Milliman Data Science Survey

Eamonn Phelan, FSAI, FIA, CERA
Karl Murray, FSAI
Gillian Tucker, FSAI
Eamon Comerford, FSAI, FIA, CERA

Milliman

Life and health insurers and reinsurers are taking a growing interest in Data Science. Here we discuss the results of a recent and wide-ranging survey of activity in this area.

Data Science is a term sometimes given to the broad array of activities which are now undertaken in relation to extracting value from data, including techniques such as data analytics, predictive analytics, machine learning, data mining and artificial intelligence. It is certainly not a new discipline, but its widespread application to insurance is relatively new, particularly its application in the field of life and health insurance.

In this note, we share the outcome of a recent survey, conducted amongst Irish-based life and health insurers and reinsurers, looking specifically at current and planned future applications of data science to life and health insurance business. A total of 22 companies participated, operating in both the domestic and cross-border markets. It provides us with a comprehensive snapshot of activity in the life and health insurance market in Ireland today.

Our survey focusses on overall data science strategy, data collection, process and technical application, resourcing and governance, as well as benefits and challenges, and has yielded some very interesting results. While current levels of activity are still relatively low, this is about to change, with companies planning to use data science in lots of different ways.

Key Findings

- Over 75% expect to be using data science within the next 3 years, with over 35% already making it a point of focus.
- Most common uses of data science right now involve either assessment of customer behaviour or assumption setting.
- There has been limited use of external datasets so far.
- Actuaries and risk roles are currently most heavily involved in applications.
- There is limited standardisation thus far around collection and use of data.
- Data science is seen as a way to deliver major benefits in increasing customer engagement, increasing customer satisfaction and retention, increasing sales, improving fraud detection and achieving process efficiency.
- The biggest challenges facing companies involve a lack of infrastructure and technology, cyber risks, regulatory expectations, a shortage of talent, data quality, and access to data.

Scope and Strategy

In the first part of our survey we asked participants about their current strategy on data science and its areas of application.

QUESTION 1: HOW WOULD YOU DESCRIBE THE OVERALL STRATEGY FOR DATA SCIENCE IN YOUR ORGANISATION?
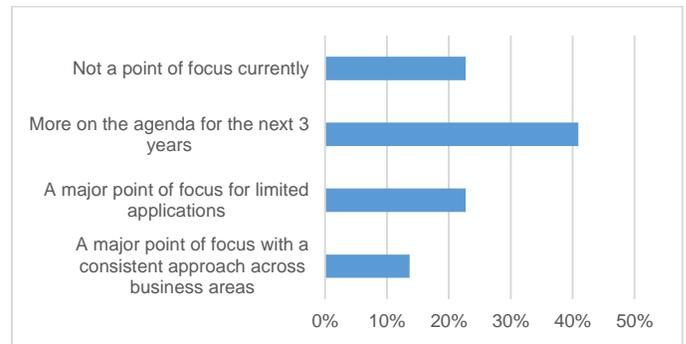
Over 35% of participants in the survey are currently active in data science with just over 40% reporting that data science is on the agenda for the next three years. This indicates that data science within the life and health insurance sector is beginning to grow and we expect to see a lot of activity over the next three years as more companies turn their focus to it.

QUESTION 2: FOR WHAT BUSINESS DECISIONS OR APPLICATIONS IS DATA SCIENCE CURRENTLY USED AT YOUR COMPANY?

Overall, there is a broad range of applications for which data science is currently used. The most popular application of data science amongst participants in our survey is in sales and marketing activities, in particular in understanding customer behaviour. Another popular use is for actuarial assumption setting purposes.

In the "Other" category, participants mentioned the use of data science for monitoring distribution channels and optimising technical provisions. One company indicated their participation in innovation labs external to the company as a form of knowledge transfer.

Data Usage

Next, we asked participants about their approach to sourcing and managing data for data science applications.

QUESTION 3: DOES YOUR ORGANISATION HAVE A DEDICATED DATA ARCHITECTURE/INFRASTRUCTURE FOR DATA SCIENCE?
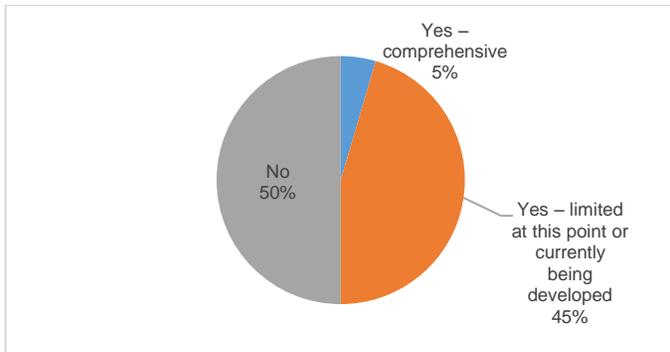

Half of the participants in the survey have some kind of dedicated data architecture/infrastructure for data science, though for most of these it is limited at this point or currently being developed. We expect to see a lot of development in the near future.

QUESTION 4: HOW WOULD YOU DESCRIBE YOUR CURRENT ACTIVITIES RELATING TO SOURCING AND ACCUMULATING DATA FOR DATA SCIENCE APPLICATIONS?
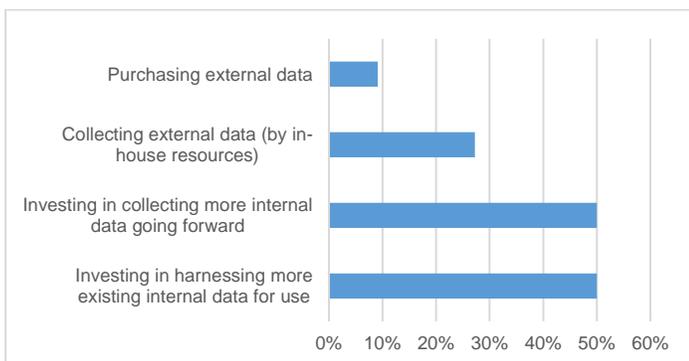

We expect to see a lot of investment in data sourcing going forward, with half of the companies surveyed intent on investing in collecting and harnessing more internal data. There is currently a limited use of external data among the survey participants. We expect companies to focus more on collection and use of external data to supplement their internal data as their data science activities continue to develop over time.

QUESTION 5: WHICH OF THE FOLLOWING SOURCES OR METHODS HAVE YOU USED (OR PLAN TO USE IN THE NEXT 3 YEARS) TO CAPTURE DATA FOR ONWARDS DATA SCIENCE PROCESSING?
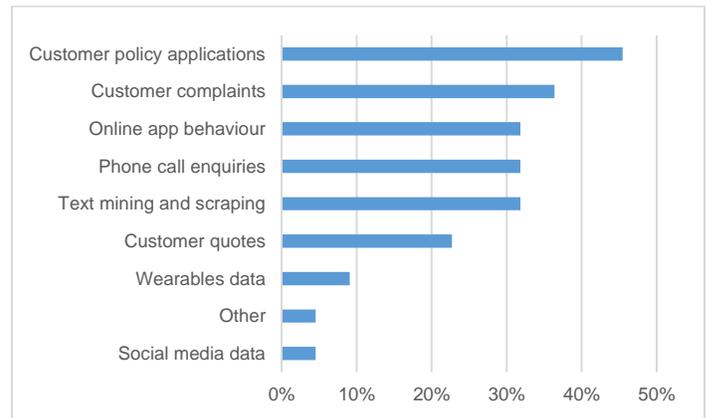
A broad range of sources have been used (or are planned to be used) to capture data for data science purposes. The most popular source amongst our participants is customer policy applications. Other popular sources include customer complaints, online app behaviour, phone call enquiries, and text mining and scraping.

QUESTION 6: DOES YOUR ORGANISATION HAVE AN APPROACH TO SATISFYING GDPR REQUIREMENTS SPECIFICALLY ADAPTED TO DATA SCIENCE ACTIVITIES?
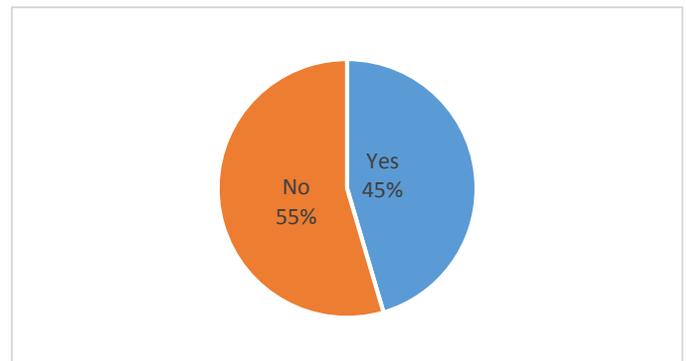

45% of participants in the survey currently have an approach to satisfying GDPR requirements specifically adapted to data science activities. For most companies that do not have such an approach, data science is not currently a point of focus or is more on the agenda for the next three years. We expect these

companies will, in due course, develop an approach to satisfying GDPR that is specifically adapted to data science activities.

Data Science
# Process, Architecture and Tools

Participants were then asked to provide information about some more technical aspects of data science in action.

QUESTION 7: WHICH OF THE FOLLOWING TYPES OF TOOLS OR TECHNIQUES HAVE YOU USED (OR PLAN TO USE IN THE NEXT 3 YEARS) IN THE APPLICATION OF DATA SCIENCE?
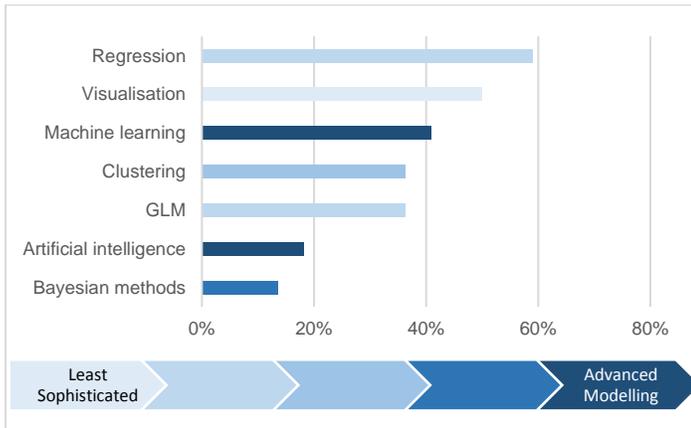
The tools/techniques in the above chart have been colour-coded according to level of modelling sophistication. The least sophisticated techniques, such as regression and visualisation, are currently the most popular approaches being used among the survey participants. While more advanced techniques, such as Artificial Intelligence and Bayesian Methods, are currently less popular but may grow in popularity as the use of data science develops.

QUESTION 8: HAVE YOU USED ANY OF THE FOLLOWING SOFTWARE IN DEVELOPING DATA SCIENCE PROCESSES?
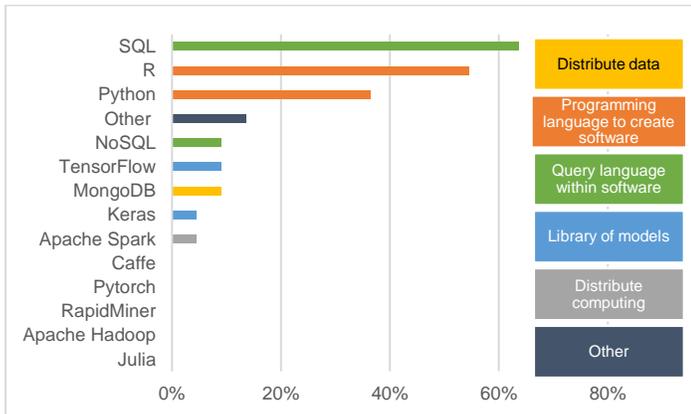

The top types of software being used to develop data science processes amongst the survey participants are query languages (SQL) and programming languages (R and Python).

Currently, software used to distribute data, distribute computing and house libraries of models are the least used.

In the "Other" category participants also mentioned their use of Microsoft Power BI and Tableau.

QUESTION 9: FOR THE FOLLOWING AREAS INDICATE THE LEVEL OF TRAINING OR UPSKILLING REQUIRED BY INDIVIDUALS IN YOUR ORGANISATION (5 BEING A HIGH AMOUNT OF TRAINING OR UPSKILLING)
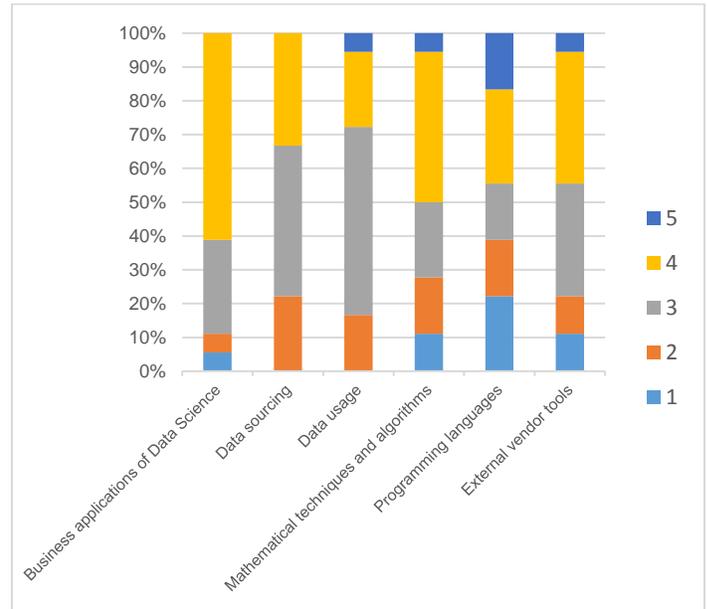

The average level of training and/or upskilling required is highest for business applications of data science, with over 60% of respondents to this question reporting a level 4 (on a scale of 1 to 5). For all areas, the average level of training and/or upskilling required is 3 or higher, so companies will need to start investing in training/upskilling in these areas soon.

Resourcing and Governance

We also asked participants about more practical aspects of data science applications, including their approach to resourcing and governance of projects.

QUESTION 10: TO WHAT DEGREE ARE THE FOLLOWING AREAS OF THE BUSINESS INVOLVED CURRENTLY WITH DATA SCIENCE APPLICATIONS?
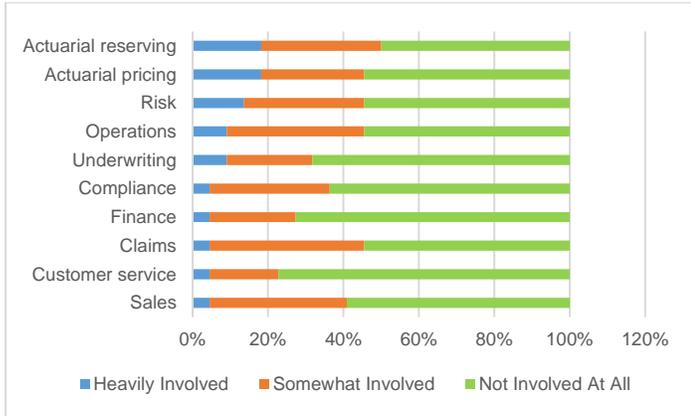
Actuarial and risk roles are the most heavily involved in data science applications, at the moment, amongst our survey participants. Claims and operations roles are also seeing some involvement. One participant also mentioned that their investment managers are involved with data science applications.

Given the answer to Question 2, we would also expect an increased involvement over time from customer service, underwriting and sales functions.

QUESTION 11: DOES YOUR ORGANISATION HAVE INTERNAL STANDARDS GOVERNING THE USE OF DATA SCIENCE?
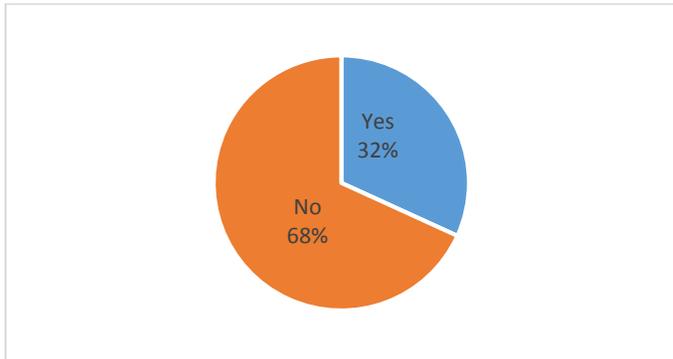
Given the relative newness of data science applications for life and health insurers and reinsurers, it is perhaps unsurprising that there has been limited standardisation around data science to date. Over time, we would expect to see internal best practice frameworks emerging for data science processes.

QUESTION 12: DOES YOUR ORGANISATION HAVE A TEAM DEDICATED TO DATA SCIENCE OR ARE RESOURCES EMPLOYED ACROSS MULTIPLE TEAMS?

In answer to this question, only a quarter of participants said they have a team dedicated to data science whilst other participants said that either their data science resources are spread across existing functions or they have not carried out data science projects to date.

QUESTION 13: APPROXIMATELY HOW MANY FULLTIME EMPLOYEES WOULD YOU ESTIMATE ARE CURRENTLY DEDICATED TO DATA SCIENCE ACTIVITY WITHIN YOUR ORGANISATION?

On average, two fulltime employees (FTEs) per company are currently dedicated to data science activity across the 22 survey participants.

In companies which have a team dedicated to data science the average number of FTEs is closer to four.

One company stated that they have up to five FTEs involved in data science right now.

QUESTION 14: DO YOU OUTSOURCE ACTIVITIES IN RELATION TO DATA SCIENCE?
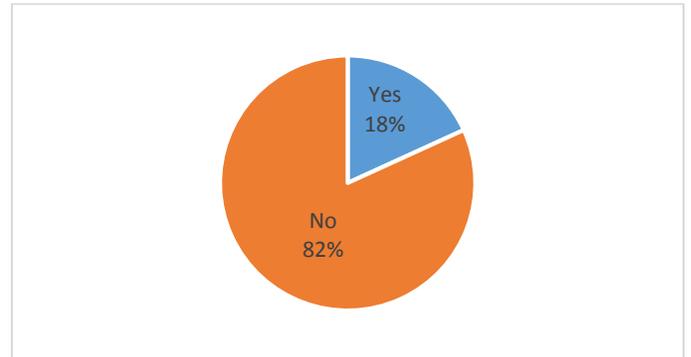
Our participants indicated quite a limited use of outsourcing to date in relation to data science. This likely reflects the reliance, to a large extent, on internal data at present. As firms come to use more external data, and require expert input, we expect to see more use of third-party advisor support.

Furthermore, as developing in-house capabilities and a dedicated data infrastructure/architecture can be time-consuming and expensive, there may be benefits to be gained through external support.

QUESTION 15: TO WHAT EXTENT ARE THE ACTUARIAL, RISK, COMPLIANCE AND INTERNAL AUDIT FUNCTIONS INVOLVED IN ENSURING THE ACCURACY/QUALITY OF THE DATA AND DESIGN OF YOUR DATA SCIENCE PROCESSES?

Actuaries are most involved in oversight around the area of data science, followed by other risk professionals. Participants indicated very limited involvement from compliance and internal audit at the moment.

Data Science Benefits and Challenges

In the final part of our survey we asked participants to comment on the benefits and challenges related to the use of data science.

QUESTION 16: WHAT ARE THE MAIN POTENTIAL BENEFITS ARISING FROM THE USE OF DATA SCIENCE FOR YOUR INSURANCE BUSINESS?
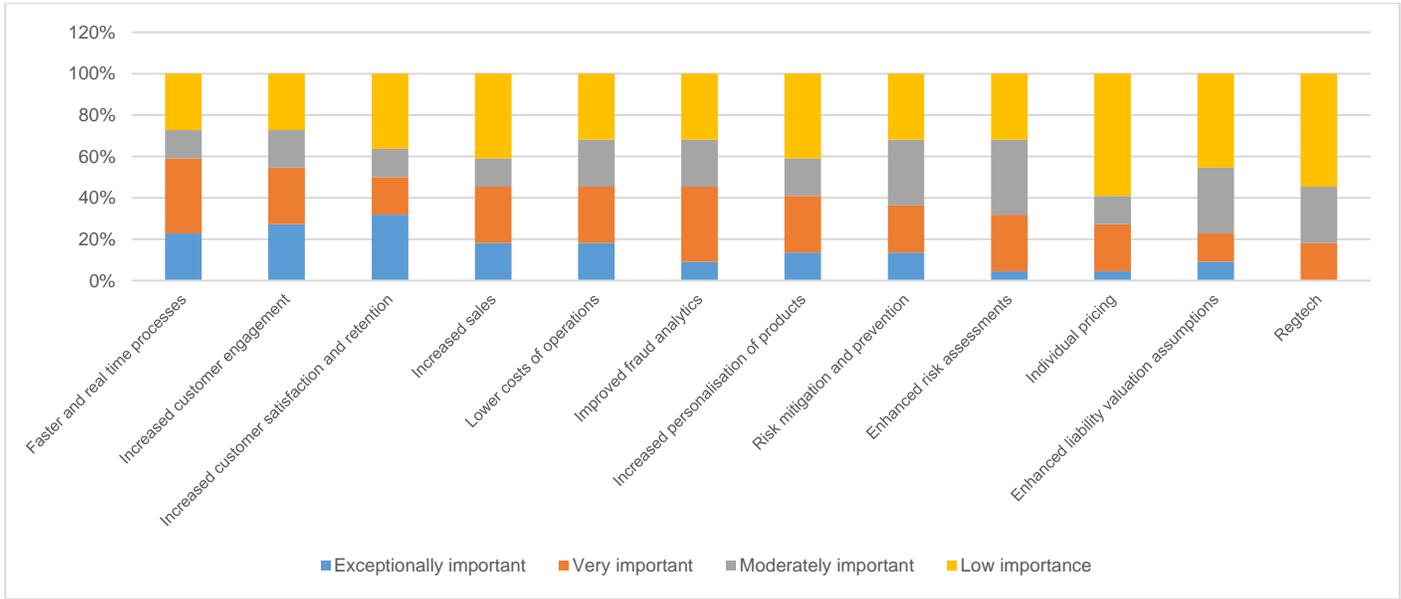
The most important benefits cited by participants included faster processes, increased customer engagement, increased customer satisfaction and retention, increased sales, lower costs and better fraud detection. Improved quality of sales was also highlighted as a potential benefit. What is clear from these responses is that there is a broad range of areas in which companies expect to be able to deploy data science techniques in order to drive the business forward, and we expect to see this continue to expand as more and more companies start to actively engage.

QUESTION 17: HOW RELEVANT ARE THE FOLLOWING CHALLENGES FOR YOUR ORGANISATION?
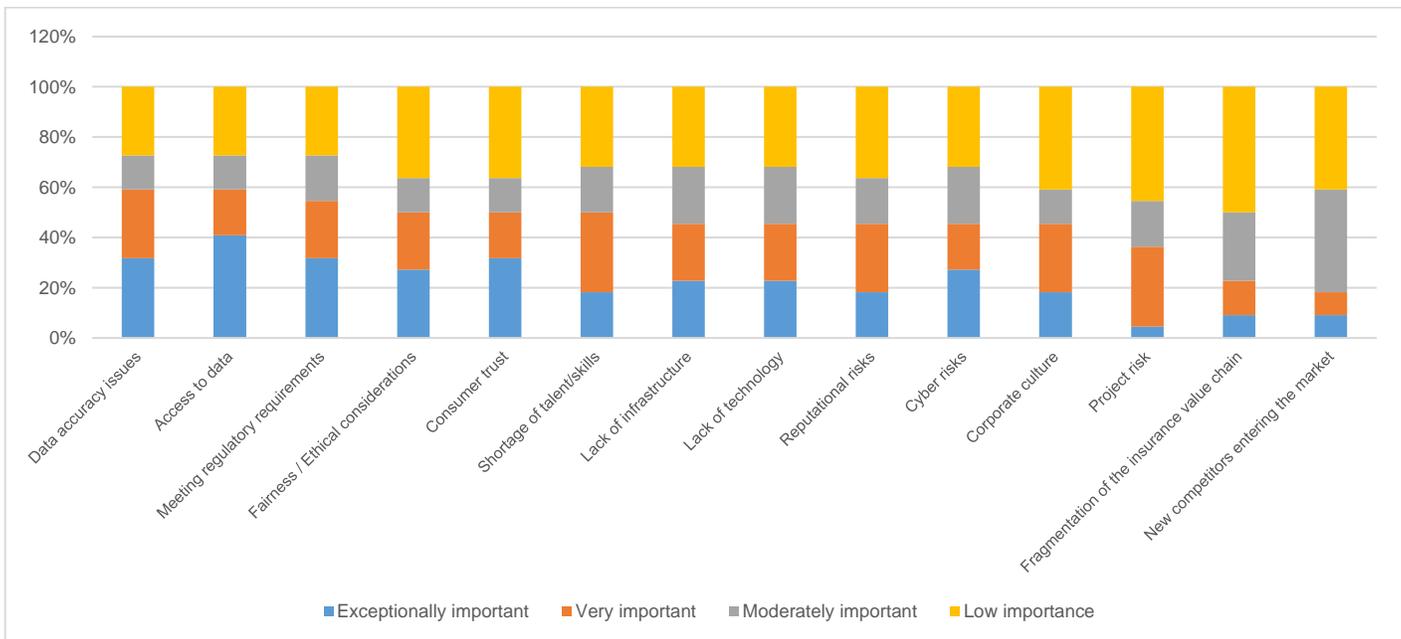
The biggest challenges facing participants involve data accuracy and access to data, meeting regulatory expectations, fairness and ethical considerations, consumer trust, a shortage of talent, a lack of infrastructure and technology.

The landscape of key benefits and challenges to companies using data science techniques will evolve over the coming years, as more and more firms actively engage, as this is when these will become more tangible for them.

How Milliman Can Help

At Milliman, we have been actively working with our clients for many years to effectively harness the power of data science in order to help meet their business needs.

We can assist you with all aspects of your data science initiative including providing advice on:

- Best practice frameworks for data science processes
- Collection and processing of data
- Identifying applications for data science techniques
- Identifying suitable tools and techniques for particular circumstances
- Implementing solutions
- Understanding the implications of results
- Constraints and practical challenges

For further information please contact your usual Milliman consultant.

EUROPEAN MILLIMAN LOCATIONS

CONTACT

Eamonn Phelan
eamonn.phelan@milliman.com

Karl Murray
karl.murray@milliman.com

Gillian Tucker
gillian.tucker@milliman.com

Eamon Comerford
eamon.comerford@milliman.com

Milliman is among the world's largest providers of actuarial and related products and services. The firm has consulting practices in life insurance and financial services, property & casualty insurance, healthcare, and employee benefits. Founded in 1947, Milliman is an independent firm with offices in major cities around the globe.

milliman.ie