Generative Artificial Intelligence at Keble

Executive Summary

Generative artificial intelligence (GenAI) has been widely adopted over the last year or so, not least in academic contexts including at Keble College. The University of Oxford has issued some guidance around use of these tools, and has stated aims to increase "AI literacy" among staff and students, but nevertheless many practical questions remain open. This short report:

- (1) considers the University of Oxford's current position on *generative artificial* intelligence, and
- (2) presents guidelines, in the form of ten practical tips, that our staff and students may refer to when using generative artificial intelligence tools.

Examples of how GenAl tools might be used as part of undergraduate academic work, and a short guide to relevant terms related to generative artificial intelligence, are appended.

Dr Thomas Player Lecturer in Study Skills, Keble College January 2024

Introduction

Today, many Keble students and staff regularly use generative artificial intelligence tools (GenAI) as part of their academic work. This is a new situation:

- ChatGPT, an online text generation tool launched in late 2022, was by January 2023 the <u>fastest-growing piece of consumer software in history</u>.¹
- A <u>survey in April 2023 by Cambridge student newspaper *Varsity* found that almost half of Cambridge students had used ChatGPT to complete university work.²</u>
- In a <u>call for evidence run by the Department for Education in summer 2023</u>, 78% of responding educational institutions reported using GenAl in an educational setting.³

Many believe that GenAI will become an unquestioned fixture of the academic environment, in the same way that pocket calculators, spellcheckers, and search engines have. Already, almost any part of an academic task *could* be done using GenAI (although by no means *should* – see examples in appendices). This fast progress motivates this report, which aims to present guidelines to this new technology that Keble staff and students may refer to when using GenAI.

¹ Krystal Hu, "ChatGPT Sets Record for Fastest-Growing User Base." *Reuters*, 2 February 2023. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

² Michael Hennessey, "Exclusive: Almost Half of Cambridge Students Have Used ChatGPT to Complete University Work." *Varsity*, 21 April 2023. https://www.varsity.co.uk/news/25463.

³ Department for Education, "Generative AI in education. Call for Evidence: summary of responses." November 2023. https://www.gov.uk/government/calls-for-evidence/generative-artificial-intelligence-in-education-call-for-evidence

Section 1: GenAl at the University of Oxford

At the University of Oxford, <u>we summatively</u> assess the work of our students relatively <u>infrequently</u> as compared to other UK universities. For many of our undergraduate students, summative assessment happens almost entirely at the end of the third year. Some graduate students are formally assessed less than once per academic year. Perhaps due to this, and the rapidly changing nature of GenAI, the University of Oxford has been somewhat slow to publish rigorous formal policies on appropriate use of GenAI in academic work.

The University of Oxford's main policy on AI is contained in <u>its plagiarism policy</u>, where plagiarism is defined as "[p]resenting work or ideas from another source as your own ... by incorporating it into your work without full acknowledgement", including "the use of material generated wholly or in part through the use of artificial intelligence". Despite the prohibitive wording of the policy, the upshot is that if acknowledged fully, AI generated content *can* be used as part of academic work. The policy goes on to clarify that in *assessments*, AI can only be used with specific prior authorisation, or when agreed as a reasonable adjustment regarding, for example, disability.

In July 2023 the Russell Group universities issued a statement, backed by our Vice Chancellor, that aims to ensure students and staff are "Al literate", including the principle that staff should be able to support students to use GenAl appropriately.⁶ In January 2024, the Centre for Teaching and Learning published a guidance note for students on using generative Al tools to support learning which advises that students can "make use of generative Al tools ... in developing your academic skills to support your studies" and emphasises the importance of attribution or acknowledgement when used.⁷ The University has also recently provided guidance on ways tutors might incorporate GenAl into their practice, and online workshops.⁸

Much of the University's guidance on GenAI leaves practical questions open, perhaps inevitably because of the fast-moving field and wide range of tools. GenAI is, however, a live issue with great relevance to Keble's academic work. Our students produce work that is marked on a weekly basis, and undergo termly formative assessments (collections), while our tutors deliver a high number of contact hours and mark dozens of pieces of work each week.

It is in this context that Keble students and staff are already using GenAl with little central guidance, or avoiding doing so because of this lack of clarity on its proper use. Tutors are regularly questioned by students about appropriate use of GenAl, and many are also wondering how they might use it in their own work.

Section 2 of this document provides ten tips on these and related questions, focused on practical steps staff and students can take now and thoughts for the future. We can use (and review as needed) this advice at Keble over the coming terms and years.

⁴ Graham Gibbs, Harriet Dunbat-Goddet, "The effects of programme assessment environments on student learning" *The Higher Education Academy*, February 2007. https://www.advance-he.ac.uk/knowledge-hub/effects-programme-assessment-environments-student-learning

⁵ "Plagiarism", University of Oxford, accessed January 2024.

https://www.ox.ac.uk/students/academic/guidance/skills/plagiarism

⁶ "New principles on use of AI in education", Russell Group, 4 July 2023. https://russellgroup.ac.uk/news/new-principles-on-use-of-ai-in-education/

⁷ "Use of generative AI tools to support learning", University of Oxford, January 2024.

https://www.ox.ac.uk/students/academic/guidance/skills/ai-study

⁸ "An introduction to the use of generative AI tools in teaching", Centre for Teaching and Learning, University of Oxford, accessed January 2024. https://www.ctl.ox.ac.uk/ai-tools-in-teaching

Section 2: Advice for using GenAl at Keble

Learn from each other

GenAl tools are very new, and will continue to evolve. We have lots to learn from each other about best and worst practice.

Talk to each other – students, tutors, and supervisors alike – about how you are using GenAl in your work, or why you are avoiding it.

Be critical

Many users see GenAl as a philosopher's stone, speedily transforming inputs into high quality outputs. While speedily is certainly correct, high quality is debateable. And although GenAl has its uses, it's perfectly possible that feeding your ideas through GenAl makes your work worse.

- 2. Read text that is produced by GenAI closely before using it in any context. The quality of the work you produce is, ultimately, your responsibility alone.
- 3. Check facts carefully, exercising even more caution than you would with sites such as Wikipedia. You have every reason to be cautious: GenAl is known to present falsehoods as facts.

Use the best tool for the task

The quality of a GenAl response depends critically on the type of task you ask it to do. Current GenAl can excel at editing, organising, and performing repetitive tasks, but can fall down on computing mathematical or logical problems, creating original ideas, writing concisely, and providing up to date results – proceed with caution (see appendices)!

4. Try GenAl tools out, and use your own judgement to decide how high quality, appropriate, and correct the responses are. For example, with repetitive or procedural tasks it can be a real timesaver. Don't forget to check the final outputs by hand.

Acknowledge what you use

If you choose to use GenAl in your academic work then don't be shy about this — acknowledge it. Just as there's no one way you could use GenAl as part of your workflow, there's no one correct way to acknowledge its use, and you will have to develop and use your own academic judgement.

- 5. You can acknowledge GenAI in writing (footnotes, acknowledgements, methods, discussion), and/or by discussing it with tutors, students, supervisors.
- 6. Usually, a simple citation will not be appropriate; in the same way that you wouldn't cite Google, a citation to "ChatGPT, 2024" isn't going to cut it.
- 7. Unacknowledged direct use of text produced by GenAI is plagiarism, as laid out in the University's policy.

Practice writing on your own

GenAI can be used at many different stages in academic work: reading, generating ideas, drafting, writing, editing, proofreading. It is best, however, to avoid GenAI is when writing final drafts.

- 8. As with traditional notes, you should avoid plagiarism by not copying and pasting GenAI text directly into your writing, especially for final drafts. It is much harder to plagiarise if you write in your own words.
- 9. Developing your own writing voice is important. Writing coherently, fluently, and quickly is important for exams, for your academic career, and for the world of work. This skill takes practice don't rely on GenAI and miss the chance to practice it.

Remember confidentiality

Anything you put into a GenAl prompt can be stored and become part of its training data. This important facet of GenAl raises wider societal impacts, but for our purposes a practical concern is confidentiality.

10. *Never* enter any confidential or private information into a GenAl tool. This is particularly important for things like lists of names, contact details, or academic results.

Appendix 1: Example, Philosophy (tutorial/reading)

This example is based on a tutorial question and suggested reading from the syllabus for a prelims course in Moral Philosophy sat by Keble students in Michaelmas Term 2023. It demonstrates how GenAl could be used to summarize set texts as part of tutorial work.

Prompt

Can you summarize J.S. Mill's book Utilitarianism, specifically Chapters 1 and 4, in relation to the below question?

Q: What is the most significant objection to the proof of the principle of utility? Is it sound?

Response (generated using ChatGPT 3.5)

Certainly, in J.S. Mill's "Utilitarianism," Chapters 1 and 4 play a crucial role in establishing and defending the principle of utility. Here's a summary with a focus on the question of the most significant objection and the soundness of the proof:

Chapter 1: General Remarks

- Mill introduces the principle of utility as the foundation of his moral philosophy, asserting that actions are right in proportion as they promote happiness and wrong as they produce the reverse of happiness.
- He argues for the greatest happiness principle, emphasizing that the moral worth of an action is determined by its contribution to overall happiness.
- Mill introduces the qualitative distinction between higher and lower pleasures, suggesting that higher pleasures should be given greater weight in moral calculations.

Chapter 4: Of what sort of Proof the Principle of Utility is Susceptible

- Mill addresses objections to the proof of the principle of utility, particularly the criticism that happiness cannot be the standard of morality.
- He argues that the proof of the principle is empirical and based on induction, drawing from human experience and introspection.
- Mill responds to the objection that the pursuit of pleasure is base or ignoble by asserting that happiness is the ultimate goal that people naturally seek.

Most Significant Objection: A significant objection to the proof of the principle of utility may be the contention that reducing morality to the pursuit of pleasure oversimplifies ethical considerations and neglects other important aspects of morality, such as justice or individual rights.

Soundness: The soundness of Mill's proof depends on one's acceptance of the utilitarian framework. Critics argue that the principle may be too reductionist and fail to account for the complexities of moral reasoning, while supporters maintain that it provides a clear and rational basis for ethical judgments.

In summary, a potential objection revolves around the perceived oversimplification of morality in the utilitarian framework. The soundness of Mill's proof is a matter of debate and depends on one's acceptance of the underlying utilitarian principles. Chapters 1 and 4 provide the groundwork for understanding Mill's utilitarianism and his responses to objections, contributing to the ongoing discussion about the validity of the principle of utility.

Comment

A Philosophy tutor writes that many parts of this summary are vague, inaccurate, imprecise, or wrong – the opposite of what is required for philosophical reading and writing. They write that it "reads like someone who has never written a philosophy essay and as if they don't know anything about utilitarianism. It is everywhere vague and imprecise ... It's also just factually wrong about a decent amount of the claims, but that might be a consequence of the fact that it wasn't given the [Roger] Crisp guidebook [to Mill on Utilitarianism], which explains the theory more clearly, in more detail, and in a more professionally appropriate style than does Mill's original presentation and yet is written at a more accessible level than much of the professional secondary literature. If this stuff were in the first essay of their first term, I would think that it is about the level I might expect. If it were in the second, I would be disappointed. Were it submitted as part of a preliminary exam, it would receive a low 2.2—50-54, or thereabouts"

Appendix 2: Example, History (exam/essay planning)

This example is based on a question taken from the Preliminary Examination in History, *History of the British Isles 4:* from 1500 to 1700, Trinity Term 2023. It demonstrates how GenAl could be used to respond to exam questions, perhaps as part of revision.

Prompt

Outline an essay answer to the following question, which could be written by a first-year undergraduate student of history in one hour:

To what extent did Tudor rule before 1547 reconfigure the workings of politics?

Response (generated using ChatGPT 3.5)

- I. Introduction
 - A. Definition of Tudor rule (1485-1547)
- B. Overview of political landscape before Tudor rule
- C. Thesis statement: Tudor rule significantly reconfigured the workings of politics before 1547.
- II. Consolidation of Power
- A. Henry VII's establishment of a strong monarchy
 - 1. End of Wars of the Roses
 - 2. Centralization of power and royal

finances

- *III. Administrative Reforms*
 - A. Creation of the Court of Star Chamber
 - 1. Enhanced royal justice
 - 2. Curbing noble power
 - B. Development of the Privy Council
 - 1. Advisory body to the monarch
 - 2. Increased administrative efficiency
- IV. Foreign Policy and Diplomacy
 - A. Henry VII's diplomatic strategies
- 1. Marriage alliances (e.g., Treaty of Perpetual Peace with Spain)
 - 2. Expansion of international influence

- V. Religious Changes
 - A. Impact of the Reformation
 - 1. Break from the Catholic Church
- 2. Establishment of the Church of England
 - B. Dissolution of the Monasteries
 - 1. Redistribution of wealth
 - 2. Weakening of ecclesiastical power
- VI. Economic Reforms
 - A. Enclosure Acts
- 1. Changes in land ownership and agriculture
 - 2. Economic impact on society
- VII. Social Changes
 - A. Shifts in social hierarchy
 - 1. Rise of the gentry
 - 2. Impact on traditional power

structures

VIII. Conclusion

- A. Recap of key points
- B. Emphasis on the transformative nature
- of Tudor rule on political structures
- C. Closing thoughts on the lasting impact of Tudor rule before 1547.

Comment

A History tutor writes that the GenAI "does appear to have generated a brief outline of a plan which could be used to answer this question, organised efficiently according to different themes and aspects of 'politics' e.g. consolidation of power, administrative reforms, foreign policy, religious policy, socio-economic policy. It thus would provide the student with a sensible sense of the types of issues they might want to cover for this question, yet provides too many points which could not all be covered in a one-hour exam or single 2,000-word essay. In this sense it is expecting first year students to cover a little too much ground. However, using this as a basis for work might well give students an idea of the way they might structure such an essay. For its comments on the introduction, AI has noted the need for students to define some of the key terms in the essay question, yet there might be some issues here in the words it has picked for definition e.g. it asks them to define Tudor rule, though I would suggest students also offer a definition of 'reconfigure' and of what we might mean by 'politics' in this essay - things that AI has not picked up on here. The latter for example would allow them to consider such things as popular politics. It also suggests that an overview be given of the political landscape prior to Tudor rule, which would perhaps be best integrated in the paragraphs dealing with various themes to assess changes and continuities. Al has also given the student a thesis statement, something which is perhaps more worrying as I would hope students advanced their own arguments in relation to the question, and arguments with a little bit more nuance e.g. instead of saying that Tudor rule did or did not reconfigure politics, to offer an assessment of whether we see variations in this according to the various themes examined, or any issues of change over time in the period 1485-1547."

Appendix 3: Example, Chemistry (exam)

This example is based on a question taken from the Preliminary Examination in Chemistry, *Physical Chemistry*, Trinity Term 2023. It demonstrates how GenAl could be used to respond to exam questions, perhaps as part of revision.

Prompt

Answer the following questions, where the number of marks is given after each part in square brackets.

- (a) What is the meaning of the term state function? Give an example of a thermodynamic quantity that is a state function and one that is not. [3]
- (b) For a process $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$ write down the overall change in a state function, f, using as few terms as possible. Write down the overall change in a quantity, g, that is not a state function. [3]
- (c) Evaluate the overall entropy change in the following process, treating the gas as ideal:

A sample containing two moles of helium gas at 298.15 K and 1 bar is cooled to its normal boiling temperature of 4 K, condensed to a liquid and then cooled further to 2 K, where it undergoes another phase transition to a second liquid form called He-II. This liquid is suddenly vaporized by a beam of laser light and the resulting gas is brought to a temperature of 298.15 K and a pressure of 0.5 bar. [3]

Response (generated using ChatGPT 3.5)

(a) **State Function**:

- A state function is a thermodynamic quantity whose value is solely determined by the current state of the system and is independent of the path the system took to reach that state.
- **Example:** Internal energy (U) is a state function as it depends only on the current state of the system, defined by temperature, pressure, and volume. On the other hand, heat (q) is not a state function as it depends on the specific process.

(b) Overall Change in State Function (f):

- For a state function, the overall change is path-independent, and it can be expressed as $\Delta f = f(final) f(initial)$.
 - **Example:** $\Delta U = U(final) U(initial)$.

Overall Change in Quantity (g) that is not a State Function:

- For a quantity that is not a state function, the overall change depends on the specific path taken, and it cannot be expressed solely in terms of initial and final values.
- **Example:** Work (w) is not a state function. The overall change in work (Δw) cannot be determined by w(final) w(initial) due to path dependency. (c) **Evaluation of Overall Entropy Change:**
 - 1. Cooling from 298.15 K to 4 K (Condensation):
 - Use the formula $\Delta S = nCln(T2/T1)$, where C is the molar heat capacity.
 - 2. Cooling from 4 K to 2 K (No Phase Transition):
 - Use the formula $\Delta S = nCln(T2/T1)$.
 - 3. Phase Transition to He-II at 2 K:
 - Utilize the entropy change associated with the phase transition.
 - 4. Vaporization at 2 K by Laser Light:
 - Apply the formula $\Delta S = q/T$.
 - 5. **Heating from 2 K to 298.15 K:**
 - Use the formula $\Delta S = nCln(T2/T1)$.
 - 6. Overall Entropy Change:
- Sum up the entropy changes from each step, ensuring to consider the appropriate signs for processes like condensation and vaporization.

Comment

A Chemistry tutor writes that "part (a) is completely correct and would be awarded full marks. The first part of the answer to (b) is correct, although the question is written such that $f_4 - f_1$ is the expected answer rather than f(final) – f(initial). The second part of the answer to (b) is true, but misses the point of the question, which is that the change in g could be expressed as the sum of the changes at each step in the process. The method presented in (c) is incorrect and misses the point of the question, which is that entropy is a state function. Thus, we can ignore most of the steps given in the question and just calculate the difference in entropy of the gas under an isothermal reversible expansion between the initial state (when T = 298.15 K and p = 1 bar) and the final state (when T = 298.15 K and p = 0.5 bar)."

Appendix 4: Example, French (translation)

This example is based on a question taken from the Preliminary Examination in Modern Languages, *French IIA: Unseen Translation into English*, Trinity Term 2023. It demonstrates how GenAI could be used in studying languages.

Prompt

Translate the following passage into English:

Deux femmes feuillettent des catalogues de vente par correspondance, l'une en face de l'autre, dans le train Cergy–Paris. La plus jeune commence solennellement : « Ma mère ne s'est pas remise d'une histoire, qui est arrivée dans son immeuble. » L'autre la regarde avec intérêt. La narratrice continue donc. Elle construit le récit devant nous (nombreux voyageurs debout, plusieurs se mettent à écouter), avec un personnage, une vieille femme qui a des ulcères aux jambes, un lieu, l'immeuble de la mère de la narratrice, des péripéties : disparition de la vieille, absence de bruit derrière sa porte, puis des gémissements, intervention de la mère auprès du gérant pour faire ouvrir la porte, refus de ce dernier, appel ultérieur à la police. Les actants du récit se distribuent en « bons » (la mère) et en « mauvais » (le gérant). L'issue fatale est prévisible dans le ton et la conduite du récit, la jeune femme multiplie les parenthèses lourdes de significations, « on ne pouvait enfoncer la porte, massive, c'est un immeuble ancien », les indications temporelles, « avant-hier », « hier », menant à un présent d'horreur. Elle s'interrompt, « bon », relance en feignant la surprise « et voilà que », petits mouvements de langue, geste de la main. Jouissance de la narration visible sur son visage, aux yeux baissés, levés épisodiquement sur la première destinataire de l'histoire, la jeune femme assise en face d'elle (mais destinataire maintenant fictive, la véritable étant la foule de gens agglutinés dans l'allée centrale du wagon).

Response (generated using ChatGPT 3.5)

Two women flip through mail-order catalogs, facing each other on the Cergy–Paris train. The younger one starts solemnly, "My mother hasn't recovered from a story that happened in her building." The other looks at her with interest. The narrator continues, constructing the story before us (many standing passengers, several start to listen), with a character, an old woman with leg ulcers, a place, the narrator's mother's building, and twists and turns: the disappearance of the old woman, silence behind her door, then moans, the mother's intervention with the manager to have the door opened, the manager's refusal, a subsequent call to the police. The characters in the story are divided into "good" (the mother) and "bad" (the manager). The fatal outcome is predictable in the tone and conduct of the narrative; the young woman multiplies parentheses heavy with meanings, "the door couldn't be broken down, it's massive, it's an old building," temporal indications, "the day before yesterday," "yesterday," leading to a present of horror. She interrupts herself, "well," restarts, feigning surprise, "and here it is," small movements of the tongue, a hand gesture. The enjoyment of storytelling is visible on her face, with eyes occasionally lowered on the first recipient of the story, the young woman sitting across from her (but now a fictional recipient, the real one being the crowd of people crowded in the central aisle of the carriage).

Comment

A Modern Languages tutor writes: "I think it's worth a low to mid II.1 mark. The syntax and vocabulary have for the most part been very well understood. There's a strange error/omission/ 'aux yeux baissés, levés épisodiquement'. 'Levés' is omitted in the translation and 'épisodiquement' is translated as if it went directly with 'baissés', so falsifying the sense of the sentence. The sense of "et voilà que' has not been understood at all: 'and here it is' won't do. There's another omission that loses sense. The translation doesn't translate 'donc' in the sentence 'La narratrice continue donc'. But the 'donc' flows on from the previous sentence: the narrator continues her story BECAUSE the woman has looked at her with interest. The translation doesn't convey that causal relationship. There's clumsy repetition in the translation: 'la foule de gens agglutinés', 'the crowd of people crowded' isn't very good. Americanism: 'catalogs'. And about half a dozen unconvincing words or phrases: e.g. translating 'une histoire qui est arrivée dans son immeuble' as 'a story that happened in her building', when we would more likely say 'something that happened'; similarly, we'd more likely say 'heavy with / pregnant with meaning' rather than 'heavy with meanings'."

Appendix 5: Key terms

Artificial intelligence (AI) concerns intelligence of computer programs or machines, and as an academic discipline it is roughly 70 years old. Since the early 2010s advances in particular computer methods known as deep learning have accelerated the field. AI technology is widely used in the modern world, and is built into day-to-day tools including search engines like Google), and speech interpretation systems like Amazon's Alexa.

Generative artificial intelligence (GenAl) is a subset of the above technologies that came to particular public prominence in late 2022/early 2023. GenAl tools are capable of quickly producing novel text, images, or other outputs based on users' prompts (often a short sentence). The GenAl tool should produce outputs that have similar characteristics to the large body of data on which the GenAl model has been trained.

In short, training an AI model means writing a computer program that is fed large amounts of data (text, images, music) and that analyses these data to spot patterns. The program can then mimic these patterns in future outputs, with varying degrees of fidelity and originality. GenAI tools are normally accessed via a web browser or an app, with most requiring the user to create an account and some requiring paid subscriptions.

Some academic-related tasks that GenAI can be used for are in the below list, which is both non-exhaustive and subject to change over the coming months and years.

- Writing Generating ideas for essay answers, parsing questions, writing essay plans, writing sections of essays (or entire essays).
- Reading and note-taking Analysing and summarising set texts, translating text, organising or writing notes, generating references and bibliographies.
- Science and maths Writing, or de-bugging, computer code, performing mathematical calculations, generating images and figures.
- Editing and marking Formatting text, adapting text for style/tone, spellchecking, marking work.
- Administration Writing reports, performing peer review, writing emails, organising lists.

The range of tools available is also advancing quickly week-to-week, and an exhaustive list would be impossible and quickly outdated. At the time of writing, current examples of GenAl software include:

- Large language model (LLM) chatbots, which generate text outputs based on text inputs. The outputs can be refined in conversation with the chatbot, often in an instant messaging style. Current examples include ChatGPT and Bard. These tools are the main focus of this report.
- **Text-to-image** systems, which as the name implies generate images based on a text input. Current examples include DALL-E and Midjourney.
- Image editing tools, which can fill in gaps, or remove objects, from images. Many tools, including <u>Adobe Photoshop</u> and <u>Google Pixel cameras</u>, have recently built these tools into their software.