

# Milliman Data Science Survey for Non-Life Insurance

Eoin Ó Baoighill, FSAI  
Gillian Tucker, FSAI  
Divey Hiranandani  
Gavin Maher



Non-life (re)insurers are continuing to take a growing interest in data science. Here we discuss the results of a recent 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 modelling, machine learning, data mining and artificial intelligence. It is certainly not a new discipline, but its widespread application to insurance is relatively new and it is one of the key challenges that Milliman believes will bring new perspectives to insurance companies

In this note, we share the outcome of a recent survey, conducted amongst Irish-based non-life insurers and reinsurers, looking specifically at current and planned future applications of data science to non-life insurance business. The survey included domestic and cross-border companies and some Managing General Agents (“MGAs”). It provides us with an indicative snapshot of activity in the non-life insurance market in Ireland today.

Our survey focusses on overall data science strategies, data collection, processes and technical applications, 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.

Throughout this briefing note we have drawn comparisons to a similar data science survey carried out on Irish-based life and health insurers and reinsurers operating in both the domestic and cross-border markets. The results of this survey were published in a briefing note released in 2019<sup>1</sup>.

## Key Findings

- Over 83% of companies surveyed expect to be using data science within the next 3 years, with 50% already making it a point of focus.
- Most common uses of data science right now is for pricing and sales and marketing activities.
- There has been some use of external datasets so far. 60% of those surveyed collect external data with internal

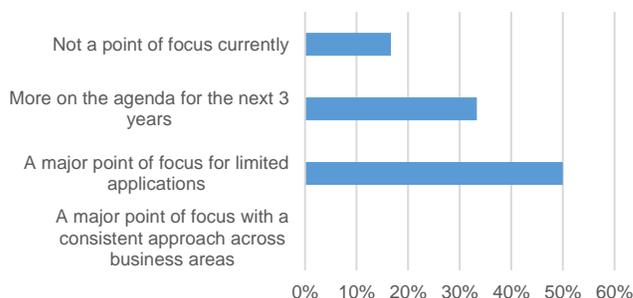
resources while 33% purchase external data. This use of external data is much greater than that of the participants in our life and health survey, who had equivalent figures of 27% and 9% respectively.

- 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 risk mitigation and prevention, enhancing risk assessments, improving pricing, creating faster and real time processes and lowering the cost of operation via automation of processes.
- The biggest challenges facing companies involve consumer trust, corporate culture, 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?

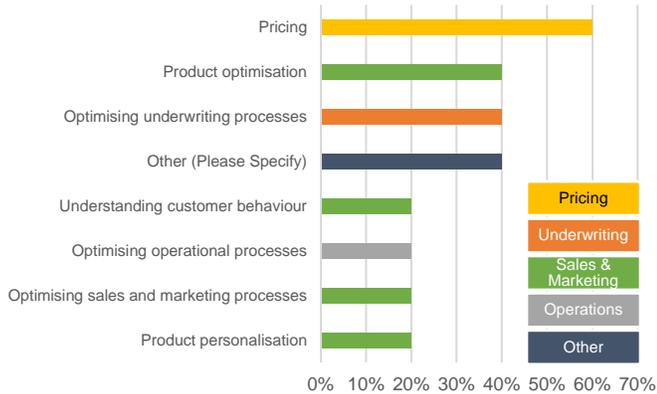


50% of participants in the survey are currently active in data science with over 30% reporting that data science is on the agenda for the next three years. This indicates that data science within the non-life insurance sector is growing and we expect to see more activity over the next three years as more companies turn their focus to it. The non-life sector seems slightly ahead of the life and health sectors. In our life and health survey, just over

<sup>1</sup> <http://ie.milliman.com/insight/2019/Milliman-Data-Science-Survey/>

35% of participants were currently active in data science with 40% reporting that it is on the agenda for the next three years.

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



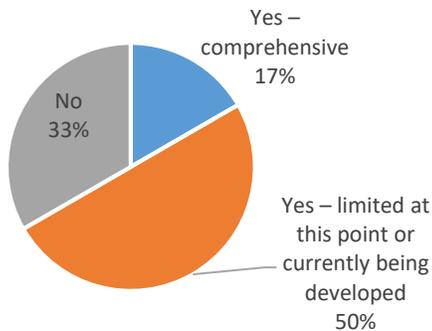
Overall, there is a fairly broad range of applications for which data science is currently used, but there are many other possible applications for which it is not being used – the absence of the use for fraud detection among our survey participants is particularly surprising. The most popular applications of data science amongst participants in our survey is in pricing and sales and marketing activities.

Those participants who answered “Other” do not yet make much use of data science, but expect to over the coming years.

**Data Usage**

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

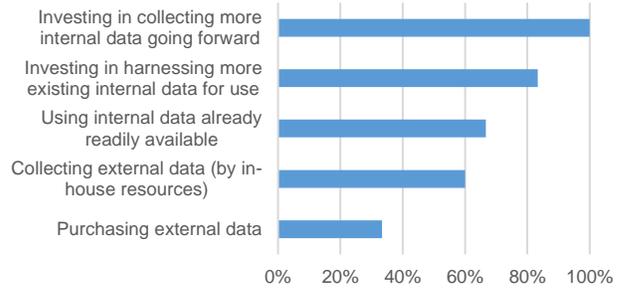
**QUESTION 3: DOES YOUR ORGANISATION HAVE A DEDICATED DATA ARCHITECTURE/INFRASTRUCTURE FOR DATA SCIENCE?**



Over two thirds 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.

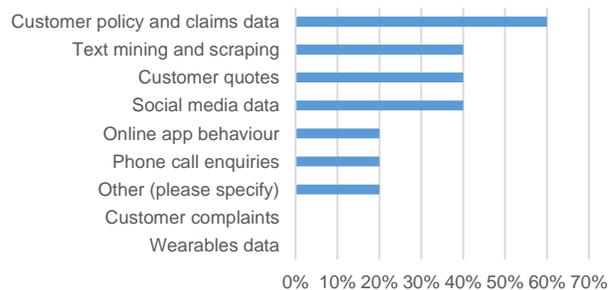
**QUESTION 4: HOW WOULD YOU DESCRIBE YOUR CURRENT ACTIVITIES RELATING TO SOURCING AND ACCUMULATING DATA FOR DATA SCIENCE APPLICATIONS?**



All companies surveyed are investing in collecting more internal data and most companies are investing in harnessing more existing internal data. There seems to be a clear trend that companies are aiming to make sense of their own data and to extract value from it.

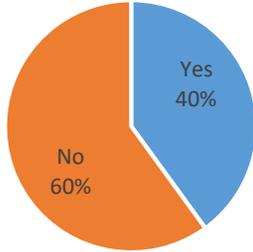
External data is currently being sourced to a lesser extent., but is still being sourced by a higher proportion compared to the participants in our life and health survey (60% of non-life companies surveyed are collecting external data, compared to 27% of life and health companies surveyed. 33% of non-life companies surveyed are purchasing external data, compared to 9% of life and health companies surveyed). 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 and claims data. Other popular sources include customer quotes, social media data and text mining and scraping.

**QUESTION 6: DOES YOUR ORGANISATION HAVE AN APPROACH TO SATISFYING GDPR (“GENERAL DATA PROTECTION REGULATION”) REQUIREMENTS SPECIFICALLY ADAPTED TO DATA SCIENCE ACTIVITIES?**

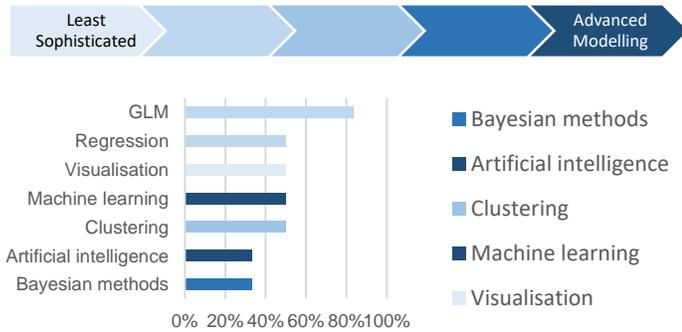


40% of participants in the survey currently have an approach to satisfying GDPR requirements specifically adapted to data science activities. For most companies that do not have such an approach, data science is not currently a point of focus or is more on the agenda for the next three years. We expect 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**

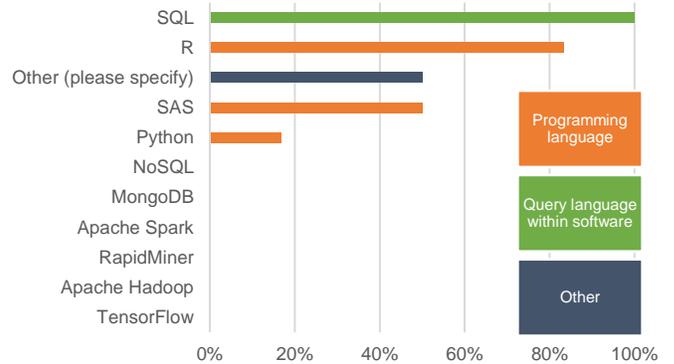
Participants were then asked to provide information about some more technical aspects of data science in action.

**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. Generalised Linear Models (“GLMs”) are the most popular technique used by the survey participants, and are also one of the least sophisticated techniques. 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?**

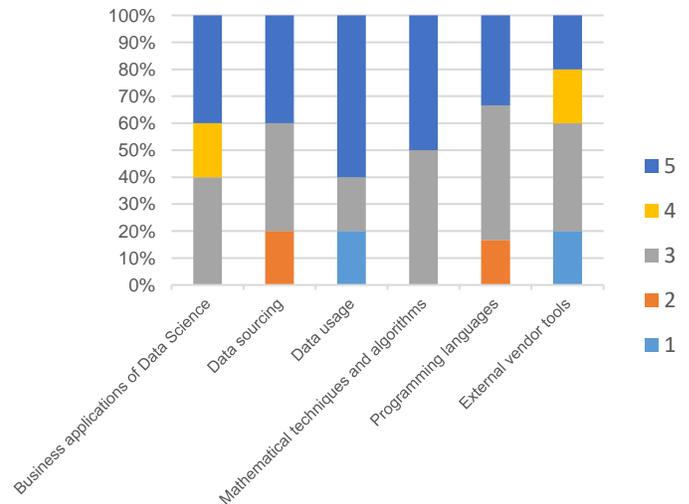


The top types of software being used to develop data science processes amongst the survey participants are query languages (SQL) and programming languages (R, Python and SAS). We may see an increase in popularity for some of the open source/free platforms such as R and Python as opposed to the more expensive platforms such as SAS over the coming years.

Currently, software used to distribute data and distribute computing is not being used by the participants in the survey.

In the “Other” category participants also mentioned their use of a custom-built system.

**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)**



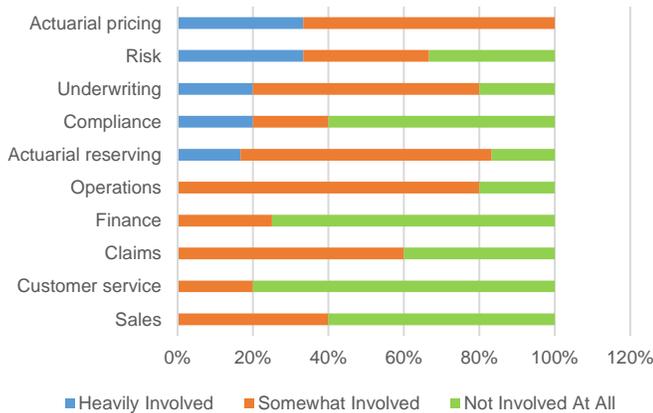
The average level of training and/or upskilling required is highest for business applications of data science and mathematical techniques and algorithms. For all areas, the average level of training and/or upskilling required is 3 or higher, so companies

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. Underwriting and compliance roles are also seeing some involvement.

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

### QUESTION 11: DOES YOUR ORGANISATION HAVE INTERNAL STANDARDS GOVERNING THE USE OF DATA SCIENCE?

All participants in the survey answered “No” to this question.

Given the relative newness of data science applications for non-life 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.

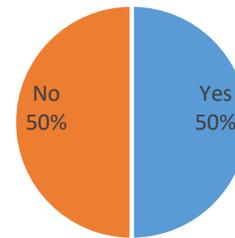
### QUESTION 12: DOES YOUR ORGANISATION HAVE A TEAM DEDICATED TO DATA SCIENCE OR ARE RESOURCES EMPLOYED ACROSS MULTIPLE TEAMS?

In answer to this question, all participants said that their data science resources are spread across multiple teams and only one of the participants said they also have a team dedicated to data science.

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

One company stated that they have up to twenty fulltime employees (“FTEs”) involved in data science right now. This participant had stated that their data science resources were spread across multiple teams. Among the other survey participants, the average number of FTEs is closer to one.

### QUESTION 14: DO YOU OUTSOURCE ACTIVITIES IN RELATION TO DATA SCIENCE?



Half of our participants indicated outsourcing in relation to data science. This is much higher than the participants in the life and health survey, of which 18% indicated outsourcing in relation to data science.

For the half of non-life participants not outsourcing, there is likely a reliance, to a large extent, on internal data at present.

### 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?

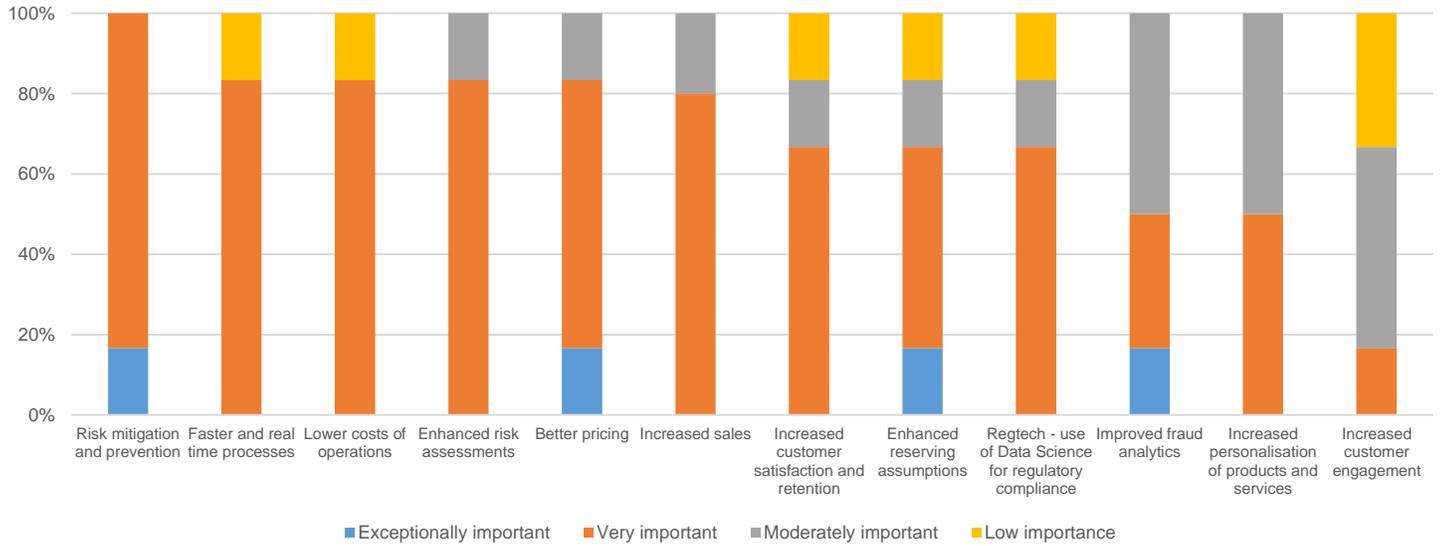
On average, risk management and internal audit are most involved in oversight around the area of data science.

Participants indicated lower involvement from the actuarial function, as opposed to the participants in our life and health survey where participants indicated actuaries were most involved. Limited involvement was also indicated for the compliance function, which was similar for the life and health survey.

## Data Science Benefits and Challenges

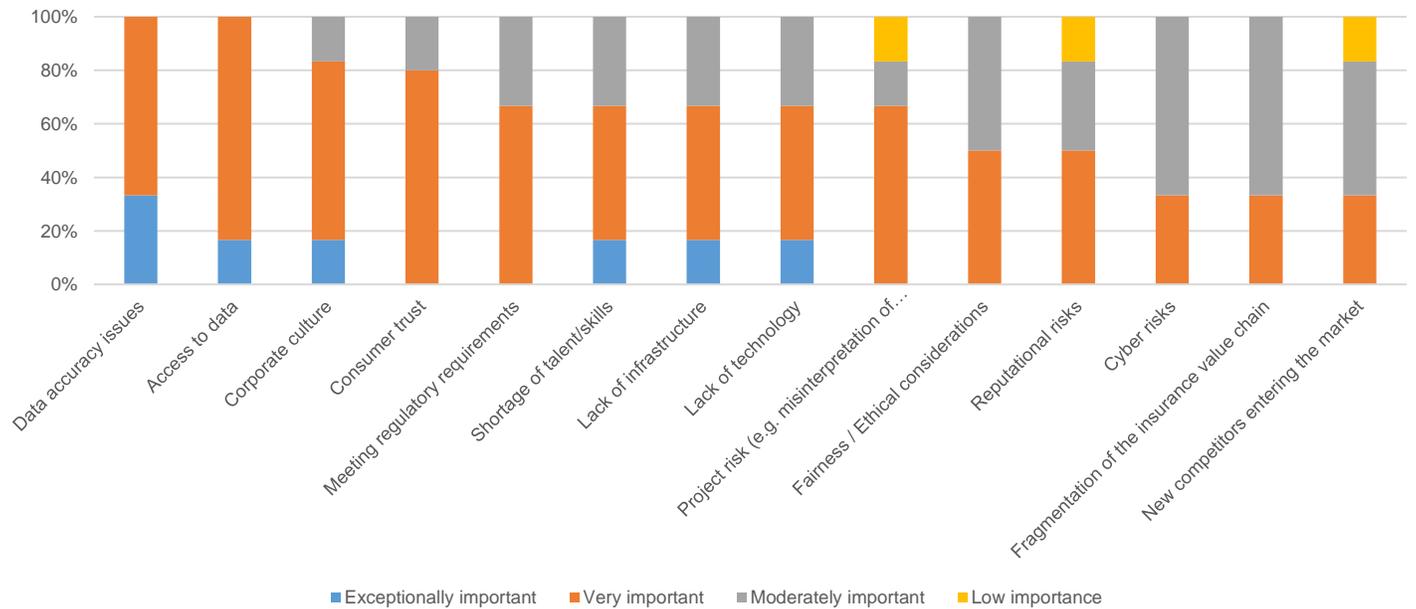
In the final part of our survey we asked participants to comment on the benefits and challenges related to the use of data science.

### 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 risk mitigation and prevention, enhancing risk assessments, improving pricing, creating faster and real time processes via automation, and lowering the cost of operation. Increased 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 consumer trust, corporate culture, data quality, and access to data.

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
- Data visualisation techniques
- Implementing solutions
- Automation of processes
- Understanding the implications of results
- Constraints and practical challenges
- Staff training in data science

For further information please contact your usual Milliman consultant.



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.

[ie.milliman.com](http://ie.milliman.com)

### CONTACT

Eoin Ó Baoighill  
[eoin.obaighill@milliman.com](mailto:eoin.obaighill@milliman.com)

Gillian Tucker  
[gillian.tucker@milliman.com](mailto:gillian.tucker@milliman.com)

Divey Haranandani  
[divey.haranandani@milliman.com](mailto:divey.haranandani@milliman.com)

Gavin Maher  
[gavin.maher@milliman.com](mailto:gavin.maher@milliman.com)