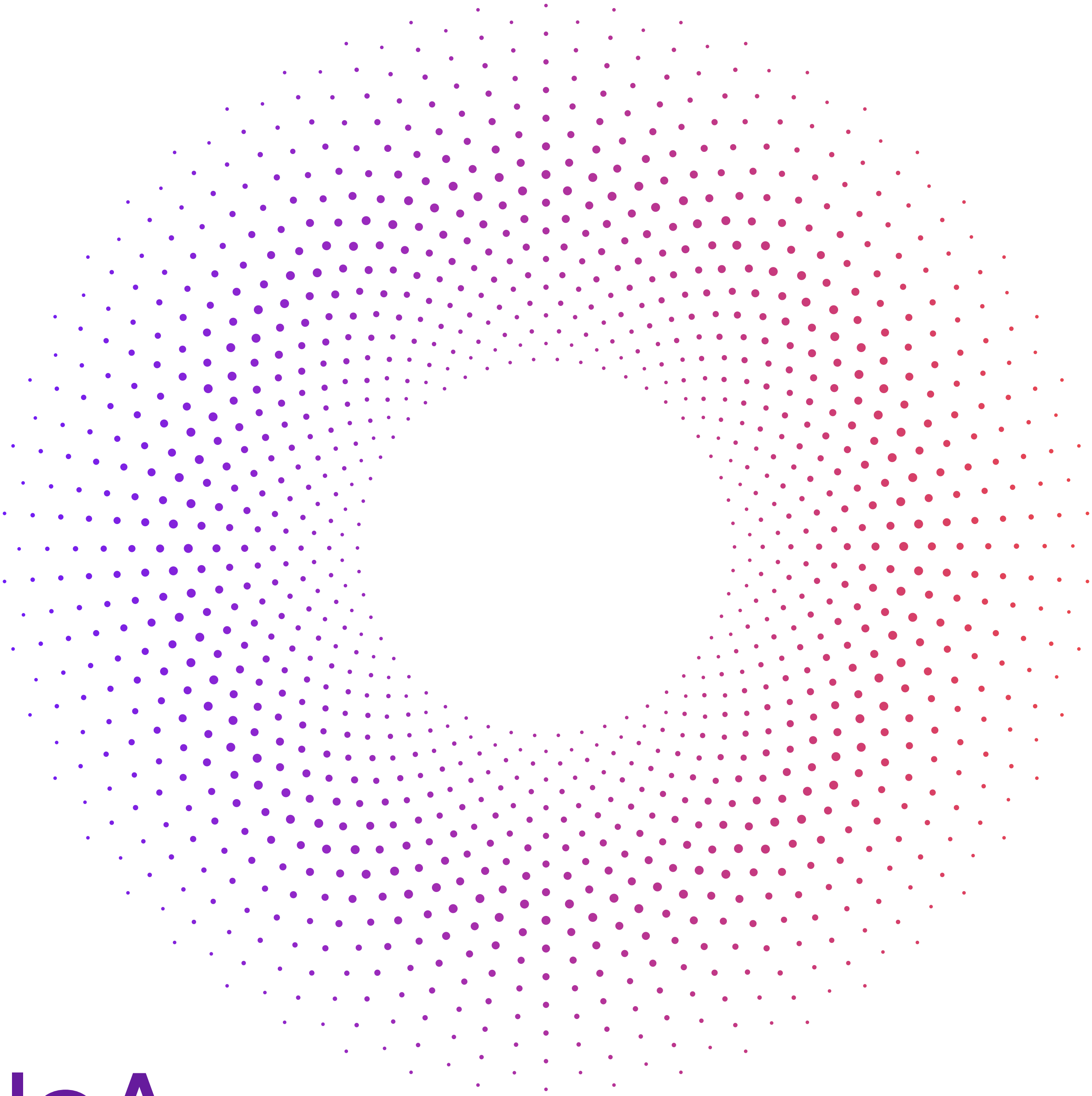
Institute of Analytics

The Global Body for Analytics

ioaglobal.org



IoA GENERATIVE AI POLICY

24 April 2023

Following recent improvements in the technologies behind text, image and other types of generation, we have created this generative AI policy to guide our staff in the use of new tools.

Note: The growth in these technologies currently goes beyond exponential and it may be necessary to update these guidelines as new issues emerge.

Table of contents

1. About generative AI and the technology behind it

Generative AI is any type of AI system that can generate text, images or other types of media in response to prompts. At the time of writing, these tools use large language models, which produce a result based on a set of training data. The technology is not the same as search engine Eigenvector algorithms, meaning that the results may be even less reliable than a Google fact search. The technology is not the same as traditional Natural Language Processing (NLP) chatbots, which often had human input to optimise the results for a particular use case.

The training data is not usually linked to the internet or any other real-time updates in most models. Therefore, any content will have no knowledge of events since the model's last extraction date, which can be anything from months to years old.

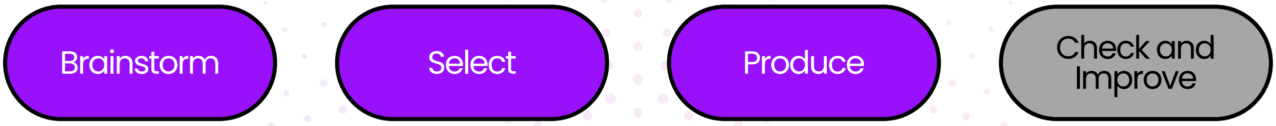
These are narrow AI uses, meaning that a 'text generator' will produce beautiful natural language but is not optimised for factual accuracy. An 'image generator' will produce an impressive image but will have no ethics controls. They are not general AI models that mimic human thought, models of human ethics etc.

1.1 Streamlining processes

These technologies streamline and improve processes and therefore we welcome plans to incorporate them into work in productive and ethical ways.

1.2 Pre-generative AI

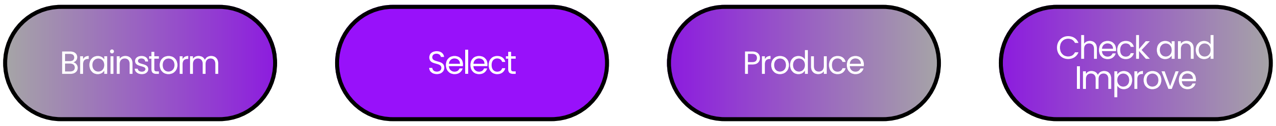
We have traditionally brought in machines towards the end of processes. If we imagine the humans involved in the production of a new process, represented in purple below, the machines might be brought in at the end to help to present the information in a more professional way, through spellcheck, suggested grammatical improvements, format and design.



Design stages without generative AI, with human-based processes in purple and machine based processes in grey.

1.3 With generative AI

These new technologies prompt a different way of working with content. As shown below:



Brainstorm

Select

Produce

Check and Improve

Design stages with generative AI show shared roles for machines and humans, with human responsibilities indicated by purple and machine capabilities indicated by grey

We might use the machines to brainstorm suggestions before beginning a new project (eg show me 10 ways to optimise SEO) or to summarise the contents of large reports that cannot be manually inspected. These suggestions could be incorporated with human-generated content. The review and selection process must ALWAYS be carried out by a suitably qualified human. The production may be a mix of human and machine collaboration. For example, the generation of documentation to accompany a coding process would be done through inputting the human designed code into the machine, which then produces a natural language summary of what the code did. The human may write a report and then ask the machine to extract the key bullet points. The human may write content for a social media post and the machine will improve the language, suggest suitable hashtags and emojis etc.

Checking and improving in the final stages will be a combination of human and machine review. For example, a machine could generate slide show content from a report and suggest ways to visually display it effectively, but a human should also review the final result.

2. Guidelines for all uses of generative AI

Generative AI should NOT be used for the production of the following types of media:

2.1 When the results need to be accurate

The technology that generates the content could be wildly inaccurate. We do not know where the machine found its responses and so, if accuracy matters, human design and research should be used. (NB some of the emerging tools may explain their thought process with verifiable data trails, which have more accountability but should still be checked)

2.2 When ownership is an issue

Copyright over ownership of anything produced by generative AI remains poorly defined but, as a user, you do not own the copyright on anything that you produce with these machines. The grey area is whether the owners of the data that the machines were trained on have any rights to the output. Therefore, if ownership is an issue and the output will be monetised in any way, avoid these technologies completely.

2.3 When the origin of the document is not stated

Whenever machine-generated content is the predominant method of production, it should be stated on the media object that the content was generated by machine.

A predominantly machine-generated media could be:

- A slide show almost entirely generated by AI
- A summary of a human-written document that has not been checked for accuracy/relevance
- An entire article, report, etc. written by the machine

A non-predominantly machine-generated media would be:

- A social media post that was written by a human and copy-edited by the machine
- A summary or bullet points of a report etc that has been checked by the original author
- A slide show that has been generated from a report the author has checked for accuracy

2.4 When personal data or sensitive data is involved

Never put personal data (names, addresses, phone numbers) or sensitive data (data about someone's religion etc) into a generative AI tool. It is against the law in many countries and is unethical practice as once entered, it will then be shared as training data. Do not enter any text in the prompt box where proprietary ownership is important (eg do not enter any form of course content). We will lose ownership rights.

3. Guidelines specific to text-producing generative AIs

Those generative AIs that produce text response, such as the use of ChatGPT or Bard, create a number of additional concerns listed below.

3.1 Do not use these technologies when the results need to be accurate

The technology that generates the content works with a ‘best guess’ approach and you do not know where it got the information from. Some information will be valid, but it could also be wildly inaccurate. If accuracy matters, we shouldn’t use generative AI.

3.2 Do not use these technologies when the results cannot be/are not checked by a human

Text-based generative AI should never be used to complete a task that the human making the request could not do themselves. It is essential that all output be reviewed before incorporating the work into any kind of workflow.

3.3 Do not use these technologies when recency is an important variable

Unless the machine that you are using is in internet browser mode, assume that it can only incorporate information that is widely known up to the date of their extraction but not beyond that. The extraction date may be several months or years previous, and it is essential to use human alternatives when the main subject is emerging, such as new laws or new developments in technology unless you are in a specific browsing mode of AI.

3.4 Do not use these technologies when the contents cannot be verified

In general, text producing generative AIs should only be used to automate the text production that the human requesting text could produce themselves. Any output must ALWAYS be verifiable and verified by expert human oversight.

3.5 Do not use these technologies where nuance or depth of analysis is required

The machines may be able to generically summarise common ideas shared, but they are not able to wrestle with high level or original thought.

4. Guidelines specific to embedded AI

4.1 The use of embedded AI in IoA services

Embedded AI is the use of any AI tool, such as a chatbot, with any customer-facing service such as our website or training materials. If any member of staff would like to propose embedding an AI tool, this raises much more serious ethics questions and the proposition needs to be brought to the attention of the senior leadership team at the earliest opportunity.

4.2 Concerns IoA staff may have about non-IoA services that use embedded AI

We are here to support our staff during this difficult transition to a rapidly-evolving world of AI-generated content. If you have concerns about the potential of generative AI in the field of online harms, cyber security, trust in social media etc please raise your concerns in the first instance with your line manager and we will attempt to address these and support you during this transitional phase.

5. Suitable use cases

We welcome the use of generative AI to streamline operational processes and evidence our adoption of data-driven best practice. We recommend the use of generative AI in the following circumstances providing that the conditions above have been met:

- Producing social media content
- Producing images for visual effect
- Documenting processes (particularly coding or analytics)
- Supporting coding processes/skills improvement in tools like Excel etc through the use of OpenAI guidelines, tutoring support, breaking down code into natural language etc
- Producing presentation content from a pre-designed report
- Summarising reports that would otherwise not be read
- Brainstorming and planning new processes
- Brainstorming ideas for internal campaigns (eg to promote adherence to cybersecurity principles)
- Playing devil's advocate (by testing out the opposition to an opinion you strongly hold)