



Checklist for an Organisational **Generative AI Policy**



Why develop a separate policy around these tools?

At the IoA, we strongly recommend creating a unique policy around the use of generative AI tools. Although many organisations may well be mature in terms of their ethics approach to analytics, the large language models (LLMs) are creating a host of new issues and concerns that may not be covered under a more generic ethics code. They may have the surface appearance of familiar tools, such as previously seen chatbots, voice assistants or search engines, but they have a different technical process 'under the hood', and therefore will behave in unexpected ways.

As well as promoting trust in the machines and safeguarding valuable resources in the company, clear guidelines will also encourage confident uptake of these tools to streamline work and drive efficiencies.

Use the checklist below to help shape and review your proposals to encourage confidence in generative AI tools and encourage ethical and safe uptake of these new technologies.

1) Include a brief definition of the kinds of tools that are involved in generative AI technologies

There are some obvious ones that have been in the headlines lately, like ChatGPT-4, but there are many tools built on the large language learning model technologies that include data not often considered to be language. Make sure your policy covers text, image and voice tools and others that might be specific to your field of work (eg tools that work with DNA, MRI Scans, music etc)

2) Consider pre-existing legal constraints around the use of these technologies

Some laws already exist to restrict what can and cannot be done with black boxed technologies like Transformers. For example, Article 22 in the GDPR forbids the use of any black boxed technology in a situation where the outcome will have a real-world impact on people (such as identifying them in police CCTV footage). List all the ways that you might use the technologies. What are the consequences? Are they reversible?

3) Make space for your policy to cover emerging risks

The speed of development in these generative technologies is exponential, and we are witnessing a race to get tools and services to the market right now. This means several things: the first is that there are a large number of unknown unknowns around these technologies. They are being released without full testing and the nature of LLM technologies is that the developers cannot stop the machines from giving the wrong advice, saying the wrong thing or producing random and unhelpful results. We are expecting a flood of ethical issues to be reported as these technologies become adopted and put into practice. Because the scale of development is so rushed, many governments around the world have either already banned or are discussing legislating or creating regulations around these technologies. We also expect to see legislative changes coming in the very near future. Make sure that flexibility is built into your policy to respond as needed.

4) How will employees raise concerns over the use of generative AI?

It is our strong recommendation that all use of generative AI is overseen by a suitably qualified human. If human oversight is native to your policy on the use of generative AI from the outset, reporting should not be an issue as the humans will have the authority to override any machine proposals or suggestions. If you have any other model where the production process is in any way automated, ensure that you have clearly communicated how to raise concerns at the earliest instance to the teams involved. We are braced for a large number of unanticipated consequences of using this emergent technology. Ensure that your pipeline for receiving, escalating and reviewing concerns is in place before you launch your new policy.

5) Ensure that everyone knows the risks of embedding these technologies into products or services

Some organisations may want to embed these technologies into services. For example, they may wish to embed some kind of recognition technology or an intelligent agent into machines that they are designing to interact directly with customers. Be aware of the risks of using these machines. Ensure that you inspect training data that you have developed yourself rigorously for bias. Never extrapolate the results of one test to a wider population without solid grounds to do so. Check any claims of generalisability from 3rd party providers against your own use case. The margin for error when you remove direct and immediate human oversight of these tools remains very high while we are all in the early adoption phase.

6) Who wrote the policy?

Make sure that you have shared your policy with a representative sample of your own team. Check that they understand it. If appropriate, share your draft policy with potential stakeholders, such as trusted customers. Is there anything in your policy that surprises stakeholders in a negative way?

By welcoming participation, you involve the wider team in what will become an important way of working for them going forwards, they will feel a greater sense of ownership and will be more likely to adhere to the ground rules that you are laying out.

7) Where will the policy be kept? Is it easily searchable?

If the policy is inaccessible, people will not seek it out. If you are reliant on sharing it as an attachment in one email, staff will have trouble retrieving the document somewhere further down the line. Ensure that your policy is easy to find in an online shared drive or your company's internal (or ideally external) website. Make sure that the key terms are easily searchable in the document for quick checks to support adherence.

8) How frequently should the policy be reviewed?

We strongly recommend short review periods on the policy for the time being as the exponential curve of these technologies is mystifying to even those with a strong understanding of the technologies behind the tools. We recommend finding a channel to stay up to date with the emergent issues, guidelines and legislation while these technologies begin to embed themselves in our work processes.

Unsure where to start?

We're happy to help. Please refer to our own IoA Generative AI policy for general common risks and best practices working with these technologies. We have a list of use cases when these technologies should NOT be used, and some suggestions for how they could streamline processes and push efficiencies.

About the IoA

The Institute of Analytics (IoA) is the leading global body for Analytics & Data Science professionals. We promote ethical data practice and evidence-based decision making to transform business and society and help our members stay ahead of the curve of digital reform to thrive in the data age.

ioaglobal.org

hello@ioaglobal.org