muse ADVISORY

20 years of independent thinking

INSIGHT

AI Governance

Jo Fellowes 7 Minute Read

Jo Fellowes explores where AI is currently being used and developed in the pensions industry, and considers the changes in Trustee governance required to oversee data security and cyber risk as AI use continues to evolve.

Artificial Intelligence (AI) is driving massive change across many industries and businesses. Whilst there are significant potential benefits available, the risks are real as well. Although our industry has never been one to lead in terms of the use of cutting edge technology, AI is already being used or developed by many DB and DC pension scheme administrators to manage the following types of activities:

- Data collection - sourcing missing data, e.g. from imaged files.
- Data cleansing - filling gaps, formatting the data.
- Data analysis and evaluation - the significance (so what) of any patterns, trends and anomalies, evaluating and producing insights.
- Data management action planning - identifying the next steps/ options as a result of the evaluation, although there will likely be other considerations beyond just data, including impact on liabilities, insurer risk premium, legal view on whether the action is proportionate, etc.
- Managing member enquiries - the use of web chat/ chat bots to answer member queries, responding to information requests from DC members at-retirement, or to answer a phone call.
- Call monitoring and insight - providing quality assurance overlay e.g. checking that the right security questions have been asked and passed on each call, analysing member sentiment, predicting member behaviour.
- Workflow generation - assessing and logging cases following written or email communications from members or their representatives.
- Document interrogation - sourcing knowledge from scheme documents/ Trust Deed & Rules e.g. during transition, to produce benefit specifications, following unusual benefit queries etc.
- Administrator training - via monitoring administrator key strokes and optimising paths, creating AI agents who can deliver training and test administrator understanding, and providing a quality assurance overlay and real time feedback.

We should recognise that in the administration industry, not all that is being labelled AI, really is. Often, it's a really good step forward in automation, using a programmed set of parameters (if this, then do this), but not all of it is 'self-determining'.

AI is also being used and developed by many advisers. Whilst there are some areas of overlap with administrators, for example document interrogation, and production of insights, some advisers are also developing or using AI to support:

- Minute taking
- Drafting template policies, letters, reports, etc
- Preparing for meetings - for example by producing brief summaries of all previous discussion notes and emails and/or summarising key messages in pre-reads.

Supplier selection and oversight

As AI evolves, you may wish to consider how it factors into your conversations with your administrator and other advisers. And if you're selecting a new adviser, you may want to explore their AI capabilities and development roadmap as part of your selection process. To assess where you are now, can you answer the following questions about your administrator and advisers?

Q: Do you know whether your administrator/ adviser has an AI solution in place, or under development, and for what purposes?

Q: Do you have assurance that that AI solution is a segregated instance of AI, rather than Open AI, and that it adopts any UK or EU data residency limitations you may require?

Q: Do you understand how that AI solution is being used to deliver your services, and how that usage is being governed?

Q: Do you understand what your administrator or adviser contract says about AI usage? For example, whether your supplier can use AI to service your scheme, now or in the future, without telling you.

Would you want to know or have a say in the decision to use AI, in any or all circumstances? Who is liable if AI goes wrong?

AI solutions

Open AI – using models such as ChatGPT, Copilot (Microsoft) or Gemini (Google), etc., the AI is hosted in the provider's cloud and any data shared with, and interrogated by the AI leaves your personal or organisation's environment, and helps to inform AI models and results going forward. Information shared would not be restricted to the UK or EU. The results you receive from open AI can be based on all other accessible data and prompts (subject to how you have specified your own prompt), which could render some results unreliable.

Corporate segregated AI – uses dedicated licensed AI models, which run in a private cloud. This means the data is isolated, is not trained on public information or prompts, and as such the model can be fine-tuned to the corporate's own needs. The data residency of the AI model is determined in the corporate's enterprise agreement with the AI provider. The agreement may also specify where backups are stored (if in another region) and whether engineers from other regions could access the systems. The model requires robust governance and user training.

On Premises – uses AI models run on internal servers, inside the organisation's data centre. The organisation manages the hardware, storage and security – providing maximum control and data privacy. It comes with a high upfront cost, plus a need for ongoing maintenance and in-house expertise, along with the need for robust governance and user training.

Q: Do you understand whether other open AI channels can be accessed by employees through your administrator's or adviser's IT infrastructure, e.g. through an app download or an internet browser, and how the supplier governs that usage?

An AI policy, if one is in place, and training, can help suppliers manage their employee AI usage.

This is important because even if the supplier has a dedicated instance of AI, employees will still be able to access open AI (if not through the corporate IT infrastructure, then through their personal devices).

What does governing AI mean?

Visibility: Transparency is key. Whether a scheme is implementing a data management tool, a chat bot, or an automated calculation, understanding what drives the algorithm is of utmost importance. There needs to be documented evidence of how algorithms operate to ensure that bias is not creeping into any decision making process.

Taking the example of data cleansing in the previous section, visibility is about knowing:

- What rules will AI follow, and who will determine them/ sign them off?
- How will they be coded?
- Are there any areas of ambiguity? How will they be treated?
- What sampling/ testing of the application of the rules will be undertaken?

Taking another example, minute taking:

- How is sensitive information kept confidential i.e. to what extent is the Trustees' recording kept private to the team working on the client account or the business unit/ region that the team sits within?
- If a global organisation, will recordings be held on servers outside of the UK (or EU)?
- What is the retention/ deletion policy in terms of transcription notes and recordings? What, if anything, will the AI remember even after deletion of a note/ recording (e.g. will it have learned voices)?

Robust data and processes: An AI tool will need a large amount of data or information to make sure that it is accurate. In order to interrogate the right data, and to integrate new data sets, robust document and data management processes will need to be put in place.

Taking the example of data collection above, robust processes would consider:

- How will you know AI is collecting the right data? Are the labels consistent (both of the file being accessed - such that it related to the right person, and of the data being accessed)?
- How will you know AI transposed the data correctly? e.g. if handwritten
- Under what circumstances will AI determine it can't find the data?
- How will you sample check/ audit what it does and doesn't find? What is the pass/ fail rate for errors, to trust the un-sampled data?

Monitoring: Oversight of the robustness and safety, privacy and data governance, transparency, diversity, non discrimination and fairness, societal and environmental wellbeing and accountability is in place.

Trust: Bringing stakeholders on the journey is going to be imperative with clear documentation and training available, so that everyone can understand the implications of what is being implemented. As humans can make errors, so can AI. This means robust audit practices and procedures should be built into any AI governance framework.

Training is important, as employees need to understand the purpose for which AI has been made available, including what data it has access to. The quality of results generated through AI is also dependent on the quality and specificity of the questions asked.

Big brother tools are also available to monitor employee AI usage on corporate devices.

What does the general code say about AI?

The General Code requires Trustees to have appropriate IT systems for record keeping (ensuring data is complete and accurate) and financial transactions, and to consider cyber risk and operate appropriate controls to mitigate cyber risk. Although Trustees aren't required to have a specific policy on AI, schemes should include AI in their cyber policy.

AI risk

AI is unlikely to be on your risk register as a risk in its own right, but it could give rise to new causes for your existing risks, including cyber security, supplier disruption, fraud and member misinformation. The question for each of these is does this new cause increase your risk exposure (the impact and likelihood) and do you need new controls/ assurance (e.g. safeguards against improper AI use, attack or poisoning, or against implementing instructions from deep fakes)? And will AI create an overdependence, leading to a decline in critical thinking/ understanding?

Additionally, AI can also provide new or enhanced controls, that reduce the exposure of your current risks (e.g. contact centre quality assurance).

And AI may provide opportunities to reach your strategic objectives faster, by providing a solution to speed up or make certain projects more efficient or affordable (e.g. data mining or insights).

Trustee Director's personal use of AI

In addition to governing administrator and supplier AI, Trustee Directors may have access to a corporate instance of AI, meaning any Trustee papers shared with that instance (for example where AI is being used to summarise key messages in the Board or Committee papers) might result in information becoming available to parties who should not be privy (e.g. minutes on the Trustees' stance in a valuation negotiation or a transaction). This can be mitigated with creation of special permissions within a sponsor instance, but only if this need is known, accepted and set up by the sponsor, and you'd need to be comfortable that any sponsor 'super user' couldn't later override them.

The risks increase further if Trustee Directors were to use an open AI for the same purposes, as they would effectively be putting those Trustee papers, and any confidential information contained therein, in the public domain. Whilst most Trustee packs no longer include any identifiable member information, such as for discretionary cases, this isn't yet the case across all schemes, and sharing such information with open AI would be a GDPR breach.

Training Trustees on effective and secure use of AI will certainly be required.

Will AI bring down the cost of administration, or adviser fees?

That depends on how and where it is being used. Going back to some of the examples:

Data collection from imaged files – whilst in its infancy it may be intentionally over cautious, and require more assurance/ validation, however, it's still likely to be materially cheaper than a human manually reviewing all files.

Member phone call analysis – AI increases the number of calls being assessed for quality and sentiment review, from a random sample to all calls. This is unlikely to reduce cost, as it will redirect resources as an administrator builds capabilities that proactively intervene early, e.g. to enable live support or transfer of a difficult call, or to capture feedback and intervene swiftly if a negative sentiment is detected.

Insight and reporting – Increases to the data insights available to schemes. This is unlikely to reduce cost, although it should support improved decision making which may create cost efficiencies in other areas. Direct costs may even increase if Trustees ask for a 'non-standard' version of the insights.

Overall, we don't envisage AI will drive administration or adviser fees down materially, but its use, when controlled and governed effectively, can help to maintain fee levels, add risk controls and expedite some project activities, which may lead to reduced costs for the scheme overall.

Summary

AI is bringing massive opportunities to improve efficiencies in the pensions industry, and can support member engagement, but it needs to be implemented in a thoughtful way, that protects schemes and their members. What is clear is that good governance needs to be at the heart of any AI implementation from day one.

If you'd like help overseeing your administrator or adviser's adoption or AI, or integrating the impacts of AI into your system of governance, get in touch with:

Jo Fellowes
Director
Muse Advisory

jo.fellowes@museadvisory.com