The Insurance Industry's 'Buy vs Build' Dilemma for AI Solutions

By Tomas Vykruta, CEO & Co-Founder, EvolutionIQ

Executive Summary

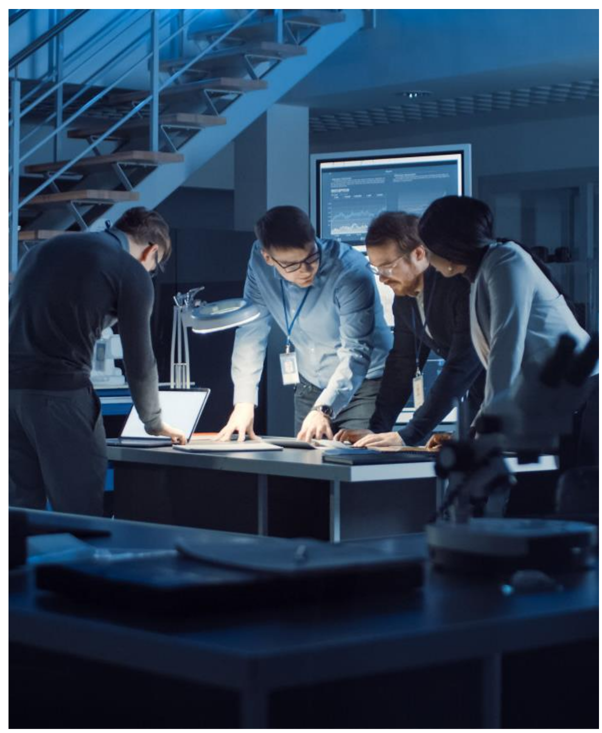
The 'Build vs Buy' dilemma for AI-powered solutions is a tough one for business leaders. This paper by EvolutionIQ's CEO and Co-Founder outlines the many steps required – and the many risks and pitfalls – that insurers will need to address when making the decision on whether to build an in-house AI platform.

The challenges include

- The substantial costs of the investment
- That as many as 87 percent of AI projects fail to launch
- The need to augment proprietary carrier data with external data sources
- Examiners and adjusters are not trained to adopt new technologies – and they approach such initiatives with skepticism at best, or often with outright rejection
- How marginal returns in AI projects are nonlinear, with the last 5 percent of engineering effort producing most of the value
- That long time horizons are required before prototypes can even begin to be tested
- The need to prepare to invest 7-figures and plan to see no return for up to two years
- After launch, a dedicated team of engineers is needed to continually fine-tune and calibrate the model
- The need for builders to develop a framework of evaluating risk vs success across the full life cycle of the initiative
- How machine learning initiatives have distinct tailend risks not present in traditional software development
- How adoption by frontline users can be a major stumbling block if the software does not immediately perform as expected
- The need to plan for ongoing fine tuning and calibration due to 'data drift', in which statistical properties in the model can change in unforeseen ways as new data continually hits the system

As Tomas Vykrutra points out, there is a need for urgency. "Advances in machine learning in just the last five years have created a situation in which the insurance industry – which has long been underserved by advanced technologies – is about to have its own Uber moment as it undergoes a once-in-a-generation shift to replace outdated claims management methodologies."

That means the decision about build vs buy isn't just academic – it's one that C-suites across the industry are making now.



Tomas Vykruta's career has spanned more than two decades within elite technical organizations. Most recently Tom was a leader within Google's Applied Machine Learning organization, leading teams from Mountain View and New York over the last 8 years as Google embraced the big data revolution.

AI Projects Are Very Different From Traditional Software Build-Outs

The 'Build vs Buy' dilemma for AI-powered solutions is a tough one for business leaders – even for those with deep pockets, significant scale, and technology teams packed with cutting-edge talent.

The reason it's so tough boils down to the fundamental nature of artificial intelligence and next generation machine learning projects: They don't behave like traditional software build-outs – and that 'bad behavior' can be quite a shock to those not prepared for it.

Taking insurance for an example, a company wanting to build a system to digest and understand the data in its vast claims blocks – and then deliver actionable insights from that data to reduce losses and cycle times – would first have to deal with the substantial cost. To develop such a technology may appear within the capability of a large, well-funded enterprise with experience in launching internal IT solutions. However, time after time studies show that as many as 87 percent of these projects fail to launch[1]. And given that an investment of this magnitude by definition means it's a priority project, a failure rate this high understandably gives pause.

But beyond the significant cost hurdle, there's the core issue of data.

In general, across any industry, AI is the ability for computers to think like a human and perform tasks in real-world environments on their own.

"Time after time studies show that as many as 87 percent of these projects fail to launch."

Machine Learning (ML) is an advanced branch of AI that mimics human reasoning by using a neural network – which is a series of algorithms modeled after the human brain – to identify patterns, make decisions, and improve themselves through experience. An ML project can only solve the problem at hand properly if it's seeing – and learning from – multiple carrier data sets and, importantly, learning to train itself to act one way in certain cases, but not in all cases.

[1] https://www.forbes.com/sites/enriquedans/2019/07/21/stop-experimenting-with-machine-learning-and-start-actually-usingit/

It is through interacting with a multitude of unique data streams that the system constantly improves.

For example, imagine your doctor devised treatments for you informed by an algorithm that had read and analyzed millions of medical records from other patients. You'd likely be happy that you're getting that kind of informed opinion, right? Well, that is until you are told that the millions of patients in that data set are exclusively white males aged 25 to 50 – and you're a 60-year-old Asian female. In that case, the doctor may get the diagnosis and cure right, but then again, he may not, given the non-varied data set.

Likewise, insurance carriers can augment their vast data stores by purchasing external data to overlay on their own, but they cannot augment it with the learnings inherent in predictive platforms that come only from having analyzed tens of thousands of claims that lie outside their own organization. What's important to underscore here is that these ML platforms are not exposing or sharing competitors data in any way as external ML platforms are bespoke to each carrier and siloed. However, the way in which the underlying system learns and evolves is shared. And it's that institutional knowledge that is nearly impossible to replicate in-house. For example, Tom Brady brought two decades of quarterbacking knowledge to Tampa Bay – but he didn't bring the New England Patriots' playbook or intellectual property with him. He simply brought the way he plays, the way he trains, and the way he learns. ML systems do the same thing client to client.

But suppose an executive team said yes to the cost – and found tremendous sets of external data that could be overlaid on their own to generate the requisite system learning. Then there's the issue of marginal returns. While the ultimate ROI is very high, marginal returns in ML projects are nonlinear – with the last 5 percent of engineering effort producing most of the value.

In my own work in AI-powered claims guidance, our team of PhD data scientists from Google, Meta/Facebook, Bloomberg and other tech leaders could not solve the problem until the end of the first six months of effort, when we launched the first version of the software. Over 15 different approaches were taken, and only one worked. Although we had made progress, we were getting zero breakthroughs in terms of complex claims analysis and insights. Only at the end of the 14 month mark did the breakthroughs happen and then continue to cascade – and we are still making major breakthroughs that continue today with substantial value being delivered. But the ability to tolerate this kind of 'success-desert' for very long periods of time is essential – and not for the faint of heart.

That's why a key factor to success in ML projects is having an R&D group focused on problems that may not be solvable.

And allocating costly engineering resources in this fashion isn't how most insurance companies are designed to operate. This R&D group also has to focus on developing both the front and the back end of the product as the goal isn't to simply sell a software suite that can be handed off. This team needs to build a rich, completely custom, front-end user interface because the AI by itself, without this, is not useful. The real goal must be to develop a world class overarching solution – and one that must be constantly maintained and calibrated like a fine watch or Formula 1 racer.

"The ability to tolerate this kind of 'success-desert' for very long periods of time is essential – and not for the faint of heart."

A Framework is Needed to Evaluate Risk

A framework of evaluating risk vs success and understanding the full life cycle of such initiatives is an absolute must-have tool for executives tasked with funding such initiatives. To illustrate, let's use the example of how EvolutionIQ developed its ML platform for long-term disability mature claims blocks and look at three discrete phases: Initial development and evaluation of the model, launch and adoption with the examiner team, and ongoing support and tuning.

The initial model development typically requires extraction of key variables from normalized claims data – a labor-intensive process that must retain and utilize the vast majority of the information stored within the various active and historical data silos.

Upon completion of data acquisition, the team can then develop and evaluate the model itself and estimate business impact. An R&D team must be in place, and must have not only expert domain knowledge in the area of statistical applied predictive modeling with large, messy, industry data sets, but additionally, expertise in claims handling and the particular business pain point being addressed.

It took EvolutionIQ's team of five PhD artificial intelligence experts from Google and Bloomberg and 83 person-months to develop the first minimum viable business model at a cost of approximately $2.3M spanning the Q4-2018 prototype and the Q2-2020 launch.

> "Over 60 technical initiatives were developed during the R&D phase, of which the majority did not lead to model improvement."

Data sparsity led to an initial failure in acceptable model results. Only when multiple carrier data sets were analyzed – separately and with each carrier's proprietary data secure and siloed – were trends identified and the first working framework began to take shape. But as mentioned, the 14 month ramp-up period of the project was nonlinear – which comes with the territory in the world of technology startups, but can generate intense pressure on budgets, teams, morale, and business projections when they happen in-house in an established carrier.

Again, not an unimaginable scenario for an established insurer to endure, but certainly not the norm – and a path that would take steady nerves to see through to completion.

We found out firsthand that a series of highly proprietary technical breakthroughs were necessary to achieve a minimal viable model. It's worth noting that over 60 technical initiatives were developed during the R&D phase, of which the majority did not lead to model improvement. It was only through relentless experimentation that the desired methods and results were achieved. Again, not an impossible scenario to replicate inside a carrier's tech center, but not the norm.

What's critically important for decision makers to understand is that the tail-end risk described above is unique to machine learning initiatives and not present in traditional software development. This critical distinction must be taken under careful consideration when budgeting and staffing. It not only creates the risk of failure to produce a minimal viable model, it also creates uncertainty in when, and if, the initiative will produce results adequate for launch.

Adoption by Frontline Users is a Major Stumbling Block in Both Buy & Build Scenarios

The next hurdle to overcome is launch and adoption of the model. Examiners and adjusters, like other enterprise employees, are not trained to adopt new technologies – and they approach such initiatives with skepticism at best, or often with outright rejection. Overcoming this challenge is a high risk effort and requires the expertise of a highly senior project manager and a technical team that has a proven track record of deploying such products in complex enterprise organizations.

If you ask the chief technology officer of most insurance carriers, they'll readily tell you they have excellent data scientists on staff and the analytics they deliver are on the cutting edge. True. And if you ask those same data scientists how their analytics are driving results, the answer isn't so clearcut. Many will likely tell you that they're producing great models, but no one is using them. Or if they are being used, they're not at scale or part of standard operating procedures, so the results are confined to narrow uses – or not used at all. That's why even after the pain of developing an AI-driven claims model from scratch, the whole ball game can be lost if the means to ensure adoption of the technology are not pursued just as vigorously.

EvolutionIQ's team already had a combined 53 person-years across over 170 successful deployments (and countless failed deployments) when work on the disability module started. The risk is too great to take a chance with an inexperienced team.

> "While it takes years to develop, launch and tune such a model, it can take just months for complete abandonment and turn-down of such a service."

If the team fails to inspire and gain trust of the examiners and adjusters during this critical phase, the project will fail to launch and a second chance will not be possible for several years. This is true whether it's an outside technology or one developed in-house.

Following successful adoption of the technology, model tuning and improvement can begin. EvolutionIQ, during the 6 months following the launch, continued heavily investing in R&D and achieved an additional 3X improvement in acceptance rates compared to the initial model. This required the careful cooperation and trust of the disability examiners as partners with the technical team.

Finally, ongoing maintenance and improvement is required following this tuning phase. Time after time, we've seen teams move on after a successful launch. But without expert caretakers (unlike with traditional software), machine learning solutions quickly degrade. 'Data drift' – which is where new data hits the system and the statistical properties start to change in unforeseen ways – and other problems are unavoidable and must be addressed by continual calibration of the model. If these factors result in a reduction of the software's business impact, the expensive new tools are usually quickly abandoned by busy teams.

While it takes years to develop, launch and tune such a model, it can take just months for complete abandonment and turn-down of such a service. Even two years after launch, EvolutionIQ continues a material R&D effort and continues to deliver new breakthroughs in business impact. This investment is only feasible due to multiple carriers simultaneously using the platform, thus, distributing the cost of the expensive AI team.

Economies of Scale

Given this advice and hard lessons learned, why wouldn't a well-funded insurance carrier simply bite the bullet in advance and say, "We're going to invest 7-figures and plan to see no return for up to two years. And when we finally nail the system, we're going to have dedicated engineers to continually fine-tune it?" The reason is simple: Economies of scale.

In the race to develop AI-driven software specifically for insurance claims, the heavy trial and error lifting that comes with creating an entirely new class of ML software has already been done. So even the best-intentioned in-house developer will find it difficult, if not impossible, to create a home-grown AI that delivers measurable results because of the wide moat that first-movers have already created.

And even if this moat could be overcome, the data issue remains – meaning the carrier's in-house technology will likely never be able to learn at the rate of external technologies that have the benefit of continually recalibrating based on widely varied data sources.

In essence, just like an automotive factory has economies of scale that make it cheaper to build a vehicle than building bespoke cars, machine learning has its own economies of scale that come from exposure to new data, and practice with it.

"The insurance industry is about to have its own Uber moment as it undergoes a once-in-a-generation shift to replace outdated claims management methodologies."

At EvolutionIQ, we estimate that clients receive a 7x to 10x return on their investment in the platform, which measurably flows to the bottom line – such as through reduced durations and losses, and through the positive impact on staffing. Examiners and adjusters have higher job satisfaction as they focus on the most complex claims, which in turn lowers attrition as burnout disappears. There's also faster upskilling of new claims team members as they have the benefit of an AI co-pilot that has years of institutional knowledge – and information about cases that need immediate attention – ready to help 24/7.

The Tough Decision

Most predictions were that Elon Musk was out of his mind to try and build a rocket company from scratch. Doing the impossible has an irresistible allure for some – and it’s those visionaries who actually get it done. But they don’t come along that often. So the real question in insurance becomes: Can we pull an Elon and build from scratch – and then actually pull it off? Or do we future-proof our business now, with proven technology already in place?

Advances in machine learning in just the last five years have created a situation in which the insurance industry – which has long been underserved by advanced technologies – is about to have its own Uber moment as it undergoes a once-in-a-generation shift to replace outdated claims management methodologies. So the Elon question isn’t just academic – it’s one that has to be asked immediately by businesses – or risk playing a forever game of catch-up.

That’s why the “buy” model of AI for predictive modeling in insurance claims organizations is proving so successful now. It eliminates multiple risks completely while having measurable ROI immediately. Importantly, a buy decision also translates into speed, as an insurer would see high impact results in live claims in an early as three months by using a proven technology.

About EvolutionIQ

EvolutionIQ is the market leading claims guidance platform in Group and Individual Disability, Property & Casualty, and Workers’ Compensation lines of insurance. Our proprietary Artificial Intelligence uses the entire claim file contents, historical claims, and external data to guide claim handlers to their most productive task across the entire claim block, every day.

To request a demo email us at sales@evolutioniq.com