EVERYTHING YOU NEED TO KNOW ABOUT

DATA & AI

THE ULTIMATE EXECUTIVE BRIEFING

"Finally. A one-stop shop for all those tech terms you've been too scared to ask about. Grab Ian's advice with both hands."

Emma Barry
Global Fitness Authority

*

"A must-read, and not just for your IT department. Data is the key to the new customer experience paradigm and must be front and centre on every leader's desk."

Jon Brady
President, Midtown Athletic Clubs

*

"The data is out there to enrich member experiences, and with it our businesses. We simply need the ambition to dig deeper, and this paper shows us where to look."

Can Ikinci
CEO, Mars Sportif

*

"A clarion call to the industry to re-invent the customer experience, basing our conclusions no longer on intuition but on hard datasets."

Doron Dickman
CEO, Holmes Place Germany

Contents

Foreword

Data is boring? Wrong. Understanding it can be frustrating, but your data is the key to your business success. It is your past, your present and, by acting on the insights of this white paper, most importantly it is your future.

At a top level, of course, most of us know data is key to improving retention figures, sales forecasting and performance metrics. The issues come in the detail: in operators not knowing what's at their fingertips, or how to interpret it efficiently to extract the value.

The first step is to take the guesswork out of the numbers game. In this follow-up to his excellent first white paper – *The Fitness Future: Rules of Engagement* – Ian Mullane therefore now leads us on a journey through AI, machine learning and business intelligence in a way that actually makes sense, even to those of us who usually see data analysis as a tech session of buzzword bingo.

> ## "We are in the era of attention – and the currency is data"

What's the true importance of understanding data? Do I need any special equipment to data-mine? Exactly how big can data get? Is my data even good enough? In this white paper, you'll find the answers to these questions, and many more besides.

You'll learn the difference between AI and machine learning, and how you can now harness their power – without having to get involved in any of the maths to baffle even the brightest of brains – and render those late nights of poring over financial documents, forever looking for fields of green, a thing of the past.

You'll learn how all data is not equal, and how so much potentially valuable data – from tapped-out text forms to emoji-based feedback to CCTV – is hiding within your business (all this before you even start on data showing what members do outside of your club).

You'll explore how partnerships can further magnify the value of your data. And you'll learn how all of this translates into real-world results – improved member lifetime value, lead conversions, non-dues spend and referrals just a few examples – as you harness the power of the information you never even knew you had to better understand and serve your community.

Whether you're taking the first steps to understand AI, machine learning and harnessing your most important asset, or you're taking internal insights and external thought leadership to the next level, we are in the era of attention – and the currency is data.

By reading the following pages, you'll better understand why you can no longer afford to lose track of yours – and how to ensure that doesn't happen.

David Stalker
President, EuropeActive

Introduction

Data. It's never been seen as the sexiest of topics – but all that's changing, and fast.

Buzzwords such as 'Big Data' and 'AI' now populate many a boardroom conversation, bringing with them an exciting vision: that of a super-brain capable of analysing every piece of information and knowledge in a business, then using it to improve performance.

It's enough to make anyone giddy with excitement, which explains why data is quickly emerging from the shadows, no longer seen as a dark art best left to the geeks but rather a mainstream topic that most people – and particularly senior management – know they need, hopefully even actively want, to dive into.

The challenge, however, is this. For so long a back-room discipline, across all levels of an organisation most people's understanding of data – collection, analysis, usage – will have gaps. The overall vision fills us with excitement, but we're not necessarily equipped to answer important questions such as: how significant are the latest developments to fitness operators, and how relevant to our business specifically? What value can I expect to create? What do I need to succeed?

> "Most people's understanding of data has gaps. The overall vision fills us with extra excitement, but we're not necessarily equipped to answer important questions."

That's why we've produced this white paper: to provide you with the information you need to truly understand what data is, in what forms you will find it, and how to drive exceptional business value from it.

I hope to strip away any barriers of complexity that hinder understanding, leaving you with a framework that allows you to quickly evaluate whether you have the data needed for a task, and whether you should bring in AI.

You'll come to understand the types of data you will find in your business, and how to work with them, along with tips on how to collect and store this valuable asset. We will then look at how to leverage your data, using AI and machine learning to produce predictions and insights that would not be possible through human analysis.

I intend to do all of this without any reference to maths or complicated technical concepts. Similar to 'extracting the value' from a car – something that requires zero understanding of the intricacies of engine design – so data + AI will increasingly become tools that require you to do no more than add fuel (data), get into the driving seat and let the AI steer you towards your desired end result.

If I achieve my aim, you, the reader, will come away understanding what the opportunity is for your business, along with actions you can take immediately to start extracting value.
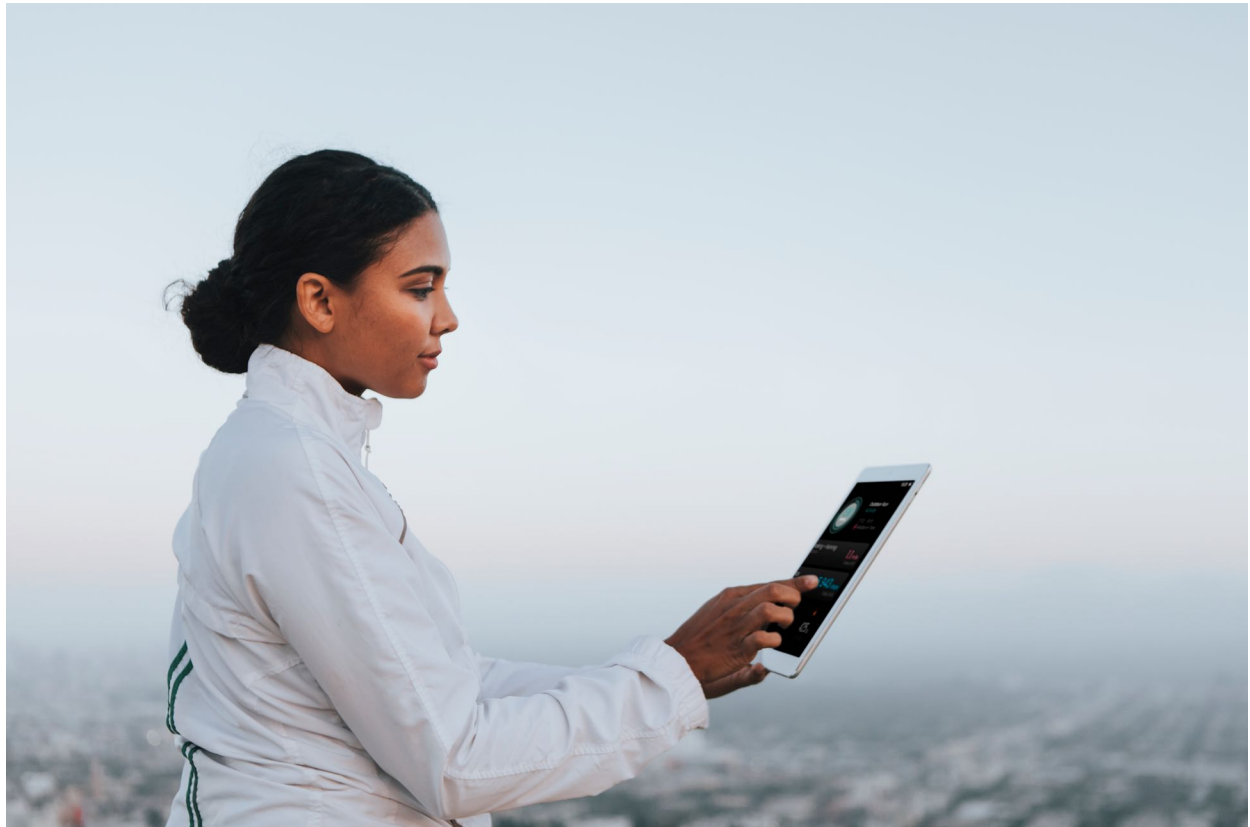
What is data, really?

The world is data-rich: market research firm IDC reported that, in 2018, the world generated 33 zettabytes of data, which is enough to fill seven trillion DVDs. That number will only rise, and it's a trend that's reflected in the fitness sector: we already have enormous quantities of data in our businesses, and the volume doubles every 18 months.

This proliferation of data underpins the existence of AI, too. Without data, there would quite simply be no AI – that invisible presence that facilitates our everyday life, from giving us a drive time and ETA at our destination, to recommending an artist we might like on Spotify.

It's important to note, however, that not all data is equal. Not all sources of data are 'insight gold'.

Let's take a look at what data actually is, and where you're likely to find it in your business.

What counts as data?

Over the years, data collection in the fitness sector has been undertaken to fulfil a requirement. We collect member details to set up their membership; we take their payment details to ensure we can charge them; we collect member surveys to understand performance. Each collection serves a specific purpose, an end goal.

Data is not just digital, either. It is not just the information that's stored within your CRM or club management system. Data can also be offline; in many fitness operations, it remains so in the form of member contracts, PT evaluations and liability waiver forms. While increasingly, most organisations will have this paper-based data converted to bits and bytes, inputting it into a digital tool, there's often data that exists only in paper form that does not make a transition.

Sometimes that's OK. Let's take statutory data collection such as liability waiver forms as an example: data we have to capture and hold, but which has little value beyond its purpose.

However, this is not always the case. I've seen plenty of rich data capture in both sales processes and PT interactions that doesn't move from the staff member's notepad, even though this data would be transformational in the delivery of member experience. More on that later.

Structured data

If we drill down beneath these specific data sources, we'll find two broad types of data existing within your business: structured and unstructured. Both can have value, and both deserve understanding.

Examples of structured data would be spreadsheet files and SQL databases, which is the data model most often used to power tools such as your club management system. Each data item fits into columns and rows, complying with a predefined data model that gives order to the items so they can be easily sorted and searched.

Structured data is where we find the most value at present. It is here that the tools we currently use are applied to produce results within our clubs. By allowing for every field to be accessed separately or together, this data model provides a powerful platform to quickly bring together information from various locations for analysis.

SQL: A common language

Structured Query Language (SQL) is a term you will often hear when data is discussed, as it's the price of entry to working with, and extracting value from, your databases.

In its simplest terms, SQL is the language spoken to share and manage data found in relational database management systems such as your CMS – that is, the backbone of the tools you use in the business every day. Indeed, while there are hundreds of different databases out there, the vast majority use SQL as their language. I have yet to find any non-SQL databases among the modern tools used by fitness operators around the world.

SQL presents us with a common language we can use across our tools to question, re-organise and change the data we have. It is SQL that is used, for example, if we need to extract a certain set of data from a club management database.

I promised we would not get technical here, but let me use an elementary example of a SQL statement to demonstrate its purpose.

Let's say we wanted to build a spreadsheet of current members from our member database.

```
SELECT Name FROM Members WHERE status = current |
```

Typing this in would see the names of all current members – and only current members – being retrieved from the Members table.

Such SQL work is fairly simply done; there's no need for a data scientist. SQL engineers are plentiful, too, as SQL is often the first skill a developer will learn. If you don't have anyone with this skillset in-house, freelancers can easily can be hired for jobs via contracting sites such as Elance.

Alternatively, for those inclined to develop the skills themselves, extensive resources are available on the internet to learn SQL, most of which are free of charge.
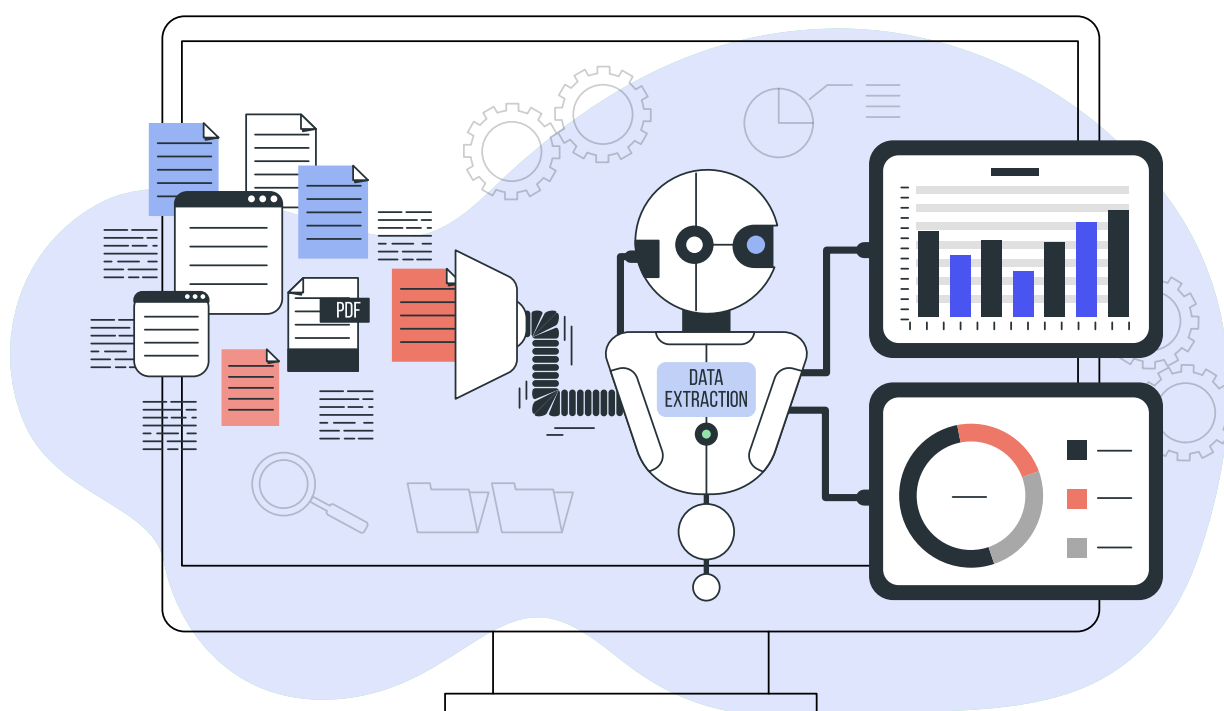
Unstructured data

Unstructured data accounts for more than 80 per cent of the data in any organisation and grows at 55-65 per cent per annum, with examples including social media posts, email, phone recordings and media files.

This is data that does not have a recognisable structure. Take email, for example. It has some structure when you consider that it's organised by date, sender and so on, but the primary 'data' in the email body is the text, and this doesn't follow a format. It is therefore considered unstructured.

> "It is when we add structured and unstructured data together that we get to that hallowed entity: Big Data"

While structured data has been the primary focus of most organisations' efforts to build business intelligence and gain actionable insights, unstructured data accounts for too large a proportion of an organisation's total data to be ignored. Indeed, it is when we add structured and unstructured data together that we get to that hallowed entity: Big Data.

Excitingly, it is in the area of unstructured data that we've seen vast leaps forward over the past decade, allowing us to gain far more value from the data in our businesses. Specifically, Artificial Intelligence (AI) algorithms can extract value from these piles of unstructured data, and this marriage can power a veritable treasure trove of insights around customer behaviours and predicted behaviour.

💡 Structured vs Unstructured

To elaborate on the difference between structured and unstructured data, let's take a look inside a typical fitness operation and explore the data pools – those with an external viewpoint, not internal such as HR or payroll – that we might expect to find.

Structured	Unstructured
Club Management System	Social media
Access control	Customer surveys
Member mobile app	Internet reviews
Website visitor analytics	CCTV
Class bookings	Member inbound email comms
Membership sign-up	Member inbound SMS comms
Fee payment	Webchat
Exercise equipment	Member assessment/induction forms
Point of sale	Telephone call recording
Member wearables data	Promotional videos

In-between data

I don't think it will come as a surprise to know that we can also have semi-structured data; it isn't simply black vs white, structured vs unstructured.

One great example of in-between data – and something AI is giving new life to – are customer surveys such as the Net Promoter Score©.

The structured element comes from the primary purpose of the survey which, in many cases, uses 'on a scale of 1–10' type questions. This score forms the majority of the analysis and is generally the one followed by business leadership.

However, in my opinion there's more value to be extracted from the unstructured aspect – that is, when we ask the customer: "Why did you give us that score?" This free-form text is gold as it provides context, detail and often insights that wouldn't have been attributed to the 1–10 score.

The reality, though, is that this unstructured data is rarely leveraged as much as it could, or arguably should, be. Why? Because if you send out surveys throughout the year, it's possible that around 30 per cent of your members will respond. If you have 10,000 members, the easy part is collating the 3,000 scores – but what do you do with 3,000 free-form text submissions?

If you're the unlucky intern, you may be asked to go through them all and pick out themes and points of note. In reality, however, this is rarely done; if it is, it's done poorly and loaded with human bias to the desired outcome.

With AI, on the other hand, every comment that comes back to you can be passed through language sentiment analysis. Regardless of what language is used (including Emoji which, go figure, is viewed as a language in itself), each comment would be assessed and given a positive, negative or neutral score.

Easy enough, you say: the intern may not like it, but they could still do it. And yes, they could. But even if we ignore the fact that every person comes with a built-in bias around what they want to see, or how they feel on the day of the task, how would they deal with the following?

"Great staff, great location, but the changing rooms require restoration."

> ## "The simple customer survey becomes a turbo-charged tool that shows which parts of the business to emphasise to which target audiences"

Overall a positive? Does the negative mean it's neutral? I would argue the value is actually in looking at this differently. By all means apply an overall sentiment, but I would want to see this broken down, with the comment also scored across three distinct areas: staff, location, and changing rooms. That way, I would be able to look up core aspects of the business, see how they're perceived by members, and understand the overall sentiment of the feedback for that specific area of the business.

When you also apply age and gender filters, you can then paint a picture of what's important to each group, how they perceive certain factors in the business, and how they describe them.

The simple customer survey becomes a turbo-charged tool that not only gives you a headline score on how you're doing with customers, but also goes beyond to show where you're powerful and where not so strong, along with additional pointers on which parts of the business to emphasise to which target audiences.

Data gifts from members

Utopia for a fitness operator is when they can gain an insight into the member's activities away from the club. There's a world of difference between a member who's stopped coming because they've decided the gym isn't for them, and a member who's taking advantage of the good weather to do park runs or bike rides.

Then there's the difference in personalisation you can give a member when – just as an example – you know their current fitness level or average cardiovascular load. The good news is that this valuable data is being collected; the challenge is that our members are managing it. Without them sharing it, we're missing out on possibly the most useful data of all.

> "Valuable data is being collected; the challenge is that our members are managing it. We're missing out on possibly the most useful data of all."

Apple Health, Google Fit, Apple Watch, Whoop, Myzone... all of these are sources of member data. These tools are increasingly not just for on-site use, but also worn 24/7 to help guide and inform their user on their current condition. This is incredibly rich data that could power everything from retention algorithms to recommendation engines, providing hyperpersonalised engagements for your customers.

When the technology is integrated within your business, such as Myzone, access is already easier to obtain. Moving forward, we can expect integration to become commonplace across the board, with data consolidators and tools providing the 'pipework' for you to gain customer data from their chosen apps (with their permission, of course) so you can deliver more personalised services to them.

Now is the time to start understanding who's using what, and why they choose to use it.

Public data

The quantity of public data available varies from country to country, and its usage even more so. However, where permitted, the integration of public data with the data within your business can produce a rich seam of value.

The US is a market that's awash with public data – and little control around how that data can be used. From salaries to purchases, political affiliations to hobbies, all of this and more is available and used daily by marketing automation and ad platforms.

An excellent example from a fitness perspective is a service I've witnessed that powers member lead generation. Merging public data with your ideal member profile, they can 'find' ideal potential members for you to target with your marketing.

Join the dots

You can no doubt think of many more forms of data, but even just from those outlined here, it's clear there's a mass of data in every one of our businesses: present in various forms; sometimes connected but most often not; some automatically generated, others customer-initiated.

> "When one set of data is connected with other data sets or tools, the value of the data multiplies"

When viewed as separate entities, separate 'islands' of data, we can understand their independent values to the business.

But when one set of data is connected with other data sets or tools – as an example, combining the member list from your club management system with your customer surveys – the value of the data multiplies. It may have been collected for one purpose, but when combined, it can be used for other purposes and its enhanced value leveraged.

Just imagine, then, what could be achieved if we went one step further still, joining every island together to form a country with a common language. We'll come to that shortly.

Key Takeaways

- Not all data is equal.
- Unstructured data accounts for around 80 per cent of an organisation's data; AI is key to unlocking its insights.
- The data currently being managed by members is potentially the most valuable of all.
- Bringing multiple sources of data together allows for greater value to be extracted.

Why is data so valuable?

I subscribe to the belief that data is business gold. I believe it is a new asset class. I've had a front-row seat to the business insights and applications that can now be delivered – things that were not possible 10 or even five years ago. And the cool thing is that it's only just beginning.

Even now, the combined data in your business is the blood of the company, flowing through tools and people to produce insights, engagements and value. Within it is the history of every decision, every outcome and volumes of behavioural information. It holds intricate lessons for improvement and detailed forecasts of our future.

When all data sources come together, it can create transparency, eliminate doubt, increase confidence and understanding, and provide better foundations for business decisions.

> "Without the right tools, you simply cannot turn your data into knowledge and insight"

Know your toolkit

But to extract this value from your data, it needs to flow through the right tools; I'd go so far as to say the application or tool is where the value actually lies. That's because, while the data powers the tool, without the right tools in the first place you simply can't turn your data into knowledge and insight.

The number of tools available to your business is growing rapidly, along with the volume of available data. Some of that data may be duplicated across the various tools in your business, but all data serves a distinct purpose within each tool/application.

Already well established in the fitness sector are tools that bring together the collection and usage of data, so one system performs multiple tasks: CMS setting up a member record, charging information and sending out newsletters based on communication preferences, for example.

Other common scenarios you'll see involve operators extracting data from one tool to use in another – a task that's become infinitely easier thanks to data exchange standards (think .CSV files). An example of this would be the extraction of data from a CMS system for use in Mailchimp. In an example like this, the data is an island within an application that uses .CSV, and we build temporary bridges to and from that data to achieve our aims.

The most common method of extracting data for everyday use is a Comma-Separated Values file, better known as a CSV. It is exactly as described: a file full of records, with each value separated by a comma.

In machine learning, we often use a CSV 'data-dump' for the historical records upload, using a CSV file to extract data from one application and upload it into Keepme.

Once the upload is complete, there's generally one more step to go through before the new application can get to work: it usually needs to be given key information that will allow it to correctly apply the data it has received. An example would be telling it to pick out the 'member number' from the data it has received, using this as the 'customer number' it requires for the task it's carrying out.

As manual as they seem, a CSV can be 'automated' when there isn't an option to use an API – more on those in just a moment. It's clunky, but a script can be written to export the CSV and then use an FTP service (File Transfer Protocol) – all of which can be automatically uploaded into the target application. When this is done, although it generally means your data will be updated less frequently (think daily rather than hourly), these downsides are typically outweighed by the value of the end product.

```
 1  Activities
 2  Date,Calories Burned,Steps,Distance,Floors,Minutes Sedentary,Minutes Lightly Active,Minu
 3  "01-07-2015","3,953","13,361","10.64","16","1,058","196","71","87","2,393"
 4  "02-07-2015","3,432","10,818","8.76","10","726","224","19","45","1,803"
 5  "03-07-2015","4,102","15,537","12.38","16","701","237","55","88","2,550"
 6  "04-07-2015","3,959","14,490","11.54","6","685","255","46","69","2,443"
 7  "05-07-2015","4,028","14,652","11.68","3","1,003","283","93","61","2,598"
 8  "06-07-2015","3,538","12,161","9.84","16","1,144","168","76","52","1,968"
 9  "07-07-2015","3,550","11,103","8.99","17","785","250","24","41","1,950"
10  "08-07-2015","3,238","10,029","8.22","10","1,204","152","36","48","1,587"
11  "09-07-2015","3,994","13,951","11.29","10","733","183","85","96","2,489"
12  "10-07-2015","3,998","12,516","9.5","9","650","194","92","92","2,508"
13  "11-07-2015","2,802","6,750","5.18","14","594","164","14","29","1,115"
14  "12-07-2015","4,156","19,471","15.39","21","472","272","92","96","2,781"
15  "13-07-2015","3,523","14,730","11.73","17","654","208","39","73","1,959"
16  "14-07-2015","3,039","9,543","7.24","16","712","195","10","38","1,409"
17  "15-07-2015","2,750","6,663","5.06","15","1,222","206","6","6","1,085"
18  "16-07-2015","3,373","14,603","11.46","12","745","198","22","69","1,813"
19  "17-07-2015","3,948","15,787","11.98","29","649","252","23","115","2,471"
20  "18-07-2015","3,917","16,652","13.35","19","543","332","52","55","2,512"
21  "19-07-2015","3,447","11,983","9.84","3","561","255","48","37","1,945"
22  "20-07-2015","3,430","12,816","10.35","15","709","174","37","74","1,834"
```

Tools that set data free

But even when extracted from one application via .CSV and placed into another – out of CMS and into Mailchimp, for example – your islands of data are, to return to my previous analogy, still like separate nations. The ability to understand each other through common language is sufficient, but there's no collaboration across national borders.

The good news is, this collaboration is now possible thanks to the latest generation of tools – tools such as AI, which sets our data free to move beyond the boundaries of where it was formed, and which is capable of combining and evolving data to deliver a new level of value.

All of which seems really cool at a high level – but what value are we looking at here?

As a fitness operator, we have the opportunity to revolutionise our business approach. (I was tempted to say 'evolve', so as not to lose credibility by coming over all sensationalist, but 'revolutionise' is accurate.)

> "Businesses that capture value in their data will be the ones with the highest retention, lowest acquisition costs and happiest customers"

More specifically, businesses that capture value in their data, deploying it through AI, will be the ones with the highest retention rates, lowest acquisition costs and happiest customers. How so? Because the understanding you have derived from the data will allow you to speak to those customers in an increasingly personal way, and each engagement will be more meaningful.

A new super-highway

But to maximise the impact of AI – or any application, for that matter – what you need is a structured, efficient method to get the data you require in the correct format, and at the right frequency. Generally, that's far more frequently than is possible using a .CSV, for example.

This is where an API (Application Programming Interface) comes in – a term that's used with such frequency these days that you no doubt already have a fair understanding of what it's all about. For the sake of being thorough, though, I'd like to offer a quick overview here.

Before APIs were invented, we were all limited in terms of the data we could extract from an application. We'd have to download all the data in a .CSV or use a report writer to export it (always assuming the data was even available, which all too often it was not). The process was very labour intensive and the data extracted would be out-of-date within minutes.

Custom interfaces could alternatively be built, but these were clunky, not to mention expensive and time-consuming to build.

The advent of the API solved all of this.

Picture an API as a doorway into an application – one that's been created specifically to allow you access to the data within that application. When provided with an API key, you can open this door and be presented with a map (aka the API Documentation) that will show you where in the application you need to go to retrieve the information you require.

Armed with this knowledge, you can build a direct integration (picture it as connecting your 'pipe' to the correct parts of the third-party application) to ensure a timely, consistent delivery of the data you require directly into your own application – all without the need for further human intervention.

"API technology makes data totally free to move… a two-way flow that allows all sources of data to be kept up-to-date"

Let's say I had an application that needed attendance data to do its magic. Without an API, this would require a staff member to run a report and then upload that report (assuming this was possible) to get the attendance data into my application. The next challenge: I would then be working with data that represented a snapshot from the moment in time when the report was run. Meanwhile, the live data would be changing every hour of every day; I would never be working off the most accurate, up-to-date data.

With an API integration, on the other hand, my new application is able to call for the attendance data automatically, at whatever frequency it needs to achieve its task. All without human involvement.

And that's not all. Thanks to API technology, which makes data totally free to move, it's now also possible to seamlessly send results back to the original application. Just as one example, let's say we use a CMS API to extract member data into an application like Keepme. Keepme then uses this member data to provide a score on the likelihood of a member staying or leaving; when calculated, this information is sent back to the CMS so it's visible on the member record.

This is a huge value-add of the API: a two-way flow that allows all sources of data to be kept up-to-date – with all connected apps continually updating each other – and, indeed, enhanced as each new combination of data sources drives new insights.

To visualise how all this works, think of the API as a beautifully engineered, wide, two-way bridge that joins data and applications together. Then compare this to the rickety, one-way rope bridges we've had to deal with in the past (and often the present), whereby getting data out was enough of an effort that the idea of sending data back was never even considered.

Insist on an API

Ideally, then, every one of the applications you use within your business should have an API, so a quick audit of your current tools' data capabilities is a useful undertaking.

> "Insist on there being an API. Without one, you could find yourself severely limited in the new data age"

List your core business tools, such as your club management and entry systems. Do they have APIs available for them? If so, request a link to the API Developers' Documentation, which you will be able to review/share when discussing new tools/ AI projects in the future. This documentation should include a list of endpoints that indicate what parts of the application are accessible. A limited set of endpoints will limit use.

No API? Ask the provider when this will be available. Ask specifically: "When is an API scheduled on the roadmap for delivery?"

Believe it or not, there are industry tools out there that have neither an API nor the ability to export your data via a .CSV. I'm not talking fringe providers, either; some of the biggest providers have tools that lack this capability. They may suggest they

are legacy platforms, but if you find yourself owning one of these, it's time to move on. Do not let your data be held prisoner, as the value lost could become significantly limiting over time.

(Incidentally, don't be fobbed off by proprietary data extraction/integration tools, either. These may be a possible stopgap until an API is available, but they're very old-school, far from optimal, and most definitely still a short-term solution.)

Certainly, for future application purchases, I cannot recommend strongly enough that you insist on there being an API. If the answer is "it's in the pipeline", it's a red flag. Push for more certainty, as without an API, you could find yourself severely limited in the new data age.

Tools that help connections

Let me briefly revisit a previous point: that the value of data can multiply when two applications are joined together.

The good news is, this doesn't need to be a technical undertaking.

Step forward Zapier, an excellent little service that connects two or more apps, allowing you to create workflows – which they call Zaps – which happen automatically when the conditions are met.

More specifically, a Zap is when a trigger event in one application starts a workflow that, in turn, produces the desired outcome or action in another. By way of explanation, let's dive straight in with a few examples.

- Connecting your company's Linkedin account with Slack, so every time a post is made, everyone is aware. Every time someone comments on the post, it also appears in the Slack channel.

- Ensuring all files are automatically stored in the right place in the cloud.

- Automatically generating (and notifying team members of) new leads based on details submitted via Facebook ads.

- Getting an alert in the relevant Slack channel if there's a customer payment failure in STRIPE.

There are so many other options, too, because more than 3,000 apps work with Zapier. And as an added benefit, it's also simple to learn and use.

(Note: IFTTT is another app in the category, though from my experience it's better suited to personal rather than business workflows.)

Key Takeaways

- Data is only useful if you can extract it from your systems. APIs are key here; without them, you could find yourself severely limited in the new data age.

- The two-way flow of an API doesn't just allow you to keep data sources up to date. It also allows for the data within those sources to be enhanced.

- By joining your 'islands' of data together, AI can eliminate doubt, increase confidence and transform business performance.

What are the first steps?

Forgive me for starting a 'first steps' section with a nod to the bigger picture. It is, however, a useful way to set the scene, underscoring the importance of taking the steps I'll shortly recommend.

That's because I foresee a whole new generation of off-the-shelf applications being with us in the not too distant future. These applications will have shallow learning curves, allowing the end user to control everything from beginning to end; no need for a data scientist to make things a reality.

> "You might call it a data audit. I call it an asset audit due to my understanding of what can be done with it!"

I would strongly suggest you do not wait for my utopia, though. Far better to gain competitive advantage by starting now. But either way, one thing will be required, and that's data. Data will be the fuel for the insights, innovation and results. Without it, you will simply not get to the start line of the new (data-driven) age of insights and intelligence.

I promise you, there will be no business wishing it had less data in the coming years.

If you are bought into this, you need to start establishing your data practices. Let's take a look at some of the things you should consider.

1 Asset evaluation is a great place to start, with an excellent exercise being to look at every area of your business – every system you use – to map where data is captured.

You might call it a data audit (I call it an asset audit due to my understanding of what can be done with it!), but list the data capture points in two columns: the online and offline. Be comprehensive, as you won't want to find out in 12 months' time that you missed one of importance. Look at the tools used across the business – not just proactive collection such as membership applications and PT bookings, but also passive data sets such as entry systems.

For the offline, review what becomes digital, noting whether every item does indeed make the jump and what, if any, captured information gets left behind. Note any islands of offline data that are not digitised, so you can be aware of them later for evaluation of their value and relevance. Finally, consider any automatic data collection, such as entry systems; often, these sources are missed as they do not have human involvement.

> ## "A good policy is simply to collect everything, just so you don't later regret any omissions"

2 Give each data source a condition value: a score of 0-10, with 0 being 'currently a mess' and 10 being complete and perfect. Your focus here should be on the manual entry systems, as it will be the human element that lets you down. However, even for automated systems, check what data capture level you've chosen. Have you set it at a default level, by which you're only capturing the data you need right now? If so, is there an option to expand this?

A good policy is simply to collect everything, even if you don't have any current defined requirement for all of it right now, just so you don't later regret any omissions. After all, you never know what could be of value in the future; imagine finding out, a few years from now, that some incredible insight is possible courtesy of a new tool – but not to you, because you chose not to collect the relevant data. If there's no additional cost associated with doing so, turn it all on!

3 Any data sources scoring below 7/10 should be identified, and a plan put in place to both rectify the historical gaps and educate on what's required going forward. There may be a lot of work here, but there are freelancers on sites such as Elance who have the skills to fix data issues such as these, and who are often able to supplement with expanded data too.

4 Ask yourself: Can automation play a role? Do you currently require fields to be filled manually where this could actually be done automatically? For example, does the personal training sign-up form include gender or age, when this data is already held on the member record?

> "Everyone in the organisation needs to understand the value of what's being collected, and their role in ensuring it's done correctly"

5 Establish responsibility. What gets measured gets managed, as they say, and this is no different. You would certainly have someone looking after the cash in the business; data should be no different. A Chief Data Officer maybe a little too grand a title (though 25 per cent of Fortune 500 companies now have them), but you want someone who can own the business data performance and report on data conditions.

In some businesses, the owner/principal would be best, at least until culturally it becomes an accepted behaviour. The senior marketer is also a good choice, as they are usually the current beneficiary of good data practices and will have cause to access most of the data sources within the business.

6 Implement a data policy. Do not get this confused with GDPR or customer confidentiality, though you can undoubtedly integrate these if you wish. Your data policy should describe the importance placed on data within the business, explain how diligent you expect staff members to be, and encourage them to report any data-related issues.

Everyone in the organisation needs to understand the value of what's being collected, and their role in ensuring it's done correctly. In my experience, the best way to do this is to make clear the benefit to the business of them doing so, rather than focusing on rules or reminding them that it's part of their job. I suspect they all know they should be doing it, but it's the difference in diligence around it that will determine whether or not you're working with quality data in the future.

Key Takeaways

- Do a comprehensive and honest audit of all your data sources, then make a plan to plug the gaps.
- If there's no additional cost associated with doing so, collect all the data your systems allow.
- Build a culture of data responsibility within your business.

What if my data is imperfect?

One of the most common questions I hear when businesses look to work with their data is: what if my data is imperfect? My answer is always the same: imperfect for what?

With almost every data project, it's inevitable that the data will not be optimal unless it was collected specifically for the purpose or project you're now trying to implement. In the future, that sort of specificity may become possible, even the norm, across the board. For now, however, we're taking what we have and applying the new tools and methodologies to the data – and regardless of data condition, it results in giant leaps forward.

That's because – with the advent of AI, Business Intelligence (BI) and Big Data – it's now entirely possible to improve the datasets that already exist within companies. Whether it be ensuring a consistent date format or backfilling missing gender fields, data teams now have a plethora of tools and understanding which allows them to improve the data prior to it being used; in my experience, unless there has been a vigilant, consistent approach to data collection, there will always be improvements that can and will be made.

To illustrate the type of challenges that can be found with a typical dataset from a club management system, let me provide an example of a recent piece of work we were involved in.

Pre-data cleanse

Data duration: 3 years
Member records: 12,659
Complete Records: 0
Duplicate Records: 2,187
Membership Types: 63
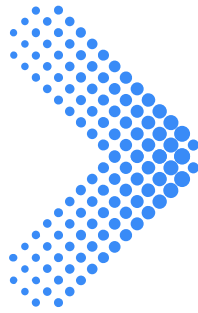
Post-data cleanse

Data duration: 3 years
Member records: 11,201
Complete Records: 11,019
Duplicate Records: 0
Membership Types: 7

Ninety-five per cent of this work took place without the need for client consultation; the tools we have were able to cleanse and supplement the records. Where we did need to work with the customer was on Membership Types, to clarify the difference between a number of apparently very similar seasonal membership deals.

Quality over quantity

Often, people still think in terms of data islands. They point to holes in the data from one application, when in fact those same holes can be filled using the data that already exists in another system.

One interesting example would be missing class attendance data from the CMS. Armed with the CMS data, plus the site attendance data from the entry system, in most cases it can be determined whether or not the member attended a class based on time of entry, time of exit and any other 'evidence' within the data.

> "It's better to have 12 months of comprehensive, clean-ish data than three years of rubbish"

Then there's the question of quantity: how much data do I need? That, again, will depend on the outcome you're looking to deliver, but it's better to have 12 months of comprehensive, clean-ish data than three years of rubbish. Interpolation can be achieved from a small clean data set to make up for lack of duration.

Advances in methodologies have helped here, driving a decrease in the quantity of data required. For example, in 2019, Keepme needed three years of member data to provide the Keepme Member Retention Score. By January 2021, only one year was required.

💡 Whose job is it anyway?

The first task within any data project is to understand if you actually have the data needed to complete the project.

If you're working with an external supplier, that is their responsibility. You are signing up for the delivery of business value that's powered by your data. Your responsibility is to provide the data (and some $ for the service); their responsibility is to provide the value.

If your data is not up to the task, they will let you know – but I advise you to allow them to do that evaluation. The tools and skills they have at their disposal are extensive, and they include automatically detecting and dealing with missing data.

Find your unique identifiers

All of this aside, it's a sad fact that even perfect data is often useless. If you cannot link the data to the individual member through a unique identifier such as a membership number, then it's unlikely to be of any use. Those customer feedback screens that allow you to choose the face and rate the service are a good example. They may be great at giving you a high-level customer service view, but the data is pretty useless beyond that.

However, don't overthink this. A unique identifier can be flexible. Let's say you've been using Mailchimp for all your member email campaigns. This is potentially valuable data, and the actions of the member (did they open, click, share, unsubscribe etc) can be linked through the email address. Mobile phone numbers are also used, along with home addresses – though the latter less so because of multi-residency.

💡 How good is your data?

A simple way to evaluate the quality of the data generated by any given tool in your business is as follows.

- Use the report writer or export function to produce a .CSV file.
- Open the file in Google Sheets or Excel.
- Select the contents of the sheet, and from the Data menu, create a filter.
- Go through each column, and you will quickly be able to see how many records have empty fields or how many have data in the incorrect format.

In some cases, you might go a step further and use 'find and replace' to insert the correct values – where known – before uploading the data back into the host system. Indeed, an interesting by-product of the data work we've inevitably had to do for our customers has been their ability to use the cleaned-up output to re-populate their platforms with new data.

Remember, though, that fixing without educating users means you're just kicking the can down the road. Report your findings and prescribe how you expect it to be done moving forward.

If you're dealing with a real mess, don't give up. As I've mentioned previously, hundreds of freelancers are available to take on data cleaning tasks, and for very reasonable costs.

Do I need a warehouse?

A data warehouse is an application (technically multiple applications) incorporating storage that pulls data from all your business sources. It's there to ensure consistent extraction and storage of data and then make it available in formats that make it easy to work with (remember SQL?)

Data warehouses have been around for 15 years now, with the idea being that, if all the data were housed in one place and in one format, we could do lots of cool stuff with it.

The reality, however, is that the data warehouse was a solution looking for a problem. It felt like insurance, because you could go to bed knowing you had 'banked' your data for when you needed it. In practice, though, there weren't many use cases that would support the cost of building a data warehouse in the first place.

> "Even perfect data is often useless. If you cannot link the data to the individual member, it's unlikely to be of any use."

If you're a large organisation that identifies as a data-driven business, with resources dedicated to technology delivery, then a data warehouse could potentially be right for you. Perhaps you've made acquisitions, or international expansion has left you with extensive unconnected data assets. If so, a data warehouse could be a useful tool, connecting otherwise unconnected applications by drawing all their data into one central repository.

I've certainly seen a few data warehouses on my travels within fitness operators. Few of them are heavily utilised, however, and most were purchased to fill data connectivity gaps between applications.

For the majority of fitness businesses, I would suggest a data warehouse is an unnecessary cost and complication. In many (most) cases, an API would in fact suffice, allowing for a more efficient, less costly integration without the need for a third layer.

Additionally, cloud services are becoming increasingly available, and 'middleware' services – Segment.io, as an example – offer a much more manageable and cost-effective solution.

Middleware sits in the business like a data traffic policeman at a busy intersection. As customer data is captured within multiple applications, it's sent to the middleware in real time, which then distributes it to other applications that may require it.

The beauty of it is that one data element, such as 'payment status', can be distributed dynamically to multiple other systems that can use it. So, the payment gateway in a club confirms a member's charge has gone through, which triggers the middleware to update the Entrance Gate application to let the member in, as well as the CMS so member services can see the member is current and up-to-date.
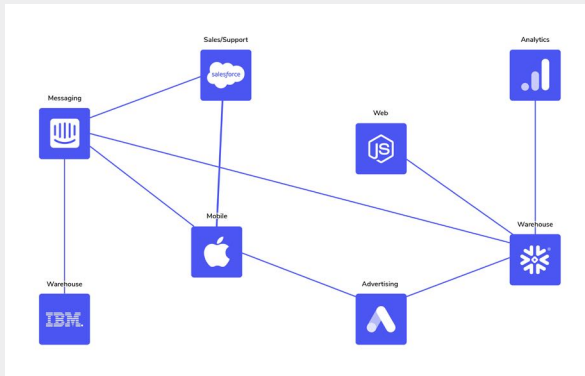
Without Middleware
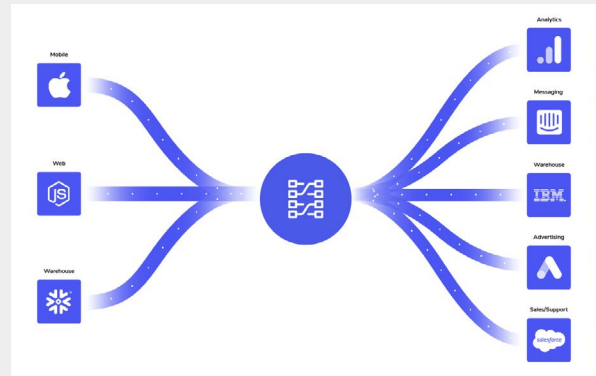
With Middleware

The link between cloud services (these middleware services such as Segment) and API is the fact that data is free to move from application to application (through API to API, or API to middleware for example) without the need for costly data warehouses – because it can all be stored in the cloud.

Key Takeaways

- Don't worry if your data isn't perfect; there are tools to fill gaps and dramatically improve it.
- It's better to have 12 months of comprehensive, clean-ish data than three years of rubbish.
- Data needs a unique identifier to be useful, but be creative in where you look for this.
- API and cloud-based services are more appropriate solutions than a data warehouse for most gym operators.

What does AI mean in the context of my business?

We've mentioned AI a few times now, but let's backtrack and start by explaining what Artificial Intelligence (AI) is.

Contrary to popular belief, AI is not new; it has been actively researched since the 1950s. However, a number of factors have converged since the start of the 21st Century, leading to a rapid acceleration in progress.

AI is now everywhere, suggesting what we might listen to (Spotify) and watch (Netflix), providing navigation guidance (Waze) and diagnosing medical problems (Health Tensor) among thousands of other things. In fact, it's so ingrained in our everyday lives that it confirms a statement made by John McCarthy – the man who coined the term 'Artificial Intelligence' in 1956 – which was that "as soon as it works, no-one calls it AI any more."

> "Think of AI as a prediction machine: it takes what we do have (data) and fills in the gaps we don't (i.e. knowledge)"

Let's put aside self-driving cars and robots taking over the world for a moment and frame AI specifically for our industry. I encourage you to think of AI as a prediction machine, because that's what it is: it takes what we do have (data) and fills in the gaps we don't (i.e. knowledge).

- Will the member leave us?
- Will they be interested in this offer?
- Will this change to the schedule affect attendance?
- What will my sales performance be three months from now?

How it gets there may be a little bit more involved, but think of it as a prediction machine and you won't go far wrong.

- Based on what they've listened to, I predict they will like...
- Based on what they've watched, I predict they will like...
- Based on my understanding of the road system and the reported traffic conditions, I predict an ETA of...
- Based on the symptoms, I predict the condition is...

The ABC of machine learning

Let's jump down a relatively shallow rabbit hole to consider how this is done.

The majority of the predictions we use in our industry come from a branch of AI called Machine Learning – a term we'll all have heard increasingly often over the last few years, but one that's actually been around since 1952.

In the simplest terms, we can explain machine learning as follows: it's when we train our prediction machine to predict an outcome by teaching it using data from what's happened in the past.

Let's say we want to predict new member sales performance over the coming months. We'd start by extracting all the historical data relating to sales, including all the leads, where they came from, whether they converted or not. We'd then feed 80 per cent of this data – complete with outcomes – into our prediction machine for it to learn how an outcome is formed. Once it's confident of its abilities, we'd then feed it the final 20 per cent of the data without the outcomes, asking it to predict these for us. This allows us to understand how accurate the prediction machine is, and to tune it if necessary. I will cover prediction accuracy in more detail later.

> **"I'm not saying humans absolutely cannot do this, but I wouldn't bet the business on the accuracy of their every forecast."**

We then hook it up to the live data source (API) so it can continue producing predictions for your business. All the while, of course, it's also consuming the new data and actual outcomes, so it keeps learning and constantly refining its output.

That, in a nutshell, is how machine learning works.

A limited narrative

So, now you know how we train a prediction machine, let me move on to explain why it's so accurate.

Suppose we stick with our sales forecasting example. A non-AI prediction would likely be based on experience, ninja skills or guesswork. I've been on the receiving end of sales forecasts that incorporated all of these things at one time or other, and while I'll bow to experience over the other two any day, what unites them remains their lack of usefulness.

Sales may be a science, but so too is forecasting, and there are so many variables to consider: human performance, market conditions, weather, pandemics (a worthy mention), competition and seasonality to name just a few.

I'm not saying humans absolutely cannot do this. What I am saying, though, is that I wouldn't bet the business on the accuracy of their every forecast. We humans like to find reasons that explain why things happen (in sales forecasting, more often than not to excuse how wrong we were), so we define a narrative that we believe dictates – or at least heavily influences – the outcome we're trying to predict.

"It's September, the kids are back at school, lead generation is up and we have a new membership type available, so let's say we'll see a 10 per cent uplift in sales."

In reality, however, there are so many more variables at work than this – many more than we contemplate – and this is then multiplied hundreds of times over when you combine all the permutations.

The ML approach

Let's now take a look at how a machine learning algorithm would approach our challenge.

We will all have seen a decision tree at some point; here's an example of one where we see it branch out in different directions depending on the decision being made.
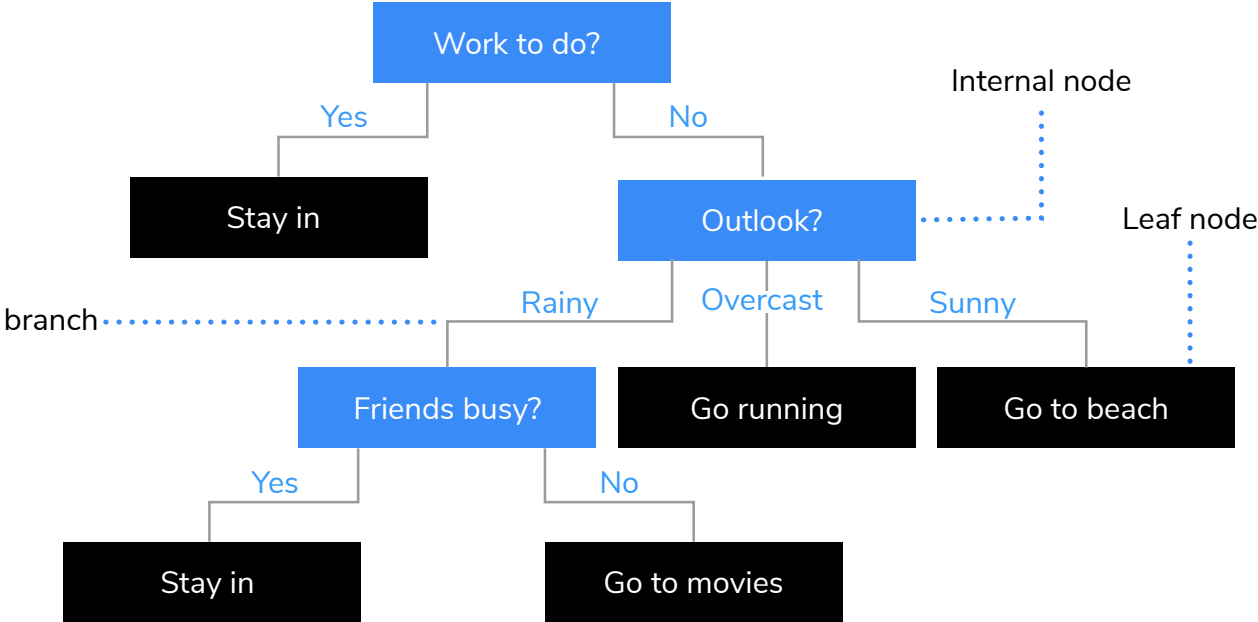
A machine learning algorithm will use a decision tree to understand the impact of each data item on the outcome. The starting point might be 50 items that we can plug into it: lead source, location and date, just as a few examples. Your prediction machine

doesn't stop there, though. It takes every item and runs every single permutation randomly. In our example of 50 data items, that leads it to generate millions of decision trees (which, incidentally, is the reason why we call this machine learning method 'Random Forest').

> ## "Don't waste your time even trying to provide your expert opinion to a data scientist!"

Most of these decision trees will not be statistically relevant when taken on their own. However, when you combine them into the forest, you're presented with a highly accurate prediction, far beyond what we may have been able to do with our Excel spreadsheet and a sprinkle of hunch from our past experiences.

In fact, don't waste your time even trying to provide your expert opinion to a data scientist. Experience can certainly assist in data collection, but when it comes to prediction, the results secured from AI + machine learning will be many times better than anything even an expert could achieve.

Humans vs AI

A LawGeex study compared the effectiveness of lawyers versus AI in assessing the quality of non-disclosure agreements. The study involved 20 lawyers, with decades of experience in companies such as Goldman Sachs, pitted against LawGeex's proprietary AI platform. Participants evaluated the risks contained in five agreements by searching for 30 specific legal points.

The results of the study showed AI achieving an average accuracy of 94 per cent, and maximum accuracy of 100 per cent. Humans achieved 85 and 94 per cent respectively (see The nuances matter in our 'How confident can I be in AI predictions?' chapter to understand the full implications of these apparently small percentage differences).

> "On average, the lawyers required 92 minutes to complete the task. AI analysed all the documents in 26 seconds."

Then there's time efficiency. On average, the lawyers required 92 minutes to complete the task. AI analysed all the documents in 26 seconds.

Let's take another example. Unlike chess, the game Go was considered beyond the reach of even the most sophisticated computer programmes. The 4,000-year-old board game is incredibly complex, with more possible configurations for pieces than atoms in the observable universe.

In 2016, however, Go world champion Lee Se-dol was beaten 4-1 by Google's Alpha Go. Initially trained with a dataset of 100,000 human games, Alpha Go then went on to play itself over and over, developing the ability to play the game with blinding speed. In one game, Alpha Go famously played a move that all the experts assumed was a mistake, but which went on to be the defining move of a game it subsequently won.

Grand Master Lee Se-dol later retired, saying: "Even if I become the number one, there is an entity that cannot be defeated."

What's the difference between AI and machine learning?

The short answer is 'nothing!' In simplest terms, Machine Learning (ML) is a type of AI and one of a number of approaches you may have heard about, such as Deep Learning and Neural Networks.

ML is all around us, powering the likes of Amazon Alexa and Apple's Siri. It's a branch of computer science that, using statistics, involves finding patterns in data to solve mainly classification problems – e.g. will the member stay or go?

There are actually two relevant subcategories in ML, and if you can stick with me for the next couple of minutes, you'll be primed to win the next trivia challenge or quieten that "I know everything about AI" friend of yours.

> "Unsupervised learning can lead to great insights, as patterns can be found that wouldn't be possible in supervised learning"

Supervised Learning is where we input the data and use an algorithm to learn what gets us from the input (data) to the output (answer). We call this Supervised Learning because, while the algorithm is making its predictions, it can be corrected by the teacher until it achieves the required level of performance.

Keepme uses a supervised learning algorithm called Random Forest for the Keepme Score, among other things.

Unsupervised Learning is where a lot of the excitement (and hot air) is often found, and is most often employed when we're dealing with unstructured data. Rather than having an output to look for, the algorithms are left on their own to investigate and present interesting findings in the data. This can often lead to great insights, as patterns can be found that wouldn't have been possible in a supervised learning approach.

Unsupervised Learning hasn't yet been used in fitness, but its application is something we're already exploring at Keepme. Ideal customer profiles are one such opportunity, whereby our AI tools could take existing data, cluster customers around every variable and develop brand new target segments.

Key Takeaways

- Think of AI as a prediction machine and you won't go far wrong.
- Attached to a live data feed, machine learning constantly learns and refines its output.
- AI's 'forest' of interlinked decision trees ensure its predictions are far more accurate than anything even an expert human could achieve.

What are the alternatives to AI?

Before the tools to implement AI became freely available, we did have other options. Let's just take a quick look at the alternative approaches we could consider.

Domain expertise

A fancy way of saying 'experience' – and for rudimentary analysis, it's a good start. If we take a look at members coming up for renewal in the next 30 days and identify those who are still active, we can fairly assume there's a strong probability they will continue. If we have 25 water bottles left in stock at the end of the month, and average monthly sales have been 28 for the previous three months, it's probably best to order some more stock.

> "Domain expertise falls down when we want to develop an understanding that spans longer periods of time"

Domain expertise on these types of outcomes makes sense, if a little simplistic and often without affording a time window to change the outcome if we're wrong. Where it falls down is when we want to develop an understanding that spans longer periods of time.

Why would we want to do this? Why not simply rely on the domain expertise as noted above? Simple. If we're already able to identify in March the members with the highest probability of leaving us in December, we then have a window in which we can try and change the outcome. This forward-looking insight is considerably more valuable than any 30-day prediction based on even the most extensive experience.

To our other example, perhaps a bulk purchase of bottles would secure us a lower cost per bottle – i.e. a solid 12-month forecast of requirements would be valuable.

There's also the fact that experts, like the rest of us, are prone to cognitive distortions that materially impact the quality of the end result. Prejudices, incorrect assumptions, the need to simplify, ignoring contradictions and, of course, emotions all affect the output.

None of these weaknesses need be of concern when AI is brought in to analytical processes.

Business Intelligence

Business Intelligence (BI) platforms in our industry fall into two camps: either too broad a capability, requiring a greater resource effort than we're prepared to make; or sector-specific and too simplistic. Certainly the vast majority I've reviewed seem to focus on presenting simplistic, standard metrics in a visual form.

Whenever I'm presented with a BI Dashboard, I go through each section and ask myself two questions: 'Why do I care about this insight?' and 'What can I do with it?'

A typical 12-month member retention chart

A stalwart of any industry dashboard, the 12-month retention chart shows us how our member retention breaks down by month. Why do I care? I care because understanding whether we're retaining members is critical to our long-term business success; arguably, it's the most important metric in the business. But what can I do with it? Not a massive amount.

I can establish whether there's an upward or downward trend. Be alerted if the previous month was considerably better or worse than the one before. But as for deliberate actions that can be driven from it, not much. The historical nature of the metric is limiting, and the one-dimensional presentation doesn't provide many initiatives for action.

Ideally, I want a forward view for the coming year, not a retrospective view on the past. This will allow me to understand where I need to investigate further and what plans I can put in place to mitigate.

Retention prediction chart

However, even if all we have is the historical performance chart, the capability to drill down would make it more valuable.

For example, if I could see not only that last month's retention was a decrease on the previous month, but that it was a specific membership category that drove the bulk of that reduced performance – members who, 12 months ago, signed up to a specific offer that's no longer available – my understanding of the situation would be significantly improved. I would at least have an insight into why we got the performance we got.

All that said, I go back to my previous point: that had we been able to see this five months ago, we would have been able to take timely action to remedy it.

Statistical modelling

Move on, nothing to see here.

There may well be operators employing statistical modelling, but if you're reading this and aren't employing it, my advice is that there's little point in doing so. A considerable step beyond BI, it nevertheless requires having a person with this skill set on-staff – and if you do have such a person, I would suggest they're either lost, or else they're a sadist who enjoys working with statistical modelling packages like MatLab.

Let me add four sentences to this section, just so I've fulfilled my promise of a sufficiently thorough explanation to ensure all options are understood.

> "If you're reading this and aren't employing statistical modelling, my advice is there's little point in doing so"

Statistical modelling is when we take a data item such as member attendance and use it to predict an outcome such as retention. As a method, it presents us with more accurate insights than domain expertise and more usable insights than the basic BI tools.

It comes down to two flavours: single and multi-variate. Single is where we take one value to model, such as attendance; multi-variate is where we take (yes, you guessed it) more than one. Accuracy is going to be better than a human, and there's certainly some value in the exercise.

However – and this is the key – if you aren't already doing this, there's no benefit in doing so with AI now available to you.

Key Takeaways

- Relying on experience alone will not give you the chance to implement timely change.
- Business Intelligence tends to lack actionable insight.
- AI has superseded the need to invest in statistical modelling.

Where could AI add value to my business?

The simple answer is 'everywhere', but let me paint a picture that will equip you to easily answer the question yourself.

Take a look at the business metrics you currently work with, the vast majority of which will be looking at what's happened in the past. Now ask yourself: "What could I do with this insight if it were predicting forward?"

> "If you currently review historical metrics, the data for an AI prediction most likely exists in your business"

The good news is, if you're currently reviewing historical business performance metrics, this indicates the data needed for an AI prediction most likely exists in your business. This a great place to start.

Now feel free to take inspiration from just a handful of the ways I've already seen AI technology applied within fitness operations.

- What will be our member retention for the coming six months?
- Which specific members will leave us in six months' time?
- What will our sales be for the coming quarter?
- Which sales channels will decrease in performance over the next quarter?
- Which members will most likely purchase PT, and at what package level?
- What impact on member retention would a class schedule change have?
- What is the optimal class schedule for member retention?
- Which classes drive the highest member retention?
- Will we hit our sales goal this year?
- What will our membership be at the end of the year?
- Create a custom audience for us, based on our optimal member, for our social ads
- Create a list of members most likely refer new members/provide a review

It's clear that, armed with these insights, a business would be able to take consistent, positive action. It isn't hard to generate considerable value from each insight.

Remember, too, that with every action taken, your AI can see the outcome of the action, so it continues to learn.

How confident can I be in AI predictions?

A question I frequently hear relates to the accuracy of AI – namely, how confident we can be in its predictions.

It's understandable when you consider that we're moving away from an expertise-led approach to effectively handing over control to an algorithm – and the bottom line is, whether human- or machine-generated, every prediction has an error possibility. The goal has to be as high a degree of accuracy as possible.

Let's take a look at how accuracy is measured in machine learning and contrast that with a human approach.

The human touch

Here's a big statement to start with: A machine learning model will always do better than any experts you have within the business. (If in doubt, I refer you back to my LawGeex and Go Grand Master examples.)

Sure, your team has decades of experience, but try this out. Sit the team down and set them a prediction challenge. Future sales performance would be a good one – i.e. what new member sales figures will be achieved in each of the coming three months?

Note that you aren't actually interested in the output here; you want to understand their inputs. Ask the team to list each data point they consider and you'll find the majority of points are put forward by every team member: previous month sales, time

of the year, performance in the same period last year etc. Equipped with these data points, they will add their experience and come up with a prediction.

I suspect this is not too dissimilar to the approach taken by most operators currently. I would hazard a guess that it has some value for short-term forecasting, maybe for the coming month, but it can be very hit-and-miss. Move beyond a month and (as previously observed in our assessment of domain expertise) the accuracy deteriorates.

"A machine learning model will always do better than any experts you have within the business"

Adapting to change

Now let's look at the AI model, with apologies that some of this material has been touched on previously. It's so fundamental to explaining the accuracy of AI predictions that we need to quickly skim over it again.

Your AI will look at every data item you feed into it: all those your team is likely to list, plus many more. It will look at every possible permutation and even develop some of its own. For example, we may look at the previous month's performance. It will look at the previous week, day-to-day comparison, quarterly and even hourly. For every new data item (they are called 'features' in data science terms), AI compares it with every other item so every conceivable permutation is considered.

As we have discussed previously, this can lead to millions of permutations being considered. Some of these permutations will have absolutely no value. Others will represent 0.01 per cent of the outcome. But when you add them all together, you get a level of accuracy that cannot possibly be replicated by a human.

You will also understand this accuracy level in the set-up phase – i.e. before deployment – due to the use of historical data to train the AI.

Neither is it simply a case of right or wrong. It is supervised learning, meaning we assist the AI in finding the optimal model. We use initial findings to address any errors and fine-tune the model further (for example, correcting a model when it 'over-fits', which is when the AI takes random fluctuations and noise found in the training data and learns these as concepts.)

Then there's the capacity to learn on an ongoing basis, even once the system is live. Every day, as new data becomes available, the model continues to evolve and learn, ensuring the accuracy is constantly refining. The 'learning' part of machine learning pays dividends.

Sadly this rarely happens with our team-based predictions. In fact, expertise or domain-based predictions very often deteriorate over time, because while the data is changing, the expert does not.

The nuances matter

What is acceptable for accuracy depends on the prediction. Forgive me if this sounds like I'm dismissing the importance of membership sales forecasts, but I think everyone would agree the impact of an incorrect diagnosis by AI analysing an X-ray of a patient with suspected breast cancer, for example, sits in a different league.

Depending on the question being asked, it can also be the case that an incorrect prediction is actually a good thing. A false negative for a gym operator posing a retention-focused question would see a member who was expected to leave end up staying – no bad thing! For the medical diagnosis, of course, a false positive would mean a tumour predicted to be benign turning out to be malignant.

> "If you're finding some predictions are wrong, make sure you determine if the wrong is a good wrong or a bad wrong"

At the risk of sounding simplistic, then, if you're finding some predictions are wrong, make sure you determine if the wrong is a good wrong or a bad wrong. Some incorrect predictions have more impact than others.

And whatever you're predicting, note that there's often a tendency to underestimate the impact of small improvements in prediction accuracy. Let's compare the improvement of two models, one from 85 to 90 per cent, the other from 98 to 99.9 per cent. Our first improvement is impressive, as it means mistakes fall by a third. But in our second example, they are actually falling by a factor of 20.

And there will always be improvements to be made – even with machine learning, as it continues to learn. Those small improvements can be material.

Don't just flip a coin

When judging prediction accuracy, there's also the question of what to benchmark against. I will predict (domain, not machine learning) that very few organisations measure – or would be prepared to measure – the accuracy of their sales forecasting. Yet they will cast doubt on a machine learning predictions if the results are not what they want to see.

Back in the early days of Keepme, we ran a model for a UK operator. They passed us their (somewhat patchy) data for the previous two years and held back the current year. The challenge was to identify which of the members would still be with them in six months' time – which Keepme was able to predict with 82 per cent accuracy (before tuning, I may add, and with limited data made available to us; our average out of the gate accuracy is between 94 and 95 per cent).

Compare this to the operator's previous probability of accurately predicting whether a member in month six of a 12-month contract would stay or go, which was effectively a coin-toss (aka no more than 51 per cent).

> "Even the jump from 51% to 82% would have provided a tremendous platform to confidently target those at risk"

Even that initial jump from 51 to 82 per cent would have provided a tremendous platform to confidently target those at risk – including members who themselves didn't yet know they were at risk – in a timeframe where the operator could change the outcome.

Delighted with the results, we sat down with the team to hear the CFO open the meeting with: "It's not very accurate then. It's a long way from 100 per cent".

Let's remind ourselves that previous accuracy was no more than 51 per cent (this assuming anyone had the time to run these figures in a multi-thousand-member business). In contrast, they now had an opportunity to be right more than eight out of every 10 times, with constant opportunities for further improvement (to reiterate, it's all about the 'learning' in machine learning).

Returning to my initial observation, then, and acknowledging that every prediction – whether human- or machine-generated – has an error possibility, the goal has to be as high a degree of accuracy as possible. With that in mind, 82 per cent (and counting) versus 51 per cent... I may be biased, but to me there's a clear winner.

Measuring accuracy

Measuring accuracy is an area where we could get lost in detail and math, but I will try and ensure we do not. I do, however, think everyone should be versed in the basics, as I suspect every one of us will be involved in a discussion on this subject as we deliver projects over the coming years.

Accuracy (ACC) is simply* measured as follows.

$$Accuracy = \frac{Number\ of\ correct\ prediction}{Total\ number\ of\ predictions}$$

The result will be a value represented as 0.x – so for example, 0.89 would be an 89 per cent accuracy.

I say simply because there are many ways of measuring accuracy – among them, ROC AUC, FPR and R², Coefficient of Determination. If you find these becoming your domain, then check your business card as you've clearly had a career change somewhere!

In binary classification models such as member retention (will they stay - positive class; will they go - negative class), accuracy is calculated in terms of positives and negatives.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Where TP = True Positive (model predicted they would stay), TN = True Negatives (model predicted they would go), FP = False Positives (model predicted they would stay and didn't), and False Negatives (model predicted they would go and they stayed.)

Once again, the result will be a percentage to show accuracy.

It's possible to accurately measure a model's performance at set-up and over time, too, although I'd certainly suggest you bring in a data scientist to do this. They will use a Confusion Matrix (you couldn't make this stuff up) to measure the model throughout its lifetime and present the error types (i.e. False Positive and False Negative). In turn, this allows them to understand whether or not the model is performing well, and to rectify any errors that are occurring.

In brief...

Let me summarise on accuracy. First and foremost, if you don't have anything to benchmark against currently, then it's safe to say any machine learning model is going to be an improvement.

Second, it's possible to accurately measure a model's performance in the set-up and subsequent life of the model, so accuracy can always be monitored – and refined.

Third, understand the impact of false positives and false negatives (see The nuances matter). If I'm predicting whether a member will leave in a retention model, then a false positive – i.e. the model reporting they will leave when in fact they do not – is unlikely to cause a negative impact. A false negative, on the other hand – predicting they will not leave, when in fact they do – has more serious repercussions. That member, assumed to be 'safe', will generally have had no intervention, no effort made to retain them, on the basis that it wasn't needed. In the fine-tuning of your model, these errors clearly need to be addressed as a priority.

💬 Key Takeaways

- Every prediction has an error possibility. The goal must be as high an accuracy as possible.

- Some incorrect predictions have more impact than others; determine if a wrong is a good or a bad wrong.

- With AI, you can fine-tune the accuracy of your model even before deployment.

- Expertise-based predictions often deteriorate over time.

What sort of results can I expect?

Before you take on machine learning, you'll want to understand the likely impact it will have on your business. It's a fair question, but I'm going to answer it by having you ask yourself the following questions.

- If I knew a member was going to leave us six months prior to the end of their contract, what value would that have?

- What would I do differently if I knew which lead would convert to membership and which would not?

- If I could predict with 90 per cent accuracy what my sales figures would be for the next six months, what could I do with that knowledge?

- If I could send a commercial offer to a member, knowing it was highly likely to be of interest, what impact would that have on our bottom line?

- If I knew which of my members to ask for referrals or testimonials, and be confident they would provide them, how would this change our approach to member acquisition?

Let's start with these, but understand that there are many, many more questions you could ask of your prediction machine. I'm repeating myself here, but take a look at any part of the business you currently view in historical terms. Look at your monthly reports and ask yourself: 'What value would it have if I could also predict forward?'

> "Look at your monthly reports and ask yourself: What value would it have if I could also predict forward?"

Results that speak for themselves

In terms of results that can be achieved by this sort of future prediction, really you have to look to our business – Keepme – as it's the first to provide these services to the fitness industry. At two years old, with much of this time of course shaped by COVID-19 lockdown, we are only now starting to get customer report results.

However, there have already been some extraordinary improvements reported by customers from around the world. These include:

- A one-year average increase in length of membership, resulting in a significantly improved member Lifetime Value (LTV).

- An 8 per cent increase in sales lead conversion in the space of just three weeks (annualised 233 per cent increase in lead conversion).

- A 35 per cent increase in non-dues spend.

- A 21 per cent rise in referrals (in hard terms, four in five members are now referring new members) leading to higher sales, lower member acquisition costs, and higher LTV.

- Plus a lot of little wins that make a difference: email open rates improving from 52 to 86 per cent, for example.

Put all this together and you'll appreciate it's not just the value of the solo prediction. It's the impact the combined predictions have on subsequent actions, outcomes – and future predictions.

> ## "Machine learning increases transparency and reduces the guesswork in likely outcomes"

In short, machine learning allows you to increase transparency around your business performance and reduce the guesswork in likely outcomes. In turn, you're able to plan and make business decisions with a far higher level of confidence.

And your members benefit too: customers can expect a more personalised experience as the business gets better at understanding who to engage with, with what, and exactly when. There's a decrease in noise as non-relevant engagements tail off, and a boost to the overall experience as the product becomes more refined to the requirements of the membership.

In short, any time a business can increase its confidence – by growing its understanding of its business performance and its customer experience – it leads to a better overall outcome for all.

💬 Key Takeaways

- Machine learning reduces the guesswork in likely outcomes, so you can plan with more confidence.
- Results can be transformational: one client increased average member lifetime by a full year.
- Your members will benefit too, as you'll learn to personalise your product to their needs.

How do I implement AI in my business?

The answer to this question will depend on the resources you have to hand.

If you have a strong technology department stacked with Python developers and data scientists, then availing yourselves of some of the Open Source AI tools available – such as TensorFlow from Google – will get you started.

For the 99.9 per cent of operators who don't, read on.

Do you have a problem to solve?

I see too many examples of AI solutions being asked to look for a problem. Likewise, I see operators enamoured of a concept (by which I mean the buzz around artificial intelligence) without ever being likely to see value from its implementation.

As we've discussed already, there are plenty of areas with potential, but you need to ask yourself some important questions before you take the leap, namely: "If I had this insight/value/prediction, would it be valuable?"

> "There must be a willingness not only to dedicate the time to implement the platform, but to then engage with it"

Do you have the time for implementation?

The results from properly applied data + AI projects can be exceptional, but the initial pay-offs over the first six months are marginal, and it's normally 12 months before we see the true justification.

To get to that point, there must be a willingness not only to dedicate the time to implement the platform, but to then engage with it and potentially change many of your workflows once armed with the new insight.

During the early days of Keepme, we found that after delivering a Keepme Score on every member – effectively giving our customers a window into who would leave and

who would stay – fewer than 20 per cent would do anything with it.

Let's consider that for a moment. These operators could now see, often with a runway of more than nine months, which of their members would not be with them when it came to renewal if they did not take action. Yet 80 per cent chose to do nothing with that information.

I was perplexed. I knew our sector was prone to more talk than action when it came to retention, but I was convinced this would not be the case once an accurate insight was delivered. I reached out to the customers and what became clear was that the insight didn't fit their process. Meaning, being told that Member A would be gone in six months if they didn't take action was valuable, but they didn't have a workflow to do anything about it.

> "The value is never in the insight. The value lies in the action that's taken as a result of the insight."

What are you going to do with your results?

You may feel this is a simple question, but with the above scenario in mind, it's clear the answer to the question 'what are you going to do with your results?' isn't always obvious – not even once equipped with a valuable prediction.

The answer must be this: You must have (or if you don't, then put) a process/workflow in place to take advantage of your new insights.

- If I know I'm going to miss sales targets in three months because of poor performance in my social channels, what can I do to remedy that?

- If I know which of the members who joined in the last three months have the highest probability of leaving, what will I do to change the likely outcome?

An AI implementation will give you access to a whole new level of transparency, but to gain value, you need to be very clear about what action will be taken as a result of it.

Put another way, the value is never in the insight. The value lies in the action that's taken as a result of the insight.
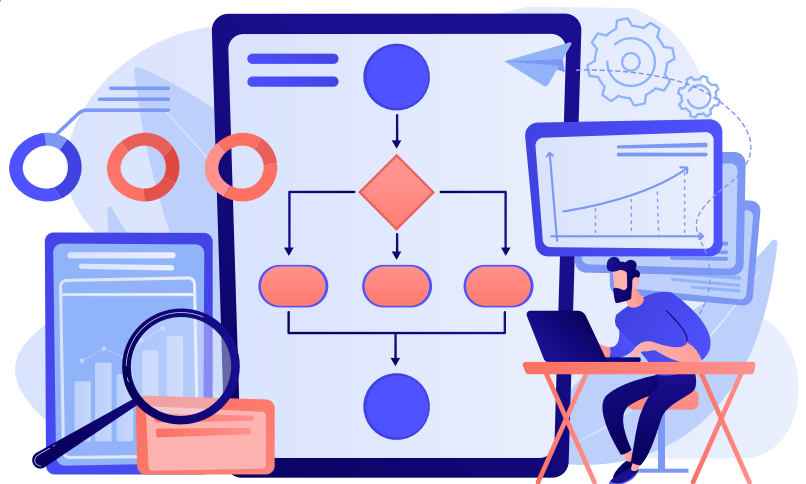
Will you add automation?

In my first paper – *The Fitness Future: Rules of Engagement* – I mentioned the power of AI and automation. Personally, I'm of the belief that processes fail mainly because they have a human element within them. I've lost count of how many times I've had a process for retention or sales explained to me, only to subsequently find that – due to the need for human intervention – it's rarely executed. I feel a whole paper coming on around the use of automation!

When you decide to deploy AI, understand what insights you'll receive and work out what actions you want to be taken with them. Then pause and consider this: Can (any of) these be automatically actioned, rather than requiring a team member to be engaged?

> ## "My advice: AI + automation is where the real results are"

This can be as simple as automatically generating a call list – all members who have become at risk in the last seven days, for example – and sending it to Member Services, who then call those members. At the next level, it can be the automatic serving of digital engagements that present customised offers to those members most likely to purchase them.

I've seen these and many more implemented by fitness operators, as well as across other services industries. My advice: AI + wautomation is where the real results are.

Who will you work with?

Before you go rushing off to find yourself a data scientist, may I suggest the following: that you first take a look at what's already available for the sector. The best solution will always be your data fed into fitness operator-specific applications.

This doesn't mean you'll end up with some cookie-cutter model similar to everyone else, as your data will drive the results. However, using a platform that's designed for the sector will ensure the downstream effects – i.e. the value that flows from the model – will be relevant and usable in your current environment.

Beware, however, the numerous solutions passing themselves off as 'AI-powered', or with AI in the brand name. Given my interest in this field, such solutions understandably catch my eye and lead me to investigate, only to find the applications in question deploy little to no AI of any type, or else they're chatbots.

💬 Key Takeaways

- Know the problem you want to solve, the questions you want answers to.
- The value is never in the insight. It's in the action that's taken as a result of the insight. You need to dedicate time.
- AI + automation is where the real results are.
- The best solution will always be your data fed into fitness-specific applications.

Summary

The question is not if data and AI will have a material impact on the fitness industry. Neither is it when. The impact is here, now.

The very term 'AI' can be a barrier to some who attach it to the sci-fi of their youth, or at the very least see it as a concept beyond their capability or needs. That would be a grave mistake. Just as the internet moved from being the domain of the military and academics to secure commercial acceptance, so AI is now widely available and in mass use – and, says Google CEO Sundar Pichai, is a force that's "more profound than fire or electricity".

Indeed, even before we saw the entry of digital players such as Peloton and Apple Fitness+, there were operators using their data + AI to gain deep insights, personalise the member experience and gain competitive advantage.

And consumers aren't just open to this – they expect it. We're so surrounded by technology that we don't think twice as we summon Siri, binge on Netflix recommendations and follow suggestions on Tik Tok and Twitter. AI is now so entrenched in our lives that we'd feel a jolt were it to be removed.

For fitness operators, then, there's never been a bigger opportunity to drive improvements in both our product and our bottom line. As an industry, we need to proactively adopt the benefits AI presents: to improve member experience, drive

business performance, and build a bridge that allows customers to feel the same level of personalisation in-person, at their club, as they already experience via myriad digital fitness offerings. (See my previous white paper – *The Fitness Future: Rules of Engagement* – for more insight around gyms' role in a blended, consumer-centric fitness ecosystem.)

This is the only way we will keep up with the digital giants. You might simplify Peloton to a connected exercise bike, or Fitness+ to digital classes; many do just that to ease their minds. But the reality is that these organisations are data juggernauts, consuming vast quantities of the stuff each and every day and harnessing it to enhance their products, improve engagement and fine-tune the user experience.

You certainly won't see these megabrands relying on already-engaged customers: the individuals they know will provide the user feedback or Net Promoter Scores they want. They use all the data available to build a detailed, honest understanding of business performance, and they have an unrelenting focus on improving the personal experience of each customer. And it's the use of AI that allows them to do this, and to do it at scale.

> "Data is the fuel in this new world; whether you choose to engage with AI now or later, you will only be able to do so if you have data"

I'm not suggesting implementing AI for the sake of having the latest technology. What I am saying is that if you want to improve retention, drive non-dues revenue and increase overall member engagement, AI will allow you to do so whether you have 1,000 or 1 million members. (Incidentally, it's also likely to be the antidote to the sector's stalled penetration rates.)

And data is the fuel in this new world; whether you choose to engage with AI now or later, you will only be able to do so if you have data. What you have may not be perfect (although it can be improved), but the first step is to acknowledge its importance. From this day forward, nurture and protect the data you do have, so it can be harnessed to deliver new business value.

I implore everyone to understand the benefits and plan to leverage this opportunity. Let's ignore the competitive threats presented by digital-only operators for a moment. Quite simply, if we keep doing what we're doing, we'll keep getting what we're getting. Add data to AI, on the other hand, and we enter a whole new world of opportunity.

Ian Mullane
Founder & CEO, Keepme

KEEPME LTD

71-75 Shelton Street, Covent Garden, London, UK, WC2H 9JQ • VAT Reg No. 334 6764 83

CONNECT WITH US

contact@keepme.ai or visit www.keepme.ai

in linkedin.com/company/keepme-ai   twitter.com/keepme_ai

keepme