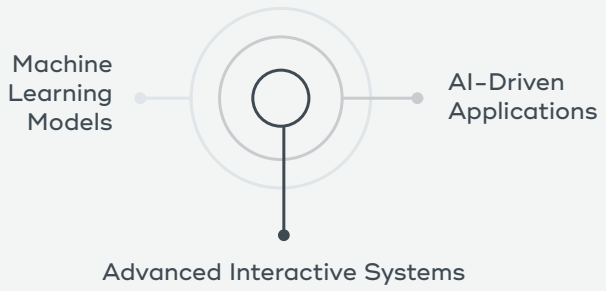


Figure 4. Narrow AI Categories.

- **Machine Learning Models**—individual algorithms meant to solve a specific use case independently
- **AI-Driven Applications**—applications that apply multiple AI algorithms and/or expert system rules to react to a specific set of environmental criteria
- **Advanced Interactive Systems**—multiple applications and models brought together to perform a complex service or interaction autonomously or in a new autonomous fashion

Within this categorization there is a great disparity, from very simple to very complex, in terms of what each is designed to do. There is also a hierarchy in that models are combined to develop AI-Driven Applications. Subsequently, models and applications can be brought together to develop advanced interactive systems that can address several different use cases but are still designed for a specific purpose and cannot evolve beyond that. Each of these is explained in more detail with specific examples in the following sections.

Machine Learning Models

As noted in the Evolution of AI section, pundits have typically used the term "machine learning" to define the discipline of self-learning models with a specific use case using non-linear mathematical methods (e.g. neural networks) defined as "deep learning." As this definition continues to evolve, this binary definition will continue to evolve as well since some linear algorithms are very complex and some multi-dimensional algorithms are relatively simple; therefore, machine learning models may be better categorized as falling across a complexity continuum. One available term to denote this range of sophistication is shallow learning and deep learning. In this category of Narrow AI, we have a variety of general linear predictive models that fall across this continuum. There are also very complex and evolving use cases such as Natural Language Processing and Facial Recognition that have an individual focus but are self-learning in nature. The way they would be represented generally across this range is as follows.

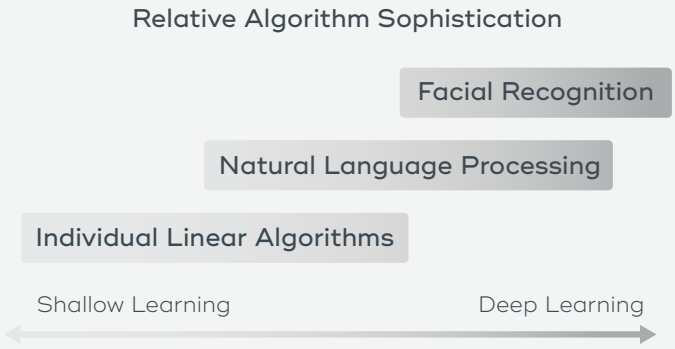


Figure 5. Shallow Learning vs. Deep Learning.

Individual Linear Algorithms

Most Machine Learning (ML) deployments involve individual predictive models that support a variety of business use cases across a broad spectrum of complexity using linear-based algorithms.

Sometimes these models serve as input to more complex predictive solutions or are combined to build out full AI-Driven Applications(see AI-Driven Applications). There are generally four types of predictive modeling techniques (which include multiple algorithm types for each), as shown in Table 1.

Algorithm Type	Description
Classification	When the output variable is a category, such as "red" or "blue" or "disease" and "no disease"
Regression	Output variable is a real value, such as {dollars} or {weight} or some value along a relative scale (e.g. 0...1)
Clustering	Where you want to discover the inherent groupings in the data, such as grouping customers by purchasing behavior
Association	Where you want to discover rules that describe large portions of your data, such as people that buy X also tend to buy Y

Table 1. Model Categorization. Adapted from "Supervised and Unsupervised Machine Learning Algorithms," by J. Brownlee, March 16, 2016, Machine Learning Mastery. Retrieved October 22, 2018, from https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/

For these categories the algorithms vary, and when in practice you can use multiple methods before determining the correct one. Denoting these as "linear" algorithms suggests that this grouping, at least for this category of ML algorithms, is reduced to non-neural network type of algorithms. Two popular examples of solutions using these "deeper" learning capabilities include Natural Language Processing (NLP) and Facial Recognition.

Natural Language Processing

NLP is an evolving capability that uses multiple AI techniques to emulate the human interactive process. Decisions are made nearly instantaneously but what goes into the decisions includes the ability to interpret situation, tone, predefined circumstantial rules, as well as experience (provided the data is available to support this). Using NLP, one can create additional value and efficiency by incorporating one or more of the following techniques.[4]

- **Automatic summarization**—This is the process of creating a short summary of a longer piece of text that captures the most relevant information. Think of the abstracts or executive summaries found at the beginnings of research papers and longer reports. This can be achieved by extracting key sentences and combining them into a concise paragraph, or by generating an original summary from keywords and phrases.
- **Natural Language Generation (NLG)**—This combines data analysis and text generation to take data and turn it into language that humans can understand. While it's been used to create jokes and poems, it's also being used to generate news articles based on stock market events and weather reports based on meteorological data.
- **Speech processing**—This is the specific technology that allows virtual assistants to translate verbal commands into discrete actions for the computer to perform. This technology allows Amazon Echo to translate your request to hear some dance music into a specific Pandora search, or Siri to turn your question about local hot spots into a Yelp search for dinner recommendations.
- **Topic segmentation and information retrieval**—These respectively refer to the process of dividing text into meaningful units and identifying meaningful pieces of information based on a search query. You leverage this technology every time you execute a Google search. Taken together, these two techniques are also being used by several legal tech companies to create searchable databases of legal opinions, allowing lawyers to more efficiently find relevant case law without having to scour briefs for hours on end.

4 Keenan, Tyler. "Natural Language Processing: Turning Words Into Data." B2C. August 16, 2016, https://www.business2community.com/brandviews/upwork/natural-language-processing-turning-words-data-01627884

teradata.

- **Biomedical text mining**—This is a subset of text mining used by biomedical researchers to glean insights from massive databases of specialized research. Some of its applications include identifying relationships between different proteins and genes, as well as assisting in the creation of new hypotheses.
- **Sentiment analysis**—This is routinely used by social analytics companies to put numbers behind the feelings expressed on social media or the web to generate actionable insights.
  Marketers use sentiment analysis to inform brand strategies while customer service and product departments can use it to identify bugs, product enhancements, and possible new features.

Multiple techniques can be applied to fit a solution and many of these use cases have evolved to using very complex multi-dimensional algorithms (e.g. deep learning techniques) which allow the algorithms to continue to improve.

Facial Recognition Models

Another very popular use case category that performs what is a simple task for people but uses complex mathematics to perform is our third group of self-learning models, facial recognition.[5] Facial recognition is a method by which software can identify someone or something based on a digital image. Today these techniques work quite well as evident on platforms like Facebook and devices like Apple's iPhone. The methods by which these techniques work is essentially a four-step process.

- **Identifying**—The algorithm must first look at an image and identify each individual face real time. This might seem easy to do, but the technology has only been available since 2016 and now is found on almost every camera, smartphone, and social media site. The typical method breaks down the pixels into gradients and then grouping those into small squares which act as a simple representation of a face and compare that to a known pattern for faces in general.
- **Posing**—This step deals with the issue that the same face will undoubtedly be positioned differently in different images, yet still (obviously) be the same face. A common method to address this challenge is to identify key aspects of any face (rather than all pixels) then reposition these to provide a standard positioning or pose based on these facial landmarks.[6]
- **Encoding**—This step identifies the key components of a face that are most relevant in identification and uses these to classify like objects. Rather than trying to compare every unknown face to every known face (too much processing involved) or defining the ideal measurements up front (too restrictive), most applications will give the algorithm the task in defining the most ideal measurements and then applying them. In this example, if a person were to look at the measurements they would see a series of numbers which would mean nothing to them. Nevertheless, the numbers and the measurements that are behind the numbers are the key to enabling this algorithm.
- **Assigning**—In this step we take the new image of an unknown person that we have encoded, find a match that is known, and apply the label of the latter to the former. This is a relatively efficient and accurate process since we have broken down images into their most important characteristics in step 3.

AI-Driven Applications

AI-Driven Applications are analytical applications that use a variety of self-learning, static modeling, and/or expert systems to address a specific purpose by providing the capability to perform a specific multi-dimensional objective. It varies from a self-learning model in that these use cases tend to incorporate multiple analytical methods to reach a conclusion.

5 Geitgey, Adam. "Machine Learning is Fun! Part 4: Modern Face Recognition with Deep Learning," Medium. July 24, 2016, https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

6 Kazemi, Vahid and Sullivan, Josephine. "One Millisecond Face Alignment with an Ensemble of Regression Trees," http://www.csc.kth.se/~vahidk/papers/KazemiCVPR14.pdf

teradata.

Game	Milestone	Date	AI Method
Chess	Deep Blue defeats reigning world chess champion, Gary Kasparov	May 1997	• Expert Rules • Simulations (possible moves and outcomes)[7]
Jeopardy	Watson defeats the two most successful contestants on Jeopardy	Feb 2011	• Natural Language Processing (DeepQA) • Knowledge Base retrieval[8]
Go	AlphaGo defeats world #1 player	May 2017	• Tree Search • Neural Network[9]

Table 2. AI Gaming Agents.

Gaming Agents

A common type of an AI deployment that is high profile is one where a virtual player can learn and eventually master various games, even against world-class human players. There are high profiles examples such as shown in Table 2.

These and other examples have been used for experimental purposes without tremendous initial business benefit because they deal with scenarios that are completely deterministic. Nevertheless, they have spawned new innovations where these innovations have been applied to other non-gaming areas. Additionally, some recent examples of AI in gaming allow these agents to break free from learning a specific game objective and instead learn different games based on rewards (see Transfer Learning in a subsequent section). This gets around one of the major AI challenges today, being able to build upon an initial skill to apply it to an entirely new set of circumstances.[10]

Recommendation Engines

There are several solutions out there that call themselves "Recommendation Engines" but what they almost all have in common is that they leverage multiple analytical methods/algorithms to provide a series of recommended offers, actions, or events. Typically no single method is used to develop recommendations as it usually involves multiple approaches combined to generate the recommendation. In the AUTOMATING INTELLIGENCE: Industrializing Analytics at Enterprise Scale white paper, six different methods for making recommendations were outlined: Attribute Matching, Popularity, Item Association, Customer Profiling, Concept Mapping, and Collaborative Filtering. These are general techniques and typically there will be a multiple of these or derivations of them used to generate a series of autonomous recommendations. From Google's search engine, to Amazon's shopping suggestions, to Netflix's movie recommendations, they all use a variety of inputs to produce analytically driven recommendations. Many of these are enhanced by improving one or more of the analytics through machine learning. Going forward, more innovative solutions are taking a page from the gaming world and using reinforcement learning techniques. Additional information on recommendations can be found in a subsequent white paper in this series .[11]

7 Greenemeier, Larry. "20 Years after Deep Blue: How AI Has Advanced Since Conquering Chess." Scientific American. June 2, 2017, https://www.scientificamerican.com/article/20-years-after-deep-blue-how-ai-has-advanced-since-conquering-chess/

8 Best, Jo. "IBM Watson: The inside story of how the Jeopardy-winning supercomputer was born, and what it wants to do next." TechRepublic. https://www.techrepublic.com/article/ibm-watson-the-inside-story-of-how-the-jeopardy-winning-supercomputer-was-born-and-what-it-wants-to-do-next/

9 https://en.wikipedia.org/wiki/AlphaGo#Algorithm

10 "Playing Atari with Deep Reinforcement Learning." Retrieved on October 22, 2018, https://www.cs.toronto.edu/~vmnih/docs/dqn.pdf

11 Casey, Thomas, "AUTOMATING INTELLIGENCE—A Recommended Strategy for Recommendations."Teradata.

teradata.

Intelligent Appliances

Ubiquitous access to broadband internet, additional devices with Wi-Fi capabilities, the growth in smartphone penetration worldwide, and the overall decreasing costs of each in relation to performance has created an environment where nearly any electronic device can be connected to the Internet. In the future, one should assume that anything electronic will be connected one way or another. These electronic devices can vary from the relatively simple (e.g., kitchen appliances, home thermostats, and personal devices) to extraordinarily complex (jet engine sensors). Conceptually, these are all electronic sensors and will not just provide instant feedback on what is happening with respect to that device but, by being connected, can understand or even talk to other connected devices. These connected appliances are typically categorized as the Internet of Things or IoT. As an AI-Driven Application, these IoT or Intelligent Appliances make up an advanced use case for Reactive Applications as they pull together multiple algorithms to deliver intelligent services as determined by the scope of the appliance. The opportunity to understand and ride the wave that is IoT and their integration using AI methods is enormous. McKinsey estimates the potential economic impact of the IoT at up to 11.1 trillion dollars per annum by 2025. By comparison, current global GDP is estimated to be 107 trillion dollars in Purchasing Power Parity terms.[12]

Advanced Interactive Systems

Interactive Systems are analytical applications that combine multiple reactive applications to build out a complex method to address a specific set of interactive use cases. These involve more than one specific set of learning actions as the interactions they support are more diverse and require additional analytical scope. That said, these systems still are not context-aware so are still classified as "Narrow"; however, the AI applications in this category garner most of the attention with respect to high-profile AI deployments. The following outlines some specific examples that fall into this group.

Digital Assistants—the Secretary of the 21st Century

A popular set of interactive AI-driven solutions can be categorized as digital assistants. Some assistants are strictly textual in nature (i.e. they focus on text-based interactions). Others engage using primarily spoken interactions. There are many niche players in this space that leverage a combination of capabilities yet almost all of them leverage some form of NLP capability (text and/or voice). There are four major players in this space that have created virtual personae to personify their digital assistants.

Company	Digital Entity	Platform
Apple	Siri	iOS
Microsoft	Cortana	Windows
Amazon	Alexa	FireOS
Google	Google Assistant	iOS, Android

Table 3. Top Digital Assistants.

These four companies are classified as Interaction Systems rather than AI-Driven Applications because they include multiple applications and techniques combined to enable a complex intelligent solution. The way these systems work is similar as they typically have at least three components in common.[13]

1. **Natural Language Processing**—Whether text-related interactions (Facebook's M) or primarily a verbal response-based system (Alexa, Cortana, Siri) the first component of a Digital Assistant is to interpret what you communicated and respond in a way that makes sense. The ability to interpret this level of communication is one of the primary determinants in how well a digital assistant is perceived.

12 McKinsey Global Institute, June 2015.

13 Assefi, Mehdi, Liu, Guangchi, Wittie, Mike P., Clemente, Izurieta. "An Experimental Evaluation of Apple Siri and Google Speech Recognition." Department of Computer Science, Montana State University

teradata.

2. **Backend Knowledge Base**—Typically there is some programming interface between the communication channel of the digital assistant (Smartphone, IoT device, messenger app) and a staged set of predefined "answers" or directions to drive subsequent actions. The NLP agent interacts by trying to interpret what is requested and will look to the knowledge base (typically external to the NLP device) to determine the appropriate response.
3. **External Application Integration**—Here is where digital assistants interact with other applications (either through hardware integration (e.g., Siri/iPhone, Alexa/Amazon Echo) or by establishing permission (i.e., Google Assistant gaining access to your calendar). Using the recommendations from the knowledge base, the agent will interact with the external application, if required, to fulfill the user's request.

In these cases, the capabilities of all three can be improved over time as the ability to improve the NLP capabilities and the decisions which they drive will continue to use its interactive history. Much like a human, past experiences will help better align perceived requests with correct actions.

Reinventing Retail—Amazon Go

As Amazon does when you visit its website, the Amazon Go shopping experience is tracking all your shopping behaviors. This will inform Amazon about how customers traverse the store, the exact placement of products that drive the most sales, and the relationship between the buying experience and actual outcomes. Using analytics, Amazon will be able to create customized on-demand discounts related to your current or prior buying behavior. Additionally, some aspects of this experience will allow users to experience the products in person but still buy and deliver online, or vice versa. This melding of channels will eliminate one of the most complex aspects of understanding customer buying decisions across multiple channels.

One of the biggest aspects of this model, however, is the elimination of direct human involvement in completing the transaction. Moving to a self-service model, all activities will be analytically driven over time. Customers enter the store, navigate based on their individual shopping objectives (ideally influenced by Amazon through product placement and strategic discounts/offers), and leave when their shopping is complete (with the transaction automatically completing). The presumption is that the experience will not only be consistent but will improve for the consumer over time, and none of it demonstrably influenced by a person.

The Fast Food Industry

Andy Puzder, CEO of fast food chains, Carl's Jr. and Hardee's, told Business Insider that, unlike human workers, robots are "always polite, they always upsell, they never take a vacation, they never show up late, there's never a slip-and-fall, or an age, sex, or race discrimination case." What this says is that there is a tremendous upside to automating the fast food industry.

At the 2016 NRA Show, robots were one of the hottest topics in restaurant technology. For example, Suzumo International has discovered a solution to bring sushi to the masses quickly and efficiently. The latest "Sushi Robot" can make up to 4,000 pieces of sushi an hour or a complete roll of sushi every 12 seconds. This can help all-you-can-eat buffets as well as sushi restaurants inside sports stadiums, schools, hospitals and more, to produce mass volumes of sushi in a small amount of time.[14]

Cafe X has created an automated barista and a "coffee shop". Cafe X is 100 percent automated from the ordering and payment system to the preparation and delivery of the coffee. This system is by far faster than any current coffee shop experience. Once the amortization of the system has been met, the cost to operate this "coffee shop" is orders of magnitude lower than the 2.5 baristas a single system replaces. Three lines can form simultaneously, and the system can deliver the order in record speed. Quality reviews suggest it meets or exceeds a typical Starbucks similar product. This is a 1-click buying experience that closely aligns the experience we see online or in person with Uber.[15]

14 "From Robots to Drones: Restaurant Tech of the Future," August 10, 2016, GrubHub for Restaurants

15 "Will No-Checkout Stores Like Amazon Go Be Commonplace by 2025?" February 10, 2017, Forbes

Self-Driving Vehicles

Whether to augment humans with a nonhuman co-pilot, revolutionize mobility services, or reduce the need for sprawling parking lots within cities, self-driving cars have the potential to do amazing things. Driving is complicated and drivers need to anticipate several unexpected circumstances. Ice, closed roads, and a child running across the street are examples of scenarios that you cannot completely account for and anticipate in coding rules. Therein lies the value of deep learning analytical algorithms; they can learn, adapt, and improve. No longer a thing of science fiction, fully autonomous cars are projected to be on the road by 2019 and could number 10 million by 2020.[16] The potential savings to the freight transportation industry is estimated to be $168 billion annually. The savings are expected to come from labor ($70 billion), fuel efficiency ($35 billion), productivity ($27 billion) and accidents ($36 billion) before including any estimates from non-truck freight modes like air and rail.[17]

What Constitutes a Self-Learning Model?

In both the Advanced Analytics and Real-Time Decisioning processes of the Automating Intelligence framework, there are examples of Machine and Deep Learning AI techniques that can be applied. For both, the "learning" aspect describes an algorithm that can independently refine itself and continue to get better at predicting whatever it was initially modeled to do. This does not, however, mean that a model just creates itself and continues to evolve independently ad-infinitum. To better understand this, one can look at a generic process for model development.

The Standard Modeling Process

There are a many ways people define this process with a variable number of steps. The model used in this section is purposely simplified to focus the conversation around contrasting self-learning with more traditional training methods (by limiting the number of steps).

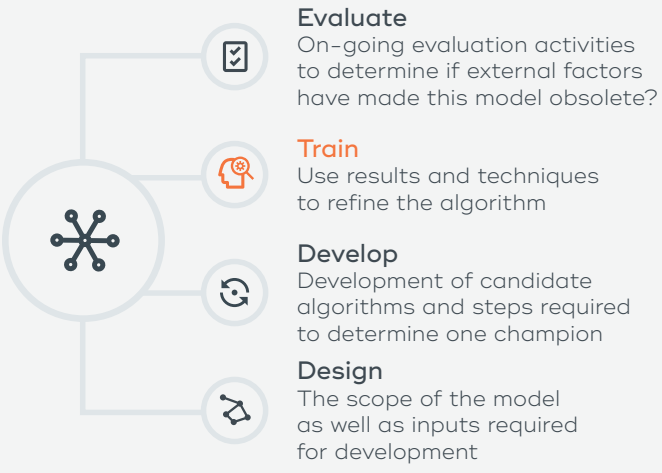


Figure 6. Traditional Statistical Modeling Process.

Using that as a backdrop, we start our process with the design of the model based upon some preliminary definition of the business need or optimization. Typically, some initial data mining activity defined the initial analytical design, including defining the necessary data source scope and creation of descriptive variables used as input to the actual development.

Next, the data scientist develops many candidate algorithms using one or more modeling techniques. Depending on the technique, a series of tests for each candidate happens with an inevitable champion chosen upon completion of this step. This champion is the model that will be deployed.

Prior to initial deployment, ongoing refinement happens where new data sources, calculation "tweaks", and additional variables may be incorporated to improve the model's predictability. The timing of this effort is ongoing and will continue if the model is viable in showing improvement in supporting the business use case.

16 BI Intelligence, "10 million self-driving cars will be on the road by 2020." Business Insider. June 15, 2016, http://www.businessinsider.com/report-10-million-self-driving-cars-will-be-on-the-road-by-2020-2015-5-6

17 "Autonomous Cars, Self-Driving the New Auto Industry Paradigm." November 2013, Morgan Stanley Blue Paper.

teradata.

Finally, there is a step whereby the model is evaluated to determine its ongoing viability. External factors can make it such that a model can no longer be refined as its original scope and definition have become obsolete based upon the new environmental conditions. This activity can be a periodic review or it can be triggered from a specific business event.

Where Is The "Self-Learning"?

All the steps outlined above are done by humans in a traditional set of statistical activities and this has been the case for decades. Even in the world of self-learning AI algorithms, three of the four steps are still done by humans. It is just the third step (Training) that is changed to a self-learning step whereby the algorithm uses data and other inputs to refine itself and essentially make itself better at performing the initial task it was designed to perform.

The point is that the model takes over responsibility to refine itself or "learn" but it still does not design, develop, or evaluate itself. These are all human-initiated activities and continue to be so for most analytical efforts. Much like the data mining process, it is still people that need to ascribe context, viability, and value to a model. People also need to determine when a model is no longer the best method to meet a certain goal by constant and ongoing evaluation. Self-learning models provide a great opportunity to improve analytics in an autonomous way; however, users of this technology should not expect that this in any way replaces the need for human involvement in the overall statistical modeling process. Additionally, it must be understood that once a model is deployed, it has essentially stopped "learning" and any changes to the algorithm require a person to initiate a new process.

Learning Methods

As noted above, the modeling process is somewhat unchanged in the world of Machine Learning other than at the training step. In this step, there are methods whereby a model or system has an automated method for running various scenarios to improve its own algorithm. There are various training methods available to support the self-learning activities used to train machine learning models. The various techniques include:

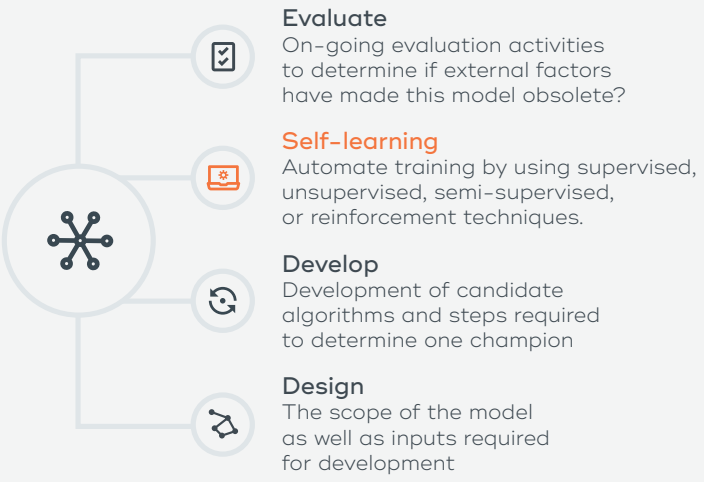


Figure 7. Statistical Modeling and Self Learning.

- **Supervised Learning**—the most popular machine learning technique whereby a model "learns" using a desired or "correct" output. For instance, if a model was trying to develop an algorithm to forecast something, it knows what it is looking to forecast from previous data and can continue to generate iterative changes to the underlying algorithm as it looks to develop a better and better version of itself,
- **Unsupervised Learning**—as the name suggests, unsupervised learning relates to scenarios where there isn't a "correct" answer. Instead, the algorithm can continue to refine itself based on the validity of the outcomes and not by comparing them to a known result, and
- **Semi-Supervised Learning**—this is a hybrid of the supervised method where one may be dealing with algorithms and scenarios which making fully supervised learning impractical. This technique can use supervised methods in a data set that is partially labeled but then apply unsupervised methods to predict the outcomes for other data elements that are not explicitly labeled but show some level of similarity with a known data point. This is a specialized learning use case.

teradata.

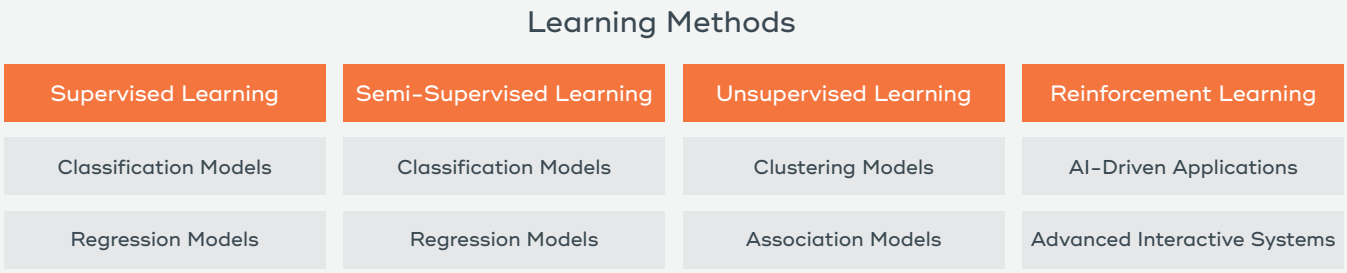


Figure 8. Machine Learning Methods.

In addition to the options above, there is also a method of self-learning referred to as "Reinforcement Learning". In this method, a more holistic approach is taken to improve an AI-driven solution to improve over time based on maximizing a particular outcome or reward. In a way, this is supervised in that people define the reward, but it is beyond simply a model improving its accuracy because the cumulative effects of maximizing the reward can drive self-defined changes. This has been particularly popular in AI-Gaming Deployments and in advanced systems like autonomous cars. This method, however, is starting to be used in broader deployments (e.g. Recommendation Engines) as its outcomes can be based on very specific real-world objectives. In terms of self-learning, then, one is left with the following alignment for each analytical category.

These learning capabilities can enable an algorithm, application, or system to improve itself. One of the key reasons we are still involved in Narrow AI, however, is because these solutions are not self-aware and cannot recreate themselves to serve a purpose outside of what they were designed to do. Therefore, a key human-initiated step still exists which is to evaluate when an AI component is no longer relevant or can be replaced with an entirely new method.

Evaluation Considerations For Self-Learning Analytics

As previously stated various times, algorithms cannot determine business priority, scope of a business problem, and they certainly cannot determine when they have outlived their respective usefulness.

Just because a model has optimized its algorithm based on one of the learning methods above, it still takes a person to constantly evaluate if there are newer and better ways to address a business challenge or maximize an opportunity. This step, identified as "Evaluate" in the process above, is where humans provide that oversight in the world of machine learning. Just as we identified three category levels in this framework for organizing Narrow AI use cases, evaluation activities take place across these same levels but with a different scope of effort for each. In the cases where multiple models are incorporated to build out an application, evaluation methods will happen at multiple levels (e.g., evaluating the models individually, modeling the collective outcomes of a group of models, and assessing a group of outcomes used to support a system or program). The following outlines the three levels of evaluation per the Automating Intelligence framework.

Model Level Evaluation

Testing for a specific model may require multiple touch points—especially if you use that model for multiple applications. In cases such as these, you should have a mechanism where you can apply a group or parent model to a specific set of use cases but provide for the ability to assign specific models or versions of models to specific individuals, ideally via a random assignment. This may require a significant duration of time to allow for results to be adequately captured but throughout this duration the actual assignments should not change.

Application Level Evaluation

An application contains one or more models to evaluate. At this level, there are multiple inputs that drive a specific set of output and to evaluate every aspect of incorporated models and rules are probably not practical at this point. Because of such, there should be a means to measure the overall effectiveness of the application by comparing it with other approaches or strategies to drive the action. In some cases, there may be two totally different approaches to solving a problem with different sets of analytics leveraged. The previous level of testing (Model) evaluates the algorithm, while this level evaluates how well a collection of algorithms perform in achieving a specific objective.

System Level Evaluation

This is similar to Application Level Evaluation although the scope is broader and more complex. In this case, we are still trying to evaluate based on maximizing an outcome, but we do this across multiple models and applications as an integrated system or solution. This level of evaluation is more strategic for business evaluations as it could be used to evaluate a strategic program and all the underlying strategies of that program. The level of testing is also against not only the status quo, but also the competing strategies used to challenge how business operates today.

These evaluation methods have varying purposes and all will be used by different people, in different contexts, within different levels of the organization. Regardless of how it is implemented or applied, the key component is to realize the human-based evaluation is still a key component to any AI-related deployment.

Broadening The Definition of Narrow AI

Looking at the situation today where Artificial General Intelligence (AGI) is still not viable, there are nevertheless expanding capabilities which seem to act more and more in the vein of exhibiting real intelligence. This is causing an ever-increasing expansion of the Narrow AI definition, and self-learning is a leading component in how this is achieved.

In looking at the various methods available, supervised learning is more about automating a step in the modeling process that has been traditionally performed by people(but now can be done in a more automated way). This does not seem to be a method (especially when using shallow learning-based models) that exhibits real intelligence in any sense, but it still has relevance as part of an automated analytical framework.

There are ongoing discussions as to whether an expansion of unsupervised learning will enable the next leap forward in moving towards real AGI. Although this seems somewhat logical, it still does not address what may be the most glaring omission in most AI solutions today as it pertains to AGI—the ability to transfer knowledge learned from one task and applying it to another completely different task. There is a relatively new discipline that tries to address this, which is typically referred to as "Transfer Learning".

Transfer Learning techniques do not supplant other learning methods; those methods are in place to learn a specific task and that action is still relevant. Instead, the idea of Transfer Learning is to take whatever is "learned" in one area (no matter how that learning took place) and give the algorithm the ability to apply that learning to a task with an entirely different scope. This happens naturally and without consciously thinking about it for people, but for analytics this frontier is still uncultivated. There are a few interesting things being done under this category around gaming. For instance, the Google company, DeepMind, built a new AI system called IMPALA that simultaneously performs multiple tasks—in this case, playing 57 Atari games—and attempts to share learning between them. It showed signs of transferring what was learned from one game to another by becoming 10 times more data-efficient than a similar AI and achieving double the final score.[18] Nevertheless, this and other similar examples are still very rudimentary in that they "transfer" knowledge across very similar scenarios (e.g., video games) which have similar reward systems and operating constructs. Therefore, the pathway to AGI is still very long...and narrow.

18 Snow, Jackie. "DeepMind's latest AI transfers its learning to new tasks." Technology Review. February 12, 2018, https://www.technologyreview.com/the-download/610244/deepminds-latest-ai-transfers-its-learning-to-new-tasks/

teradata.

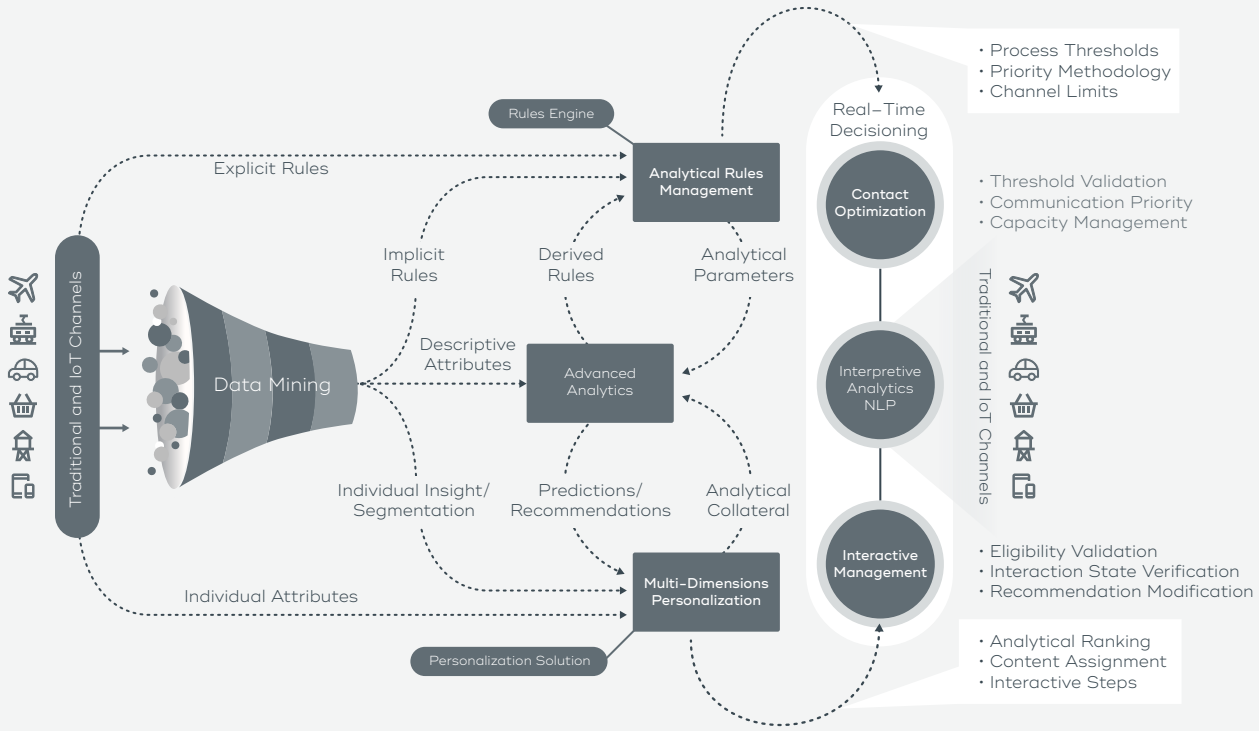


Figure 9. The Analytical Process.

Applying AI to the Automated Intelligence Framework

In the preceding white paper, "AUTOMATED INTELLIGENCE—Industrialized Analytics at Scale", a process for automating analytics based on the way humans go through their own thought processes was defined. In this definition, (see Figure 9) we derived the following four components:

1. **Data Mining**—including all the activities required to identify and define business drivers that could be influenced by some analytical technique,
2. **Analytical Rules Management**—which includes general rules management as well as personalization rules,
3. **Advanced Analytics**—including predictive and prescriptive analytics use cases, and
4. **Real-Time Decisioning**—which includes interactive rules (Contact Optimization and Interactive Management) as well as Interpretive Analytics (including Natural Language Processing).

In the first instance, data mining activities follow a logical progression to determine and quantify opportunities. Machine learning can provide a clear path to model improvement and even expand that using reinforcement strategies; however, as stated earlier in the Narrow AI section, it cannot create itself nor can it determine when it is no longer relevant or optimal to fit a specific purpose. This is a task that is still performed by people and will be relevant for humans to do for the foreseeable future.

Analytical Rules Management—Keeping Algorithms Honest

Many people wonder why they might need any sort of expert input or rules engine to be included in an AI-based solution. If the analytics are meant to be autonomous and learn, why should we involve anyone? Just let the data tell us the answer. Although there might be a time where that is possible, today it is just not practical nor advisable. Take the recent example of Microsoft's chatbot, Tay.

Microsoft launched the AI-driven chatbot in March of 2016. The idea was that Tay, representative of a teenage girl that is active on social media, would engage in casual conversations on Twitter and learn to improve its interactions through this dialog. The issues arose when she started to mimic inappropriate behavior exhibited from her followers and she was taken down a day after going live.[19]

Some would argue that interjecting bias into her algorithms defeats the purpose of machine learning. I disagree, and especially in the context of a business application. We learn from experience, but we also override these learnings when things are inappropriate to say based on the people with whom we interact and/or the circumstance. That was not applied in the case of Tay but it is why Analytical Rules Management is still important to AI solutions in the following three ways.

- **Explicit Rules:** Providing the means to ensuring adherence to explicit instructions based on a known preference (e.g. do not send me email)
- **Implicit Rules:** Determining where things might be appropriate based on some implicit knowledge gained through your actions (e.g., you showed some interest in 55+ communities so I will assume you are a senior or close to being one)
- **Derived Rules:** Using analytical segmentation (e.g., clustering) or derived fuzzy logic to apply some derived set of qualifications to you which can drive or override analytically-driven actions

These rules can be generated via a person stating their preferences, an expert (i.e., human) applying some level of business-driven expertise or known policy, or it can be generated through some predictive means (e.g., forecasting) to determine how best to drive automation. Not everything needs to, or even can be done using algorithms alone…just ask Tay.

Advanced Analytics—Teaching Machines to Learn

As is outlined in the decision-making process of the Automating Intelligence framework, Advanced Analytics is the area that deals specifically with two types of analytics, Predictive Analytics and Prescriptive Analytics. In the overview of machine learning types in the "Narrowing Down Narrow AI" section, the domain of specific machine learning models is most typically aligned to predictive analytics, whereas the AI-Application area is more aligned to Prescriptive Analytics (most notably, Recommendation Engines). Both analytical types certainly call upon machine learning techniques (as outlined in that section) and in the case of Prescriptive Analytics, multiple predictive models may serve as input to the recommendation process.

Predictive Analytics

Predictive Analytics are typically seen as independent models which align to the "shallower" end of the machine learning continuum. This is not always the case, but rather a generalization. As such, all the algorithm categories related to Machine Learning Models (classification, regression, clustering, and association) are viable approaches to use in generating predictive analytics. These models can be used for a variety of use cases and drive actions independently or used in conjunction with other models and/or rules to drive automation (as shown in figure 8). In addition to acting as input to a prescription, predictive analytics can generate fuzzy logic rules that cannot be derived through explicit or simple implicit means (i.e., I am 80 percent sure you fit a profile based on certain actions you have taken and how they apply to known customers).

Prescriptive Analytics

Prescriptive Analytics can consist of a single algorithm but are typically a combination of independent models and rules to drive the Next Best Offer/Action/Event. Given its higher level of complexity and the need to combine multiple techniques, Prescriptive techniques generally fall into a moderate- to deep-learning category as it pertains to machine learning. Given that multiple models can be used to generate a prescriptive output, these models can go through independent self-learning processes to optimize themselves within the scope of their purpose. Additionally, as an application, the inputs collectively can leverage reinforcement learning to drive ongoing improvements, thus providing a two-dimensional aspect to self-learning.

19 "Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter", The Guardian, March 24, 2016, https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter?CMP=twt_a-technology_b-gdntech

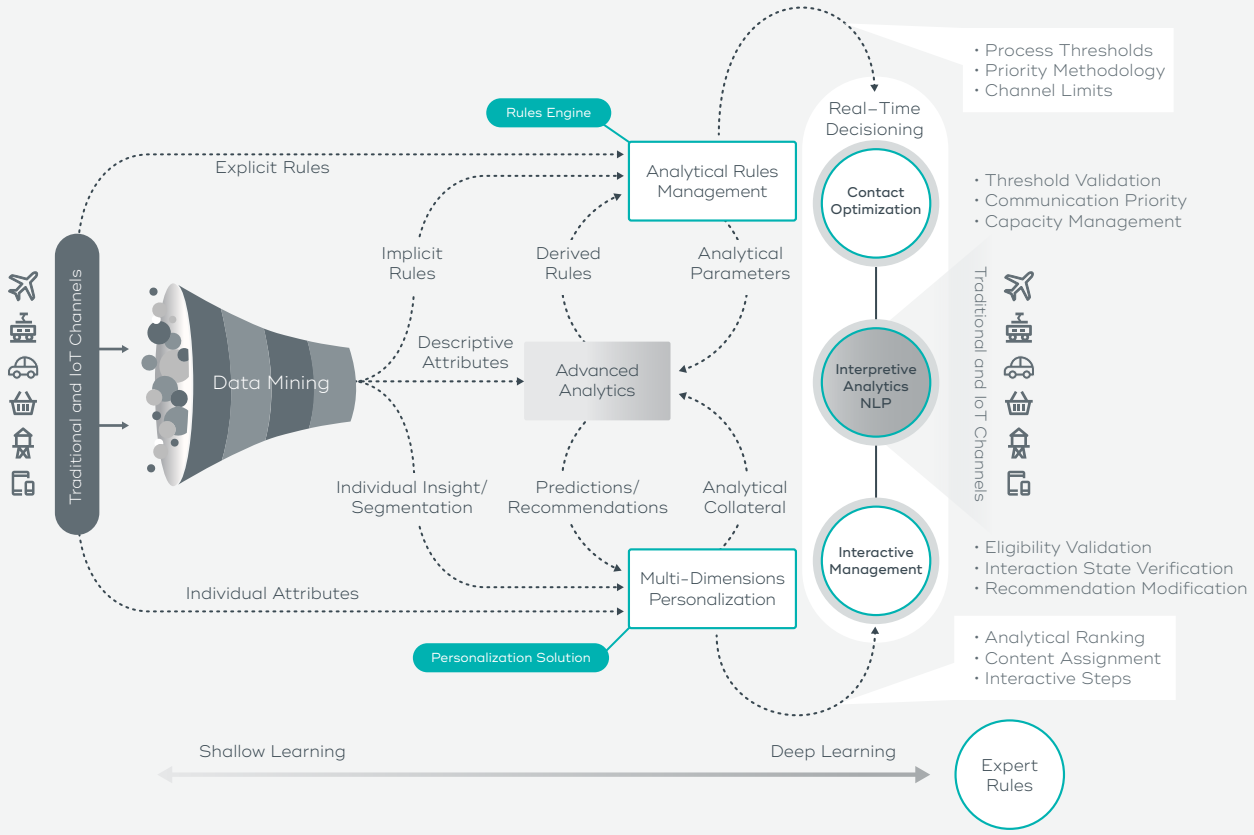


Figure 10. The Analytical Process.

Interpretive Analytics—Human Observations in Real Time

One of the most difficult things to address with analytics is to generate a two-way interaction between a human and a non-human agent (the latter presumably enabled via some AI method). In the section on NLP, many categories in this area of analytics were outlined. How these NLP methods are enacted are conceptually similar in that their analytics attempt to take in text or voice commands, make sense of them, and then look to find some set of personalized recommended next steps based on this interpretation. Interpreting the interaction and engaging, however, takes more than looking through a lot of data to define relevance (see the Microsoft Tay example). Therefore, the framework assumes some level of rules-based expertise is applied outside of the NLP analytic to augment or override what it normally would do.

In terms of AI-categorization of the rules that drive real-time decisioning, they are typically human-defined expert systems with the different rule types previously outlined.

Therefore, they are going to represent relatively shallow machine learning techniques (if they use machine learning at all). Conversely, the actual analytics used to engage (e.g. NLP, Facial Recognition) are typically very complex and today are almost exclusively the domain of some deep learning technique.

Mapping AI To The Analytical Process

In the white paper, "AUTOMATING INTELLIGENCE—Industrializing Analytics at Enterprise Scale," the analytical process was defined as a series of steps based on the way people think. The purpose was to not only automate decision making but to look at analytically driven solutions as an end-to-end process.Taking this process (as shown in Figure 10) there are multiple tasks upon which AI techniques can apply. Using the machine learning continuum as well as expert rules, AI can be overlaid on this process to represent the following:

Many people view AI/Machine Learning as the end of the analytical journey. As is evident from the diagram above, a truly automated analytical process can use various AI techniques at various points in support of the end-to-end solution. In most cases, organizations will need to deploy various algorithms and techniques to achieve the necessary results and, in some instances, it will vary as to which technique will "win" as the overriding decision maker. Regardless, the holistic view represented by Figure 9 shows how you can automate any decision-making process by incorporating AI.

Artificial Intelligence in Action

AI continues to evolve. Although there is significant hype and great desire to incorporate general machine or deep learning techniques in solving many analytical problems, there is still a natural progression in moving from less complex to more complex AI methods as it is typically not productive to jump to the most complex technique without an understanding of where the deficiencies lie between one technique and the next. A great example of this is found in fraud detection for the credit card industry.

The Nilson Report estimates that in 2016 worldwide losses associated with credit card fraud topped $24.71 billion. That represents a 12 percent increase over the previous year.[20] ACI Worldwide estimates that 47 percent of Americans have had their credit cards compromised in the past five years. Fraud is always going to occur and is a loss to the card issuer in fraudulent charges, replacement cards, and customer churn. Improved analytical techniques have constantly evolved to minimize fraud while also minimizing approval timelines. The following shows how the credit card industry used the three methods of AI and both evolutionary and complementary pieces in their fraud detection solution.

Rules-Based Intelligence

Credit decisions for many years were done by humans. Approvals were done by having people look at various inputs and deciding in real time if the charges were legitimate or not. The level of success depended upon the expertise of the approver. Companies used these experts to develop consistent rules and these continued to evolve to accommodate what people could quantify (e.g., you cannot make a purchase in person in two locations 1000 miles away within 30 minutes of each other). These rules continued to evolve and were valuable but would only be reactive, at best, in dealing with fraud detection.

Shallow Learning

Rules were helpful, but institutions needed to be able to better predict fraudulent activity. Moreover, they needed a method that could evolve with those that were committing fraud. As a result, various machine learning techniques were instituted to begin the process of understanding individual transactional patterns. Fraud not detected by rules served as the basis for this analysis and the results did improve over time; however, the complexities of fraud detection meant there were always many false positives, and these had their own challenges in terms of cost and customer satisfaction. In this situation, machine learning was helpful but the problems to solve were not always linear, especially as more transactions were initiated online. That is where deep learning came into play.

Deep Learning

Many companies realized the deficiency of complete reliability on rules and linear analytics to address fraud. PayPal, one of the early adopters of deep learning, saw a major leap forward in using non-linear techniques a couple of years ago. The deep learning algorithms can analyze potentially tens of thousands of latent features (time signals, actors and geographic location are some easy examples) that might make up a fraud, and can even detect "sub modus operandi," or different variants of the same scheme. Once the models detect possible fraud, human "detectives" can get to work assessing what is real, what is not and what to do next.[21]

20 The Nilson Report, October 2016, https://www.nilsonreport.com/upload/content_promo/The_Nilson_Report_10-17-2016.pdf

21 Harris, Derrick. "How PayPal uses deep learning and detective work to fight fraud." GIGAOM. March 6, 2015, https://gigaom.com/2015/03/06/how-paypal-uses-deep-learning-and-detective-work-to-fight-fraud/

Key Takeaways

A key component to the scenario above is that these different approaches can complement each other, augment each other, or compete against each other to identify an analytical champion (and that champion does not mean one will always work best in any situation). These approaches can be supervised or not, but to think they cannot be implemented without first identifying where to focus and how to prioritize efforts is naïve and short-sighted. This is how humans think and how we should look at AI as it tries to emulate the human experience.

A key thing to keep in mind as you look to implement these concepts is that AI is not a solution. Machine learning is not a solution. Both are important analytical concepts, but neither is a solution by itself. They are analytical techniques that may or may not be applicable to fit a specific business need. Instead of looking at specific analytical techniques as a solution, treat analytics as a process enabler that will contribute to an overall solution. That is why this framework focuses on a different type of AI, Automating Intelligence, which is focused on a holistic approach to improving business processes using analytics...and there is nothing artificial about that.

17095 Via Del Campo, San Diego, CA 92127 Teradata.com

teradata.