Research partner

IBM

Independent research by

Chartis

Demystifying Artificial Intelligence in Risk and Compliance

A Step-by-Step Guide

April 2018

About Chartis

Chartis Research is the leading provider of research and analysis on the global market for risk technology. It is part of Infopro Digital, which owns market-leading brands such as Risk and Waters Technology. Chartis's goal is to support enterprises as they drive business performance through improved risk management, corporate governance and compliance and to help clients make informed technology and business decisions by providing in-depth analysis and actionable advice on virtually all aspects of risk technology. Areas of expertise include:

- Credit risk
- Operational risk and governance, risk and compliance (GRC)
- Market risk
- Asset and liability management (ALM) and liquidity risk
- Energy and commodity trading risk
- Financial crime including trader surveillance, anti-fraud and anti-money laundering
- Cyber risk management
- Insurance risk
- Regulatory requirements including Basel 2 and 3, Dodd-Frank, MiFID II and Solvency II

Chartis is solely focused on risk and compliance technology, which gives it a significant advantage over generic market analysts.

The firm has brought together a leading team of analysts and advisors from the risk management and financial services industries. This team has hands-on experience of implementing and developing risk management systems and programs for Fortune 500 companies and leading consulting houses.

Visit ***www.chartis-research.com*** for more information.

Join our global online community at ***www.risktech-forum.com***.

© Copyright Chartis Research Ltd 2018. All Rights Reserved. Chartis Research is a wholly owned subsidiary of Infopro Digital Ltd.

No part of this publication may be reproduced, adapted, stored in a retrieval system or transmitted in any form by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of Chartis Research Ltd. The facts contained within this report are believed to be correct at the time of publication but cannot be guaranteed.

Please note that the findings, conclusions and recommendations Chartis Research delivers will be based on information gathered in good faith, whose accuracy we cannot guarantee. Chartis Research accepts no liability whatever for actions taken based on any information that may subsequently prove to be incorrect or errors in our analysis. See Chartis 'Terms of Use' on ***www.chartis-research.com****.*

RiskTech100®, RiskTech Quadrant®, FinTech Quadrant™ and The Risk Enabled Enterprise® are Registered Trade Marks of Chartis Research Limited.

Unauthorized use of Chartis's name and trademarks is strictly prohibited and subject to legal penalties.

About IBM Watson® Financial Services

IBM is working with organizations across the financial services industry to use IBM Cloud, cognitive, big data, RegTech and blockchain technology to address their business challenges. Watson Financial Services merges the cognitive capabilities of Watson and the expertise of Promontory Financial Group to help risk and compliance professionals make better informed decisions to manage risk and compliance processes. These processes range from regulatory change management to specific compliance processes, such as anti-money laundering, know your customer, conduct surveillance and stress testing.

To learn more about IBM risk and regulatory compliance solutions, visit ***ibm.biz/RiskAndCompliance*** and follow us on Twitter ***@IBMFintech***

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Table of contents

© Copyright Chartis Research Ltd 2018. All Rights Reserved

List of figures and tables

© Copyright Chartis Research Ltd 2018. All Rights Reserved

1. Executive summary

Artificial Intelligence (AI) is everywhere, in everything from self-driving cars to social media chatbots. To many it seems amorphous or even vaguely terrifying, but it need not be. At its heart, AI is a set of statistical processes – and like any other statistical process, it needs to be understood and managed in the right way. But its inherently statistical nature is often hidden behind buzzwords and hype.

AI has enjoyed some success in the finance industry, particularly in the areas of risk and compliance. Financial Institutions (FIs), responding to massive data burdens and onerous regulatory requirements, have turned to AI, using it to segment data sets and tune their analytics.

Historically, FIs have used AI techniques in areas such as credit card fraud, but gradually usage is spreading. AI is proving particularly effective in areas that involve large numbers of documents and repetitive processes, mainly in automating legal, compliance and risk documentation. It has also been useful in analyzing data sets, such as those used to detect Anti-Money Laundering (AML).

Traditionally, these use cases have focused on automating costly manual work, leaving key decisions to human beings or rules-based systems. In some cases, FIs have been able to automate as much as 90% of their processes. In an environment shaped by statistics, however, AI tools have to compete with other processes that may be more effective, or more financially justifiable. In short, AI must *work*, and success is not a given.

So how can FIs cut through the hype, fear and mystery surrounding AI to ensure they are employing its tools and techniques in the right way? To demystify AI in risk and compliance, and make it work, FIs should take three steps:

1. **Understand and define what is available**. Iterative, rules-based or explicative? To apply a statistical algorithm effectively, FIs must know which one to use, and understand its associated processes.

2. **Pinpoint the problem to solve, and its associated use case**. AI has seemingly limitless possibilities, so FIs must be able to focus on what is important.

3. **Use the right technology**. Data is at the heart of AI, and the spread of new hardware is driving AI's widespread adoption. But different AI tools require different technology combinations.

In this report we look at AI in action – how its tools and techniques are helping FIs cope with risk and compliance tasks, illustrated with a number of case studies. Then we take steps to demystify AI, providing a taxonomy of AI techniques and a straightforward guide to making them work. Finally, we look at emerging AI applications in the world of finance, before considering the crucial steps FIs can take to accelerate the adoption of AI throughout their industry.

© Copyright Chartis Research Ltd 2018. All Rights Reserved

2. AI in action: an overview

Getting past the fear and hype

AI is not a new concept. Many of the algorithms it uses have been around for 20 years or more – Bayesian inference[1], for example, is more than 250 years old. AI is commonly defined as 'a branch of computer science dealing with the simulation of intelligent behavior in computers.'[2] 'Artificial Intelligence' also tends to be used as an umbrella term covering a number of capabilities, including machine learning, Robotic Process Automation (RPA), Natural Language Processing (NLP) and rules extraction. It is also frequently mythologized, held apart from other forms of analytics, and hailed as a 'magic bullet' capable of solving any business problem.

But AI tools are simply *statistical processes*. These come in a wide variety of forms, with relative strengths and weaknesses. At the moment, the notion of an all-purpose statistical process suitable for all use cases and applications remains a myth. The common factor that differentiates AI systems from other types of analytics is their application of iterative, 'learning' processes. Using these, AI systems can adapt to inputs to enable software to perform tasks more efficiently, and often automatically. It is this ability to achieve automation that is the real hallmark of most current AI applications.

A useful way to demystify AI is to look at instances where it is being used with relative success. One notable testbed for AI – and the sector in which it has had several successes – is the financial services industry, particularly the risk and compliance areas.

AI in financial services

FIs have used machine-learning algorithms to tackle credit card fraud for a long time, and some trading firms use AI techniques to maximize returns over a set period, by employing genetic algorithms and machine learning tools that adapt to feedback from the markets. This should come as no surprise: FIs deal mostly in quantitative data, and their processes are normally intermediated by statistics. Many, in fact, sit at the cutting edge of statistical techniques.

By virtue of this technical and cultural environment, AI has established a seat at the FIs' table. But far from dominating proceedings, it has had to compete hard with other statistical processes to earn its place in the organization. These approaches[3] may produce better results, or may simply display their benefits more openly. Such scenarios can lead to an 'AI for AI's sake' form of adoption, in which AI solutions are employed in organizational niches that are already successfully occupied, in the process stripping away much of the gloss that surrounds the technology. In risk and compliance, however, the bottom line is that AI must actually work.

How AI is used in risk and compliance

AI in risk: AML/KYC

Controlling financial crime is a priority for most FIs. They are continuously evaluating the best ways to safeguard their systems, their data and ultimately their clients. Applied to certain processes, AI techniques can help to standardize manual, time-consuming tasks and make them more efficient:

- **Robotic process automation**. Speeding up routine tasks and minimizing human error.

[1] *A statistical process, devised in the mid-eighteenth century, which updates probabilities as more information becomes available.*
[2] ***https://www.merriam-webster.com/dictionary/artificial%20intelligence***
[3] *'Traditional' quant-based trading, for example, or rules-based risk analytics.*

© Copyright Chartis Research Ltd 2018. All Rights Reserved

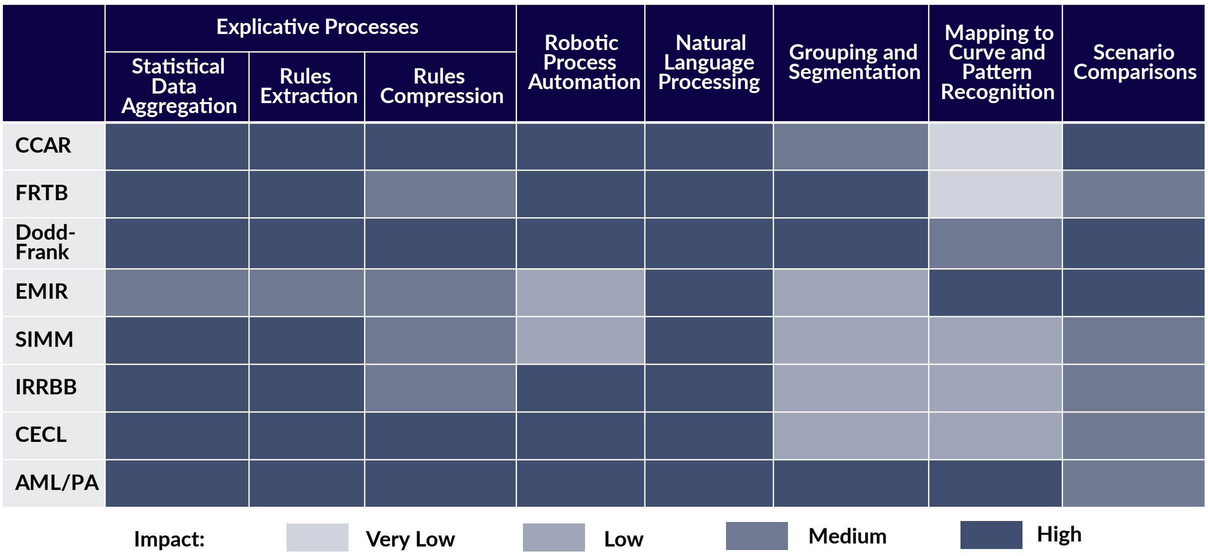
- **Text analytics and insights**. Processing unstructured data, and/or identifying relevant content, negative news, case notes and more.

- **Entity resolution and network analytics**. Determining connections between individuals in order to evaluate risky parties and networks.

AI in compliance: regulatory change management

FIs, weighed down by regulatory monitoring and reporting requirements, can employ a number of AI techniques to lighten the load (see Figure 1).

Figure 1: AI processes for specific regulations



Note: the AI tools and techniques listed here are explained in more detail in chapter 3.
Source: Chartis Research

Every reporting process FIs engage in involves many documents and repetitive manual processes. NLP and RPA are particularly useful in meeting compliance requirements (NLP, in fact, is so far the most commonly used AI process). Meanwhile, scenario comparisons are essential for Comprehensive Capital Analysis and Review (CCAR), Dodd-Frank and European Market Infrastructure Regulation (EMIR) stress tests[4].

To deal successfully with regulatory change management throughout the organization, FIs have to integrate content from thousands of regulatory publications each month. These changes involve complex interactions between different areas of the business, and have second- and third-order effects (and other impacts beyond those). FIs can restructure a portfolio based on regulations, for example, which will affect each of the assets contained within the portfolio, and potentially other portfolios. These kinds of knock-on effects will be common – and the previous example is a relatively straightforward one.

Automating the process of regulatory change management is something of a ‘holy grail’ in the use of AI. The issues FIs face in achieving it are not limited to the sheer complexity involved: regulatory compliance is a significant undertaking, with potentially huge penalties for errors (such as hefty fines for non-compliance). It is therefore vital that FIs learn from others’ mistakes, heeding the lessons of previous AI implementations and successful projects.

[4] *For more information on these methods see chapter 3 and Appendix B.*

© Copyright Chartis Research Ltd 2018. All Rights Reserved

A step back: why is AI so prominent?

Why has AI risen to prominence now, and why is it used so widely in finance, risk and compliance?

Clearly hype plays its part. FIs are certainly not invulnerable to hype, but they are also under intense pressure to spend wisely, save where possible, and avoid failure at all costs, particularly when it comes to compliance and risk management. Several developments have helped to push AI up the agenda:

- **FIs' technical demands are greater than ever**. Massive volumes of data and new regulations have forced them to develop new, often hugely complex systems. Processes around derivatives, valuation adjustments, accounting and AML have all become more demanding and complicated.

- To manage their compliance, **FIs have hired many full-time employees** to aggregate and report on the huge amounts of risk data surfaced daily within their organizations.

- **FIs are moving away from their initial panic in the face of regulation to more measured compliance processes**. After the financial crisis, the shortcomings of an industry riddled with inefficient legacy systems and heavily reliant on manual processes were laid bare. Investment is flattening and fewer major risk and compliance initiatives are being announced; FIs' focus has shifted from compliance to performance, and in particular to the efficiencies that automation promises.

Meanwhile, at the same time:

- **Technology is becoming more powerful**. Compute power is increasing and costs are falling: the cost of processing, and the cost of tools such as graphics processing units [GPUs*], which are particularly powerful when used alongside AI.

- **Data is digitized**. There are more ways for external information to be digitized and used in systemic processes, including document management and media (audio and video) storage.

- **Data has increased in volume, variety and velocity**. The vast lakes of Big Data generated by this wave of digitization present a significant integrity and control challenge for FIs – and an opportunity to feed more powerful AI algorithms.

The confluence of these trends has supported the growing use of a range of new AI capabilities and tools with multiple applications, including:

- Interpreting unstructured and qualitative data outputs.

- Identifying complex, non-linear patterns in large data sets.

- Improving the accuracy of calculations.

- Reducing complex data sets to simpler or more tractable forms.

[*] *These electronic circuits enable parallel processing to handle large amounts of data, and are normally used for processing images.*

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Case studies

How can some of the successful applications of AI in finance help to demystify the technology? Chartis Research has spoken with seven FIs implementing major AI projects at varying levels of completeness and ambition. The findings are detailed in Appendix A, but to summarize:

- Three projects had been completed. Four were still in-flight at the time of writing.
- Automation levels achieved ranged from about 50% of relevant processes to as much as 90%.
- The most common techniques used were machine learning and NLP, followed by rules extraction and compression, clustering and evolutionary algorithms. (We will look at these and other AI variants in more detail in the next chapter).

Several key themes also emerged from these case studies, reflecting broader adoption trends:

- The *iterative portions* of the AI process tend to focus on insignificant (often non-regulated), repetitive decisions, while 'important' decisions are passed to human users or rules-based systems. These systems also use case-specific algorithms that employ Markov models[5] and machine learning to automate manual tasks.
- In many of the examples involving document management, the barrier to entry for the AI algorithm was low. FIs had not previously automated these processes in any way, and had no pre-existing statistical process with which the AI tool had to compete.
- Notably, while AI systems deliver tangible benefits, FIs are wary about how much they are willing to trust to them in the early stages. They are generally using iterative and/or machine-learning techniques for only a small part of a process; much of the remaining automation is handled by rules-based systems or human users.

[5] *Models that describe probabilities in a sequence of events.*

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Example: keeping an eye on conduct – AI in trader surveillance

FIs can use AI capabilities to monitor employees and traders, by applying them to the following areas:

- **Behavior-based trader profiles**. By combining trade data with electronic and voice communications records, FIs can develop behavior-based trader profiles and track emerging patterns of behavior to predict latent risks.

- **Continuously updated risk models**. Using multiple data sources, these can detect links between employees.

- **Risk-based prioritized alerts**. The system can generate alerts based on suspicious activity, prioritize the alerts by level of risk to enable effective time management, and assess the evidence for each alert to enable efficient surveillance.

- **Analyses of voice communications**. The system can perform speech-to-text conversion, and use machine learning algorithms to analyze the language that employees use.

Clearly, despite the challenges, FIs are finding ways to successfully use AI. As part of our step-by-step approach to simplifying and demystifying AI, in the following chapters we look more closely at the tools they are using.

© Copyright Chartis Research Ltd 2018. All Rights Reserved

3. Demystifying AI: definitions and best practices

To apply AI effectively, FIs should follow three steps:

- Understand what types of AI – its algorithms and models – are available.
- Clearly define the problem they want to address, and select a relevant use case.
- Use the right technology to support their chosen AI solution.

Step 1: Understanding and defining AI – which type is most relevant and useful?

A distinguishing feature of AI is that it is iterative and can adapt to inputs. But as we have seen, this definition is somewhat blurred: many statistical algorithms can be made to be iterative[6].

To get a clearer idea of what AI tools can do, we have to place them within the wider universe of statistical techniques. A broad taxonomy of algorithms, like the one shown in Figure 2, could be based on three pillars:

1. *Iterative statistical techniques*. These provide a chain of related and repeated calculations that converge toward a solution. Most of the techniques traditionally thought of as AI – such as machine learning and evolutionary programing algorithms – are included in this category. Iterative techniques have two subsets: machine learning algorithms and evolutionary algorithms.

2. *Potentially iterative techniques*. These are statistical processes that are not necessarily iterative in nature, but which can have iterative techniques 'embedded' within them. For example, an FI could use graph analytics to determine the shortest distance in a network of transactions between two banks. But if part of the network is being processed by a machine learning system, then the graph analytics process becomes iterative.

3. *Non-iterative (rules-based) techniques*. These provide results based on basic logic rules (such as IF, THEN and so on). They can also be arranged into more complex state-dependent rules networks (so if Rule A gives a particular result, the system goes to Rule B).

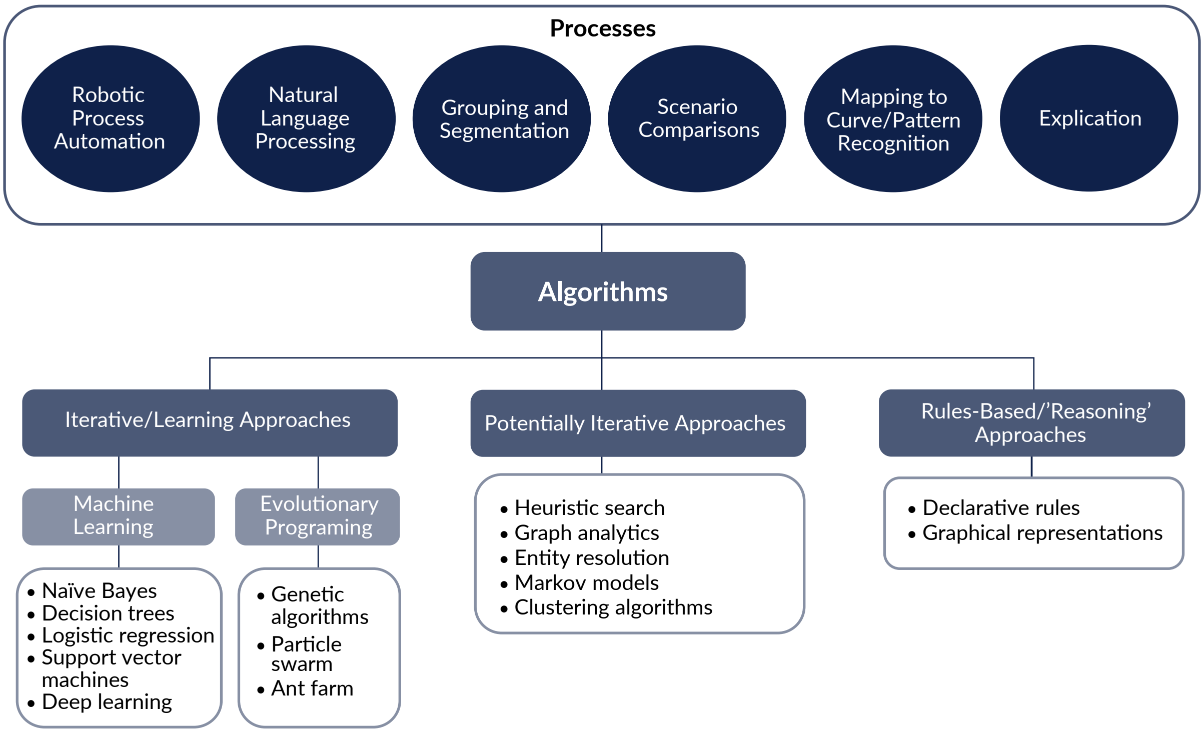
(For more information on the statistical methods outlined in Figure 2, see Appendix B.)

Once FIs have mapped out the appropriate AI algorithms and understand their respective strengths and weaknesses, they can use them for broad processes. There are several ways to approach NLP, for example, such as simple Markov models, hidden Markov models, or machine learning techniques (see Table 1).

[6] *A simple Markov model with a visible state is not an iterative process, for example, but a hidden Markov model is. For more information on Markov Models, see Appendix B.*

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Figure 2: An AI taxonomy



Source: Chartis Research

Table 1: AI processes

Process	Name	Explanation	Appropriate algorithms
Explicative processes	Rules compression	FIs can have thousands of rules, across multiple areas. This technique reduces overlapping rules down to simpler ones (reducing overlapping trading limits down to a single rule, for example).	• Entity resolution. • Graph analytics.
	Rules extraction	Can provide simpler, more efficient versions of advanced analytics. So, for example, it can be used to explain the behavior of neural networks and supervised vector machines, reducing their often complex results to a simpler set of rules.	• Entity resolution. • Graph analytics.
	Statistical data aggregation	Combining data into an aggregate to reduce the time taken to query large data sets. This makes processes quicker and potentially less costly.	• Entity resolution. • Graph analytics. • Clustering algorithms.
Robotic Process Automation		Configurable computer software used to interpret and automate existing applications normally performed by humans.	• Machine learning. • Heuristic search.

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Process	Name	Explanation	Appropriate algorithms
Natural Language Processing		Uses computer programs to analyze languages. NLP has evolved over time, employing a number of different techniques of increasing statistical complexity. These include machine learning and deep learning techniques.	• Markov models. • Machine learning.
Grouping and segmentation		Segmenting given groups of entities based on similar transactions or traits.	• Entity resolution. • Clustering algorithms.
Mapping to curve and pattern recognition		Providing adjustments to ensure that a given algorithm fits or responds to a curve. This might include mapping trading decisions that depend on buy-sell patterns, or recognizing more complex patterns in data (recognizing a potential fraudster from financial crime data, for example).	• Machine learning. • Evolutionary algorithms.
Scenario comparisons		Simulating processes to determine the outcome if a given variable is changed. Often used for stress testing.	• Machine learning. • Evolutionary algorithms.

Source: Chartis Research

Step 2: Pinpoint the problem and select the use case

Once FIs understand the broad taxonomy of AI algorithms and processes, they can apply them to given problems. FIs should first determine the problem(s) they want to solve – improving their data quality, for example, or regulatory reporting – and then focus on the specific use case or cases that feature those problems.

Problems

Algorithms and processes can be applied to a range of risk and compliance problems, including regulatory compliance, fraud, cyber security and data quality (see Table 2).

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Table 2: Applying AI techniques to specific problems

	Processes				Algorithms	
	Robotic Process Automation	Natural Language Processing	Rules Compression	Statistical Data Aggregation	Evolutionary Programing	Machine Learning
Regulatory compliance	Yes	Yes	Yes	Yes		Yes
Fraud analytics		Yes	Yes	Yes		Yes
Credit analytics					Yes	
Cyber security					Yes	Yes
Data quality		Yes				Yes
T&C extraction	Yes					Yes
Equity strategy analysis	Yes				Yes	Yes
Customer engagement/ Conduct risk	Yes	Yes				Yes
KYC	Yes	Yes			Yes	Yes
AML/Patriot Act	Yes	Yes			Yes	Yes
Trader surveillance	Yes		Yes		Yes	Yes

Source: Chartis Research

For fraud analytics, for example, FIs might use NLP to analyze unstructured information (unstructured data associated with a potential fraudster, such as scanned documentation). Fraud analytics also contain a number of process steps that can be assigned to an RPA tool. Machine learning can be used to 'learn' the workflow process for RPA, or to dynamically tune how an automated workflow responds to the types of clients that are being on-boarded.

Data: the biggest challenge

Data remains FIs' biggest challenge. For most, it is the lifeblood of the organization, and problems with its quality and reliability are a serious consideration. Analytics require good data: weak analytics with good data will normally give better results than advanced analytics with bad data. AI tools depend heavily on the quality of the data they have to work with.

AI is often being turned toward the data management and validation process itself, which is an integral part of all the problems listed in Table 2. Document management is especially pertinent, as highlighted by the case studies analyzed for this report (see Appendix A). This is the process of converting documents (essentially unstructured data) into structured formats, while preserving the information they contain. The goal of FIs in many of the case studies was to obtain segmented, cleaned and structured data sets that they could then feed into their compliance and control activities.

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Use cases

Deciding on the most appropriate use case for AI will depend on factors specific to the problem being solved:

- The **strengths and weaknesses** of the AI techniques used.
- The **decisions** that are being automated. FIs tend to be less comfortable automating an important decision or passing it to an AI process.

Measuring strengths and weaknesses, and asking the right questions

AI can potentially be applied to a huge range of use cases, from AML segmentation to document management and trade monitoring. Each use case has a different requirement, and data of differing volumes, types and time-series. Different use cases also often have very different planned outcomes: is the goal to process data, for example, or to detect fraudsters?

Take machine learning as an example. To define the correct use case, risk and compliance officers should ask themselves several questions:

- Do I need to be able to easily explain the tool to regulators/managers/other relevant parties?
- Is there a non-linear relationship between variables? (Transaction volumes do not normally increase linearly over time, for example).
- How resistant is the algorithm to over-fitting? For example, is an algorithm likely to over- or under-interpret changes in the data as being part of long-term trends?
- How much data do I have to work with?
- Should the technique scale easily – will it have to work with small and large data sets?
- How much computational power or infrastructure will I require to run the algorithm?
- How long will it take to train the algorithm, and can we afford to spare this time?

Applying these questions to the subset of machine learning techniques gives an idea of their relative strengths and weaknesses (see Table 3). Note that due to 'cross-pollination' and more advanced forms of certain statistical models, these strengths and weaknesses are guidelines rather than hard-and-fast rules. For example, random forests and gradient-boosted decision trees are fairly complex models, and often difficult to explain.

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Table 3: Strengths and weaknesses of machine learning techniques – answering key questions

		Easily explicable?	Can process non-linear relationships?	Resistant to overfitting?	Can work on small or limited data sets?	Scales well?	Low computational requirements?	Quick to train?
Machine learning technique	**Decision trees/ Naïve Bayes**	Yes	No	Yes	Yes	Yes	Yes	Yes
	Logistic/linear regression	Yes	No	No	Yes	Yes	Yes	Yes
	Deep learning	No	Yes	Yes	No	Yes	No	No
	Support vector machines	No	Yes	Yes	Yes	No	No	No

For more information on these tools and techniques, see Appendix B.
Source: Chartis Research

© Copyright Chartis Research Ltd 2018. All Rights Reserved

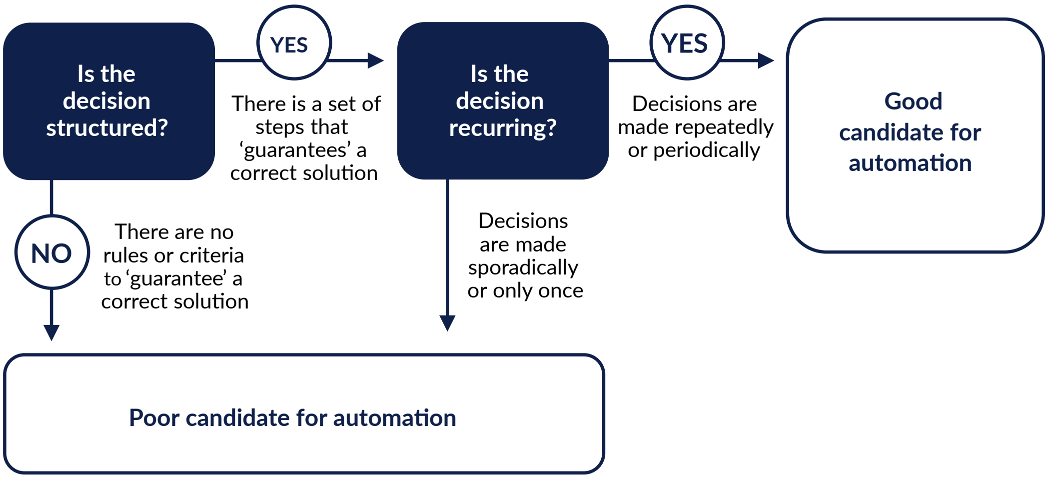
An FI may decide that for a data validation use case it wants to use a low-impact, quick to implement algorithm on a relatively small data set. Based on this analysis, its best option would be a linear regression model or a decision tree.

By asking the right questions and employing the right use cases, FIs can ensure they select the best and most relevant AI techniques. If they have not implemented AI before, however, problems may arise. They should also take into account *exactly* how they are applying the solution, by considering the *importance* of the decision the algorithm is making.

Determining the type and significance of the decision

One of AI's big strengths is that it can automate decisions that would normally be made by humans. A number of factors indicate which areas within an FI are most appropriate for automation – they are usually structured and recurring decisions (see Figure 3). During the on-boarding process in a retail bank, for example, staff carry out a number of defined checks before they create an account, and these are repeated for each new client. Iterative processes need to 'learn' from repeated steps, so AI tools are unlikely to be good at making sporadic or unpredictable decisions. Systems designed to respond to low-frequency events such as earthquakes, for example, probably shouldn't involve AI algorithms.

Figure 3: Automating decisions



Source: Chartis Research

Where the AI is applied in the process is also important. In strongly regulated firms, or those with restrictive governance structures (most FIs, in other words), machine learning processes are often not assigned to a *significant* decision – one that must be justified to regulators or stakeholders.

In such cases, AI often works better as a sorting mechanism that **feeds into** rules-based or user-based decision making, rather than one that makes the decisions itself. Rules-based systems or users take more responsibility for outcomes, so we can think of *artificial intelligence* more as *assisted intelligence*. In this case it is unlikely to provide a one-way ticket to headcount reductions – except, perhaps, in areas with more lax governance structures, where firms are less concerned with justifying their results. How much oversight there is of the process will therefore be a key factor in determining whether AI should be adopted.

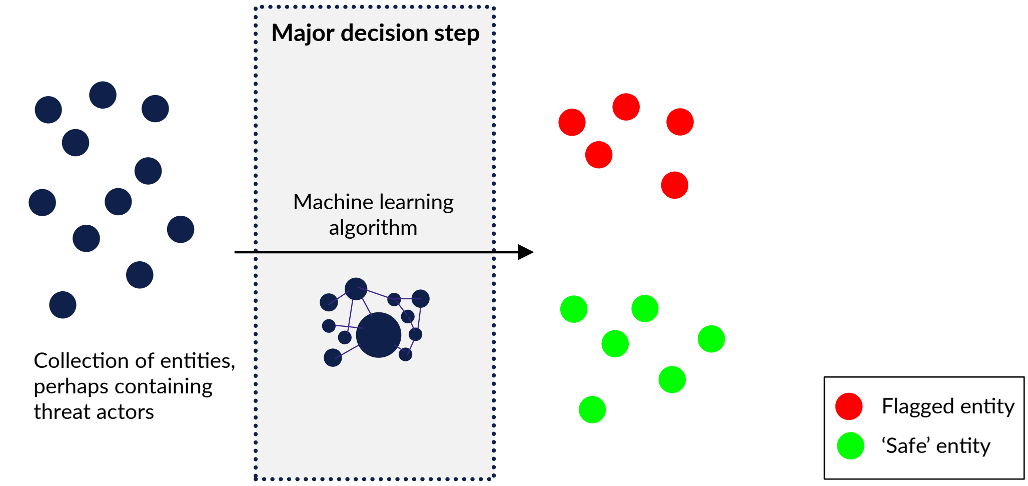
Figures 4 and 5 show this principle applied to an AML use case, in which a system must determine potential threats from a collection of entities. In this case, the algorithm is an entity resolution engine

© Copyright Chartis Research Ltd 2018. All Rights Reserved

that must determine whether entities are fraudulent or have been sanctioned. Figure 4 shows the process in which the AI tool makes the decision, while Figure 5 shows the AI tool inputting data into a rules-based decision-making process.

In Chartis' experience, FIs are more likely to choose the second option than the first. FIs want to be very confident in their systems, and want to be able to explain all significant decisions to a high level of granularity. The process of segmenting the data is less significant in that context than the process of flagging – and it can be automated.

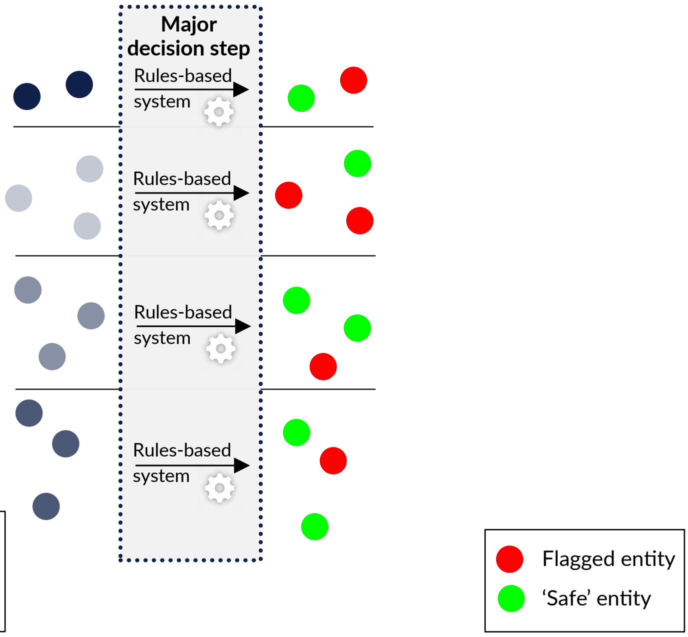
Figure 4: A machine learning algorithm as a decision-maker



The machine learning algorithm is performing the major decision step in the process, in determining which entities are 'safe' and which are flagged. This is considered to be risky, as the firm must be very confident in the algorithm and how it is being explained.

Source: Chartis Research

Figure 5: A machine learning algorithm as an input to decisions



The major decision step in the process (i.e. flagging) is performed by the rules-based systems – the machine learning algorithm is 'sorting' prior to this step.

Source: Chartis Research

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Step 3: Use the right technology

Two technological developments have enabled and driven the algorithms underpinning AI:

- **Raw materials: data**. AI requires data, and FIs have a lot of it to work with. More new tools and databases to handle it (such as NoSQL) have been emerging in recent years.

- **Hardware**. Large volumes of data and sophisticated algorithms require powerful processing capabilities. One key enabler of this is parallel processing, supported by relatively inexpensive GPUs.

Data and databases

The number and variety of new databases, both commercial and open source, is growing all the time, and they cover a range of functionalities for risk management applications.

- **Unstructured databases**. Databases such as NoSQL and Hadoop, which enable the processing of unstructured data such as video, audio and text. These are particularly useful for 'fuzzy' use cases such as fraud and financial crime, which often require ancillary information to enable FIs to make judgements.

- **Array databases**. These store data in grids of one, two or more dimensions, enabling specific data types to be arranged more efficiently, and supporting cluster algorithms.

- **Columnar databases**, which store data in columns instead of rows. By allowing parallel processing, these can enable advanced AI and machine learning.

- **Graph databases**. These employ graph structures for queries, using nodes, edges and properties to represent and store data. They underpin graph analytics, identity resolution and enhanced workflows.

Many FIs have undertaken major unstructured data projects. Hadoop and Spark – which have a fully featured stack available on open source – offer FIs and vendors an affordable entry point. However, many tools (including machine learning, graph analytics and evolutionary algorithms) are optimized with array-oriented data. As FIs more rigorously analyze the requirements of each part of their compliance systems, we expect to see technology processes branching out more, as FIs target database capabilities that meet their specific requirements in a given area.

Hardware

The spread of new hardware (particularly GPUs) has driven much of the widespread adoption of AI, by making highly accelerated versions of machine learning economically viable. These tools provide powerful parallel processing, and enable users to apply multiple processes (including the iterative processes necessary for AI) to a single unit of data simultaneously.

The impact

Figure 6 shows the impact of hardware and databases on given areas of AI, by highlighting how relevant the particular databases are to specific tools.

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Figure 6: The impact of hardware and databases on AI tools

		Hardware	Database			
		GPU	Unstructured database	Array database	Columnar database	Graph database
Iterative	Simple machine learning (regression etc.)	Low	Very High	Medium	Medium	Medium
	Advanced machine learning (deep learning)	High	Medium	Very High	High	High
	Evolutionary algorithms	High	High	Very High	Low	Low
Partially iterative	Heuristic search	Low	Very High	Medium	Medium	Medium
	Graph analytics	High	Medium	Very High	High	Very High
	Entity resolution	Low	Very High	Medium	Medium	Medium
	Markov models	Medium	Very High	Low	Medium	Medium
	Clustering algorithms	High	Medium	Very High	High	High
Rules-based	Rules-based approaches	Medium	Medium	Medium	High	Medium

Impact: Low **Medium** **High** **Very High**

Source: Chartis Research

To use the right types of AI, FIs should match the problem areas they want to address with the appropriate algorithms to use and the necessary hardware to drive it all. If a piece of this 'approach-algorithm-hardware' triangle is missing, the AI tool being used could run less efficiently, especially in the case of more hardware-intensive machine learning techniques.

© Copyright Chartis Research Ltd 2018. All Rights Reserved

4. Conclusion: accelerating the implementation of AI

Cutting through the hype to the true nature and potential of AI can make it less daunting and more approachable. Demystifying it can encourage FIs to adopt more of its tools to simplify processes and save time and money. To accelerate AI's implementation, FIs can take several actions other than the steps we have already outlined (see Figure 7).

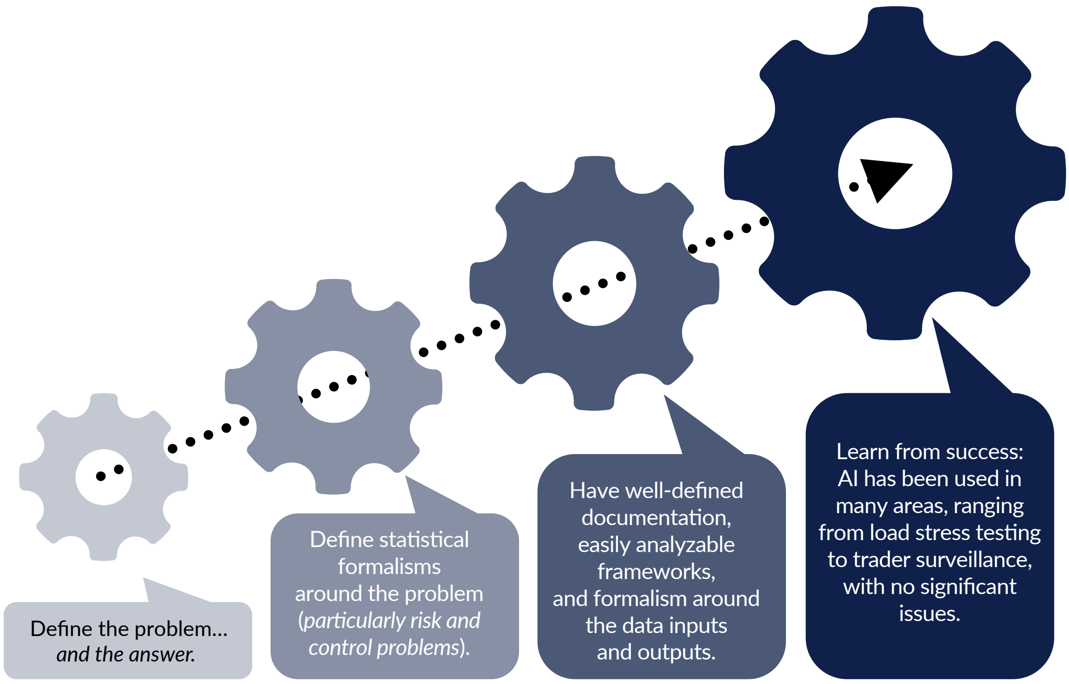
- **Define the problem**. AI should not be used for its own sake. FIs must have a good idea of the scale and scope of a problem before using AI to tackle it. To use iterative techniques, FIs must know the question they want to answer – as well as the range of potential answers – to benchmark the effectiveness of a tool.

- **Define statistical formalisms**. FIs should be prepared to rigorously test and discuss the mathematical constraints and definitions of a given AI technique with stakeholders or regulators. They should not simply apply it to a problem and expect it to iterate its way to a solution.

- **Create well-defined documentation**. FIs must ensure they have well-defined documentation, analyzable frameworks and formalized data inputs and outputs, to ensure that AI techniques are as well understood and controlled in their organization as possible.

- **Learn from success**. AI has been used in many different areas with varying levels of success. To determine how well an AI tool applies to a particular problem and use case, FIs should learn from past successes, looking at specific areas (document management, segmentation, data cleaning) to find the fastest and most efficient wins.

Figure 7: Accelerating AI implementation



Source: Chartis Research

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Finally, FIs should remember several key concepts:

- 'Artificial Intelligence' is a broad term, and to accelerate its implementation FIs must map specific strategies and algorithms to appropriate tasks.

- Good candidates for AI processes are data-intensive tasks that involve repetitive activities, or complex data structures that need to be simplified.

- Implemented properly, AI can deliver depth and breadth of insight, actionable information, and the ability to learn from data.

- Finally, and perhaps most importantly, AI requires the same discipline as any other algorithmic approach.

Contrary to what some have predicted, AI is not taking over the finance industry: it is not curing all FIs' existing problems, nor is it creating a vast array of new ones. It is a powerful tool, however, and used pragmatically it can be stable, effective and beneficial.

© Copyright Chartis Research Ltd 2018. All Rights Reserved

5. Appendix A: case studies

Tables 4 to 10 provide more information on the case studies carried out for this report.

Table 4: AI in finance – case study 1

Legal and regulatory rules management: legal document extraction at a Tier 1 European bank	
What?	Input of legal documents, and transformation into legal and regulatory rules within the bank's systems.
How?	Documents of hundreds of pages in length were converted into rules and events, which were then embedded into the bank's compliance processes.
How much progress?	Program completed. Approximately 80%-90% of the process automated.
Which techniques were used?	NLP and machine learning.

Source: Chartis Research

Table 5: AI in finance – case study 2

Trading, research and business-aligned risk and control: analysis of corporate bonds at a Tier 1 US bank	
What?	Analysis of corporate bonds.
How?	Extraction of key transactional and validation terms and conditions, which were then input into a terms and conditions database used heavily by a front-office trading team. The data was used as part of the trading, research and desk-specific risk/compliance process.
How much progress?	Program completed. Approximately 80%-90% of the process automated.
Which techniques were used?	NLP and machine learning.

Source: Chartis Research

Table 6: AI in finance – case study 3

Investment compliance: document analysis and rules extraction at a Tier 1 US bank	
What?	Document analysis and rules extraction for wealth management and private banking.
How?	Automatic extraction of rules from documents such as mutual fund offers, to remove a limited set of rules from documents. These rules could then be used to create more flexible agreements with clients. The main current focus is investment compliance.
How much progress?	Program in-flight. So far it has automated approximately 40%-50% of the process; the compliance officer still provides sign-off.
Which techniques were used?	Rules extraction, NLP and machine learning.

Source: Chartis Research

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Table 7: AI in finance – case study 4

Regulatory compliance: document management at a Tier 1 Asian bank	
What?	Digitalization and extraction of legal and compliance documents.
How?	Information defined, extracted and archived from specific legal contracts into a format that can be accessed and searched after the result, and then fed into compliance processes.
How much progress?	Still in-flight. Level of automation currently unknown.
Which techniques were used?	NLP and machine learning.

Source: Chartis Research

Table 8: AI in finance – case study 5

Regulatory and internal compliance: AML and KYC at a Tier 1 European bank	
What?	Providing AML and KYC analysis for client on-boarding processes.
How?	Segments and clusters groups of entities, which then have rules applied for Suspicious Activity Reports (SARs). Entities are clustered, and then clusters are validated via machine learning. These are then fed into rules-based AML and KYC analytics systems.
How much progress?	Currently in-flight. 40%-50% automation achieved.
Which techniques were used?	Clustering and machine learning.

Source: Chartis Research

Table 9: AI in finance – case study 6

Limits management and trading compliance: rules extraction at a Tier 1 US bank	
What?	Extracts trading rules from a number of overlapping control rules in multiple trading systems.
How?	Extracts and compresses rules from the bank's legacy trading system, reducing thousands of overlapping rules in areas such as limits management. Historically, these rules were implemented in large batches, but little attention was paid to any potential overlap, until the number of rules became so high that it created a significant burden. Rules extraction was performed to simplify the process.
How much progress?	Currently in-flight. Level of automation unknown.
Which techniques were used?	Rules extraction, rules compression, evolutionary algorithms.

Source: Chartis Research

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Table 10: AI in finance – case study 7

Trader surveillance and control: Tier 1 US capital markets firm	
What?	Segmentation of trader surveillance alerts.
How?	Parsed a number of alerts of manipulation, collusion, unauthorized trading and unethical practices in a Tier 1 hedge fund, using contextual analysis on structured and unstructured data to reduce the number of false positives.
How much progress?	Has been implemented. Automated 50% of the process.
Which techniques were used?	Machine learning, clustering.

Source: Chartis Research

© Copyright Chartis Research Ltd 2018. All Rights Reserved

6. Appendix B: statistical methods and AI processes

Tables 11 to 14 summarize the key features of the AI tools mentioned in this report.

Table 11: Iterative processes: machine learning

Name	What is it?	Strengths	Weaknesses
Machine learning	Machine learning focuses on developing algorithms that can learn from and make predictions and decisions about data, by building a model from sample inputs.	• Adapts to problems without the need for human input.	• Requires a lot of data. • Often requires powerful hardware (such as GPUs). • Can be difficult to explain the results. • Can struggle with discontinuous data, and can be trapped on local maxima on fitness landscapes: namely locating points which are better than the ones nearest them but are not highest or 'best' point.
Types of machine learning			
Naïve Bayes	A classification algorithm that assumes that the presence of a given feature in a particular class is unrelated to the presence of any other feature.	• Easy to understand. • Fast. • Performs well for categorical/ non-numeric assumptions.	• A normal distribution is assumed for numeric variables. • Results depend on an assumption of independence.
Decision trees	An algorithm that uses estimates and probabilities to calculate outcomes.	• Easy to understand/ explain. • Potential options and choices are mapped out. • Determines costs and benefits.	• Probabilities are assumptions and prone to errors. • Uses quantitative/numerical data.
Logistic regression	A probability and risk estimator used to predict a binary outcome (e.g. 1 or 0, or yes or no) given a set of independent variables.	• Relatively easy to use. • Does not assume linear relationships.	• Requires a reasonably large data set. • User needs to specify which interactions are allowed within a model.
Support vector machines	Transform linear data into a non-linear space, then map it into categories which are divided by as wide a gap as possible. Future data is mapped into the space and into one of the categories.	• Can avoid over-fitting. • Relatively easy to control.	• Requires a large amount of data. • Resource-intensive.

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Name	What is it?	Strengths	Weaknesses
Deep learning	Based on learning data representations rather than using task-specific algorithms. Includes neural networks.	• Can address complex problems and data sets. • Reduces the need for future engineering.	• Requires very large data sets. • Has correspondingly large hardware requirements. • Very difficult to explain.

Source: Chartis Research

Table 12: Iterative processes: evolutionary algorithms

Name	What is it?	Strengths	Weaknesses
Evolutionary algorithms	Evolutionary algorithms use mechanisms inspired by biological evolution, such as reproduction, mutation, recombination and selection. Possible solutions to a particular problem are treated like individuals in a population, and their fitness for a particular problem determines their effectiveness as a solution.	• Doesn't require as much data as machine learning systems. • Can work on less homogeneous and bumpier fitness-state models than machine learning systems, due to struggling less with local fitness peaks.	• Weak at linear or quadratic solutions. • Difficult to know when termination criteria are satisfied (when a solution is 'fit' enough).
Types of evolutionary algorithm			
Genetic algorithms	A population of candidate solutions is evolved to try and reach a better solution. Each candidate has a set of properties (its 'chromosomes' or 'genotype') that can be mutated and altered.	• Provides multiple local optima (doesn't get caught on local minima) - by providing a spread of options in a fitness landscape, there is a very low likelihood of a successful candidate being trapped in a low-fitness area. • The number of parameters can be very large, so it can be used for complex problems.	• Can take a long time to converge. • Very difficult to explain. • Difficult to tune due to the wide potential array of parameters.
Ant farm	Uses ant colony algorithms and 'swarm intelligence' methods to find efficient paths through graphs.	• Relatively efficient for a small number of nodes. • Can be used in dynamic situations (i.e. when the problem is changing).	• Can take a long time to converge. • Very difficult to explain. • Difficult to tune accurately due to wide potential array of adjustable parameters.

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Name	What is it?	Strengths	Weaknesses
Particle swarm	Optimization algorithm inspired by biological examples of swarming, flocking and herding phenomena.	• Information-sharing mechanisms are more straightforward than genetic algorithms or ant farms (there is no mutation, for example). • Parameters are easier to adjust.	• More constrained than other evolutionary algorithms. • More prone to getting trapped on local maxima than other evolutionary algorithms.

Source: Chartis Research

Table 13: Potentially iterative processes

Name	What is it?	Strengths	Weaknesses
Heuristic search	A search technique that employs quick, often approximate rankings to arrive at a solution.	• Quick and analytically efficient.	• Approximations can lead to errors.
Graph analytics	Models the relationships between individual objects as a graph. Relationships can be 'one-way' or directed (such as 'parent-child') or undirected (such as 'all-to-all').	• Good at managing relationships.	• Requires a graph-based data model. • Poor at delivering numerical results.
Entity resolution	Connects disparate data sources to match entities and non-obvious relationships.	• Good at connecting duplicate entities such as sanctioned individuals. • Can work with 'fuzzy matching' and incomplete data.	• Using entity resolution with incomplete or partial data sets can be more likely to produce poor results.
Markov models	Create a probability of a given event, depending on the state of a previous event.	• Good at connecting data that comes in dependent strings (such as sentence structures in NLP). • Can be made iterative and more powerful with hidden Markov models.	• Hidden Markov models are powerful, but they require training and are computationally expensive. • The success of the technique relates strongly to how well dependencies are modeled (e.g. how often two words come together in the same sentence).

© Copyright Chartis Research Ltd 2018. All Rights Reserved

Name	What is it?	Strengths	Weaknesses
Clustering algorithms	Group sets of similar objects and entities according to their similarity, and the computing distance between them.	• Break data sets into visually and analytically comprehensible groups.	• Certain data sets cannot be clustered – they do not contain correlations that are meaningful enough. • The type of clustering algorithm dictates the output – so K-means, for example, will segment data in a fundamentally different way to topographical modeling.

Source: Chartis Research

Table 14: Rules-based

Name	What is it?	Strengths	Weaknesses
Rules-based systems	Systems based on simple universal rules such as IF, OR, NOT etc.	• Fast, simple, straightforward. • Easy to understand. • Changes are easy to model.	• Only do what they are explicitly programed to.
Declarative rules	Rules created in specific formats, stating actions that occur whenever a trigger condition is recognized – often as IF -> THEN statements.	• Can effectively enforce decisions. • Straightforward to tune or adjust. • Consistent performance.	• Reliant on initial assumptions. • Knowledge extraction can be difficult and time-consuming.
Graphical representations	Representation of rules as tables or simple decision trees.		

Source: Chartis Research

© Copyright Chartis Research Ltd 2018. All Rights Reserved

7. How to use research and services from Chartis

In addition to our flagship industry reports, Chartis also offers customized information and consulting services. Our in-depth knowledge of the risk technology market and best practice allows us to provide high-quality and cost-effective advice to our clients. If you found this report informative and useful, you may be interested in the following services from Chartis.

For risk technology buyers

If you are purchasing risk management software, Chartis's vendor selection service is designed to help you find the most appropriate risk technology solution for your needs.

We monitor the market to identify the strengths and weaknesses of the different risk technology solutions, and track the post-sales performance of companies selling and implementing these systems. Our market intelligence includes key decision criteria such as TCO (total cost of ownership) comparisons and customer satisfaction ratings.

Our research and advisory services cover a range of risk and compliance management topics such as credit risk, market risk, operational risk, GRC, financial crime, liquidity risk, asset and liability management, collateral management, regulatory compliance, risk data aggregation, risk analytics and risk BI.

Our vendor selection services include:

- Buy vs. build decision support.
- Business and functional requirements gathering.
- Identification of suitable risk and compliance implementation partners.
- Review of vendor proposals.
- Assessment of vendor presentations and demonstrations.
- Definition and execution of Proof-of-Concept (PoC) projects.
- Due diligence activities.

© Copyright Chartis Research Ltd 2018. All Rights Reserved

For risk technology vendors

Strategy

Chartis can provide specific strategy advice for risk technology vendors and innovators, with a special focus on growth strategy, product direction, go-to-market plans, and more. Some of our specific offerings include:

- Market analysis, including market segmentation, market demands, buyer needs, and competitive forces.

- Strategy sessions focused on aligning product and company direction based upon analyst data, research, and market intelligence.

- Advice on go-to-market positioning, messaging, and lead generation.

- Advice on pricing strategy, alliance strategy, and licensing/pricing models.

Thought leadership

Risk technology vendors can also engage Chartis to provide thought leadership on industry trends in the form of in-person speeches and webinars, as well as custom research and thought-leadership reports. Target audiences and objectives range from internal teams to customer and user conferences. Some recent examples include:

- Participation on a 'Panel of Experts' at a global user conference for a leading Global ERM (Enterprise Risk Management) software vendor.

- Custom research and thought-leadership paper on Basel 3 and implications for risk technology.

- Webinar on Financial Crime Risk Management.

- Internal education of sales team on key regulatory and business trends and engaging C-level decision makers.

© Copyright Chartis Research Ltd 2018. All Rights Reserved

8. Further reading

- *RiskTech100® 2018*
- *Spotlight on Risk as a Service*
- *Financial Crime Risk Management Systems: Market Update 2017*
- *Spotlight: quantifying cyber risk in financial institutions*
- *Data Integrity and Control Solutions in Financial Services 2016*

For all these reports see ***www.chartis-research.com***.

© Copyright Chartis Research Ltd 2018. All Rights Reserved