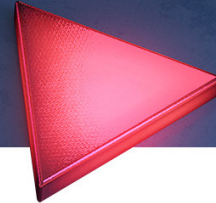




Sonasoftware[®]

The Sonasoftware AI Solution



Introduction

AI is a disruptor, transforming business processes, bringing new opportunities, and helping companies leverage the vast oceans of data they collect. There are myriad use cases, from stock control and forecasting through to identifying efficiency savings in your back office processes. The problem is, AI projects need a team of highly skilled data scientists and engineers, along with access to significant computing power. Even then, **a typical project will take around nine months to complete.** This is where Sonasoft comes in. We offer a complete AI solution, including skilled data scientists to assess your requirements, a world-class AI platform that creates models autonomously and almost 20 years experience in data storage and processing. The upshot is that we can solve all your AI needs in just 4 weeks instead of 9 months.

An AI 101

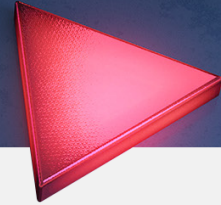
Artificial intelligence is the term for any computer system that is able to emulate some aspect of intelligence. Merriam-Webster defines intelligence as:

**the ability to learn or understand or to deal with new or trying situations.
Also, the skilled use of reason.**

At Sonasoft, we use a slightly different definition. Intelligence is knowing what to do next or in future, based on recently gained knowledge (knowledge of the present) and our experience (knowledge of the past).

Human intelligence is defined as a “general” intelligence. That is, we are able to apply our existing knowledge to solve completely new or abstract problems. We do this constantly, usually unconsciously. Most artificial intelligence is “narrow”. Typically, this means it only has a very specific application. As a simple example, a computer can be taught to recognize pictures of cats. But that same computer can’t recognize dogs, unless it is retrained.

Most AI is based on the concept of machine learning. Here, the computer is taught to recognize certain patterns in data. It then applies this learning to spot the pattern in new data. There are 3 forms of ML.



Supervised learning uses known, labelled data for training. E.g. you show a computer thousands of labelled photos of animals and teach it to identify the ones that are cats.

Unsupervised learning uses unlabelled data. The computer simply tries to identify any interesting patterns within the data. Typically, this might be used to identify clusters of similar data.

Reinforcement learning uses unlabelled data, but each time the computer identifies something correctly, it is “rewarded”. This is rather like how a human infant starts to learn.

What AI can do for your business

There are numerous ways in which AI can transform businesses. Broadly, use cases fall into 3 classes: Anomaly detection, forecasting and planning, and knowledge discovery. There are other use cases, but they are less widely used. Let’s look at each of these in turn and see what it takes to create a project.

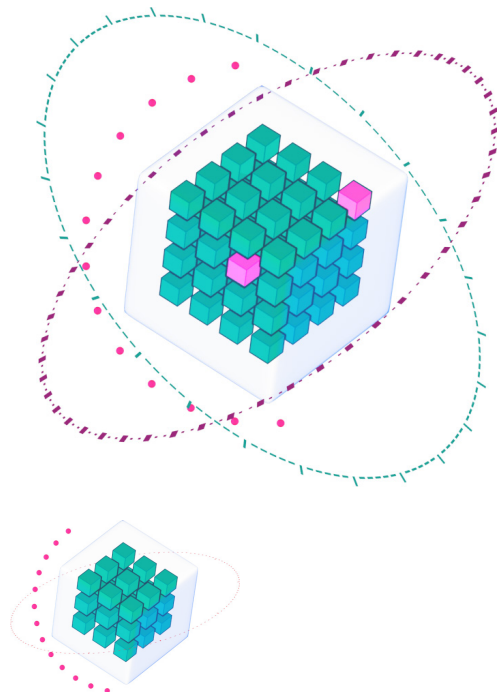
Anomaly detection

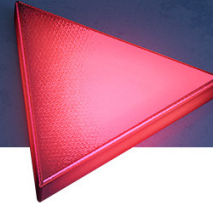
This is about finding anomalies or outliers in historical or real-time data. Anomaly detection has applications in many industries. Here’s just a selection:

Finance. Anomaly detection can be used to identify fraudulent financial transactions instantly. It does this by recognizing unusual patterns in spending behavior. An AI system is able to spot the difference between someone traveling with their credit card and someone using stolen credit card details.

DevOps. Using AI anomaly detection, you can identify an impending failure before it impacts your customers. Often, failures are presaged by changes in behavior in your backend. E.g. database queries taking a bit longer to return.

Cybersecurity. The biggest cybersecurity attacks happen when a hacker gets access to your system





by stealing a password. Typically, they will then try to access sensitive data. Anomaly detection can help spot unusual behavior like this.

Forecasting and planning

Many businesses rely on accurately forecasting future demand. This allows them to plan resource allocation, forecast profits, and streamline their business. Forecasting like this requires analyzing and modeling historical data and then extrapolating the model into the future. This can also be done in reverse. Given a future requirement, when do you need to get resources into place to meet it. Concrete examples of this include:

Just-in-time manufacturing. Here, it is critical to ensure all the required parts are manufactured and delivered at exactly the right time. This approach to manufacturing was pioneered by Japanese motor manufacturers and is so effective it allows one of Nissan's car plants to produce a new vehicle every 2 minutes.

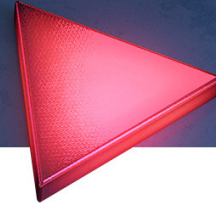
Sales forecasting. Businesses need to predict how much stock they need. This sort of forecasting is incredibly difficult as you need to predict both what products will be in demand and when the demand will come. AI forecasting can offer real insights based on past consumer behavior, current hype, and even factors such as long-range weather forecasts.

Knowledge discovery

One of the more unusual uses cases for AI is knowledge discovery. Here, you teach an AI to discover new patterns in data and to identify new knowledge.

eDiscovery. Electronic discovery or eDiscovery is used to analyze and find new information relating typically to civil litigation, patent cases, etc. It involves analyzing both the data itself and the meta-information related to it. eDiscovery may also be used during cases relating to data protection e.g. breaches of HIPAA.

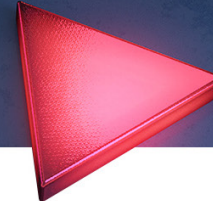
Patent discovery. The traditional approach to patents is through invention. But increasingly, there's an industry in analyzing existing patents to identify areas where there's a missing patent. The systems can even suggest exactly what the new patent should be and identify targets for selling it.



A typical AI project timeline

All the above forms of AI are applications of machine learning (ML). Creating an ML project is time-consuming and requires expertise. To give an example we will look at the steps needed to train a supervised learning model.

- 1 Scope the problem.** The first step is to define the problem and understand if you have suitable data to work with. This stage will take several weeks to complete.
- 2 Get the data.** Now, you need to process your data to get your data into a form where it can be analyzed. This can be especially problematic with historical data since often the formats are different or could have changed over time. Again, this process will take weeks.
- 3 Migrate to the cloud.** While it is possible to run AI models locally on-premises or on GPU-enabled laptops, this is an inefficient approach. So, you really need to move the data into the cloud. This is because most real-world ML models need the enormous computational and storage resources of the cloud. If your team includes people with suitable expertise, this can probably be done in just days. Otherwise, it could take much longer.
- 4 Clean up the data.** At this stage, you need to pre-process the data. This includes cleaning, filtering, and potentially manually labeling the data for supervised learning. This whole process is slow, and many enterprises tell us that preparing data for AI is one of their big challenges.
- 5 Select ML models.** Having done this, your data scientists can start trying to find suitable ML models to analyze the data. Choosing the correct model is key, and it is often based on experience and gut instinct.
- 6 Train and verify the model.** The next stage is to start training the model. For this, you need good quality training data. Typically, this means you need to further process the data you are analyzing. Most ML models involve 10-20 control parameters (called hyper parameters) which need to be iteratively fine-tuned to arrive at best accuracy. Today AutoML promises “Automated Hyper Parameter Tuning”. However, these parameters are also dependent on other choices made in step 4, like embedding type, scaler used to normalize data, shape of data, etc. This will take days to do, and typically, you will have to repeat the process many times.

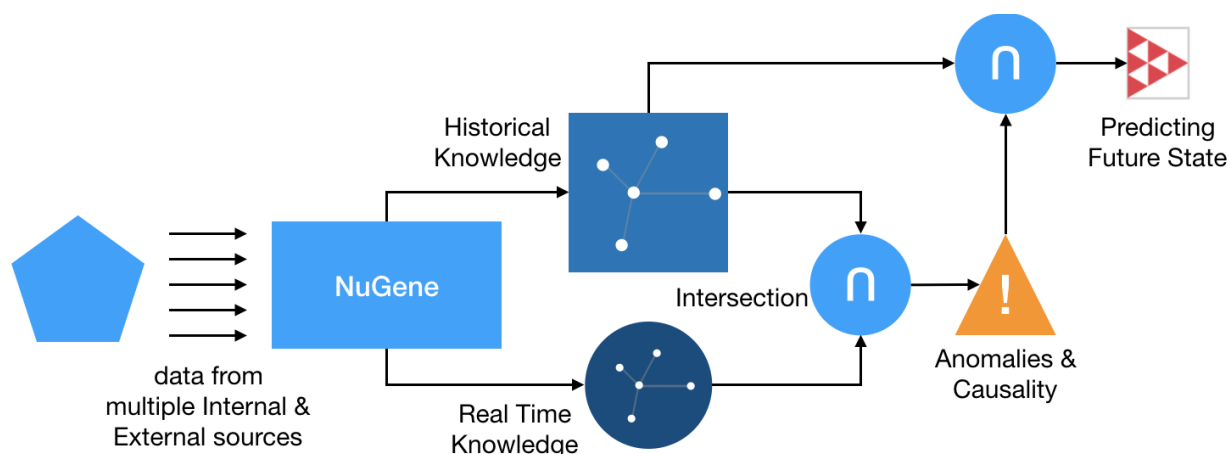


7 Validate the model is fit for purpose. Finally, you will have a trained model and can test whether it is suitable for the job needed. Validation takes days to weeks, depending on the sort of task involved.

Overall, this process will take 6-9 months to complete (from defining the problem to ending up with a working model). Worse still, this model will be so specialized, it will only be applicable for a single function. If you want to create a new model to do something else, you will have to start from scratch.

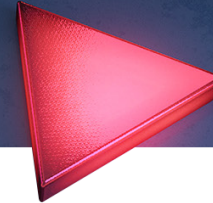
The Sonasoftware approach

Sonasoftware are experts in leveraging AI to extract meaningful business insights from your data. We offer a complete package and are able to condense your 6-9 month AI project down to just days. Our complete package consists of three key elements.



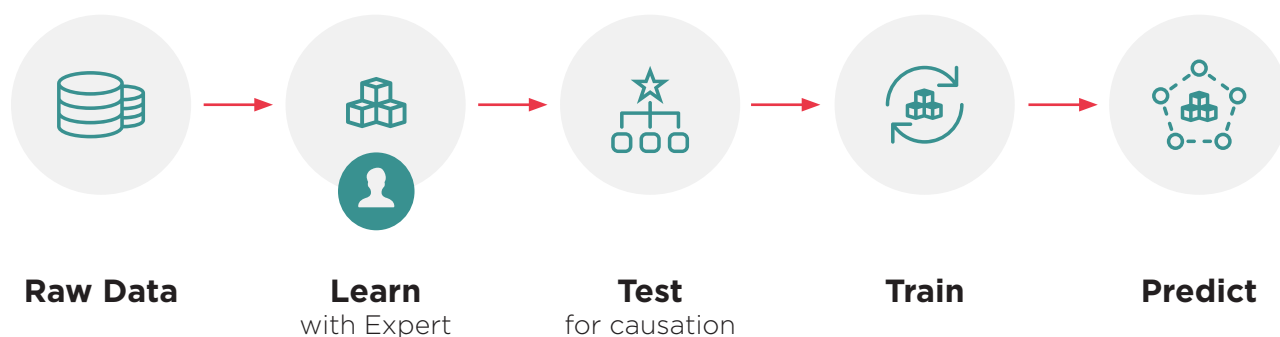
Data science

Getting your data into a suitable form for AI is often the hardest part of any transformation journey. But our large data science team are **experts** in migrating data into the cloud. We can also share our knowledge and advise you on what data you need to collect to achieve your requirements. Overall, our knowledge and expertise can save you months of wasted effort as well as saving you a fortune in hiring your own team.



NuGene

NuGene is our proprietary AI engine. NuGene differs from most AI engines in 2 key ways. Firstly, it is an autonomous AI for creating machine learning models from your data with no human input. Secondly, it is a universal narrow machine learning platform. By that we mean it isn't limited to just using a few ML models. NuGene understands all the latest forms of machine learning, and will always find the ideal one for your application.

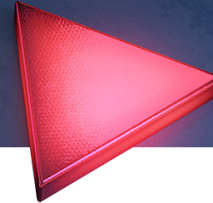


NuGene – the first Universal Narrow AI

NuGene was developed as an AI to create functional narrow AIs, hence the term “universal narrow AI”. NuGene will take your data and your desired outcome and then start to create and test machine learning models. Importantly, NuGene wants raw data. This means that it isn't being biased by the assumptions made during data preprocessing.

Like humans, NuGene observes cause and effect based on identified anomalies and lists a variety of causal hypotheses. Because it handles raw data, NuGene is free to find and test any potential patterns in the data. It then uses multiple unsupervised learning techniques to establish the hypotheses. NuGene rigorously validates its hypotheses to differentiate between correlation and real causality, allowing it to truly learn from the data. Finally, it is able to take these hypotheses and construct detailed ML models.

NuGene can also generate charts and graphs to illustrate its understanding of the data. This allows human experts to validate the model and intervene to steer it in the right direction. This avoids the issue where NuGene may lack some critical data or insight that the human experts know. However, you should be aware that this risks including bias.



Data expertise

Sometimes, you may know what you want to achieve with AI but lack the appropriate data. In such cases, we can help you set up the data gathering you need. As your data is collected, NuGene will start to process it to see if it can see anything significant. Once it has enough data, NuGene will generate the model you need.

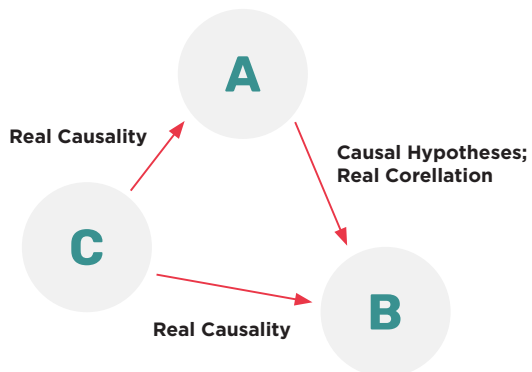
The Sonasoftware difference

Sonasoftware's integrated AI solution stands out for four main reasons.

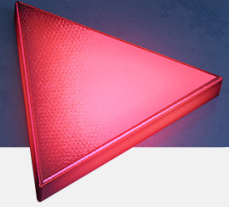
Industry and application agnostic. Our AI solution can be applied to **any** use case, in any industry, with any sort of data. This helps us to stand out against the opposition. The flexibility of our approach is down to the expertise we have in data science, coupled with the unique way NuGene analyzes your raw data.

Broad learning, not narrow feature extraction. Most data scientists talk about "feature extraction". They set out with the goal of feature engineering your data to identify a specific feature. This narrow approach can mean many important features are missed, and severely limits the learning potential of any model. By contrast, NuGene looks at the bigger picture and will find any interesting patterns in the data.

Causality, not just correlation. One of the unique differences with NuGene is that it understands the mantra "correlation does not imply causation". Having identified a possible correlation, NuGene will then test the hypothesis thoroughly before deciding if it is true causation.



Time series are critical. The central insight for Sonasoftware's AI experts was the realization that all **data is time-dependent**, something all our competitors miss. Take, for example, data for predicting bank loan defaults. Without an understanding of the wider temporal aspects like the macroeconomic climate, this data is useless. You can't compare a loan default in 2009 with a potential default in 2019. In NuGene, all data is entered as a time series. And the data isn't just numeric. NuGene also understands free text, images, sound, and structured data.



Conclusions

Sonasoftware's integrated solutions allow any business to leverage AI in just a few weeks. Moreover, our AI models are not limited to solving one problem. This makes our offering unique, and will save you months or even years of development effort. Add in the efficiency savings the AI models themselves bring and you get a truly transformative impact on your business.