Artificial Intelligence (AI) and Unit 2: Field Study / Investigation

In this support guide for students, we set out considerations you may take when exploring and critiquing the value of generative artificial intelligence (GAI) tools within sectors conducting a wide range of primary and secondary research that can includes but is not limited to:
- Participatory research with people, groups and demographic segments
- Ethnographic research into individual cultures, communities and societies
- Large and multifaceted data set analysis
- Quantitative data collection, categorisations and organising including mathematical approaches
- Archaeology through the study of human activity by recovering evidence

Context
AI is revolutionising many industries, which includes sectors working with data from both primary (first hand) and secondary (existing) sources. The analysis of increasingly larger data sets sourced using equipment, the internet, or conducting interviews or observations with people, makes it increasingly harder for humans to identify the trends, patterns and insights that the data might provide, due to the scale of the task and time it would consume. "Big data", referring to data so vast and complex that it is impossible for humans to reliably attempt to interpret, requires increasingly sophisticated ways to draw analysis that might inform the science, business or research and development sectors. The use of GAI has a potential increasing role to play in this field, but how reliable, accurate, and trust worthy are GAI tools in drawing the appropriate conclusions, to inform the right decisions are made?

AI:EPQ advice: The **sole use** of AI tools **cannot** meet assessment descriptor requirements set out for the extended project qualification. AI tool use can **appear alongside** and influence **evidence** for which the candidate has independently created (i.e. without the use of AI tools) in order to meet the assessment descriptors. Candidates are required to reference each and every instance of AI tool use within their coursework submission, and failure to do so will be considered malpractice.

Introduction
If you are interested in completing an EPQ in Unit 2: Field Study / Investigation and are also interested in critically evaluating GAI within your assignment, here are some considerations you will need to take when carrying out your project.

1. Developing a working title for a Unit 2: Field Study / Investigation
The title for a Unit 2 Field Study / Investigation should be written in the format of a hypothesis that you can test.

E.g. *'Hypothesis: The NHS is not providing patients with timely and accurate care when diagnosing common cold and flu symptoms post Covid 19 pandemic'"*

Whilst you are not permitted to use GAI to answer the hypothesis that you yourself are aiming to answer, nor are you permitted to ask GAI to do the research and arrive at a conclusion for you, GAI is capable of providing an interesting test case for a hypothesis that could be particularly market disruptive and innovative, and could give you a wonderful opportunity to become an expert in a particular area of GAI potential in the sector you hope to progress towards in the future.

E.g. *'Hypothesis: A medical Chatbot is more capable of accurately diagnosing patients with cold and flu symptoms than an NHS professional post Covid 19 pandemic'*

Alternatively, you may wish to implement GAI to do exactly what they are capable of doing, and for which a human (you) is incapable of doing. Processing "big data" that is impossible to analyse yourself, and critically considering the insights alongside your own data analysis, could be a really interesting way to explore a hypothesis that GAI is more or less capable of interpreting complex data sets. Your role could be to analyse a smaller and more manageable data set, in order to compare to a GAI analysed data set comparable to "big data".

E.g. "Hypothesis: Analysing the large performance dataset of an entire football team using a GAI is more insightful for the coach than analysing the individual performance data of each individual player."

Importantly, whilst a GAI is doing the analysis work that you simply would not have sufficient time to analyse yourself for your EPQ, you can carry out your own analysis work typical of an EPQ assignment, and use the GAI tool to provide you with the comparative information you need to answer your hypothesis, where you have identified a potential opportunity for GAI in your field of interest.

For more guidance on titles, see our titles support guide here.

2. Documenting when and how you use GAI
If you are planning to integrate the critical analysis of any use or output from GAI as part of your project, you will need to be explicit during the project write up, which includes a requirement to:
- retain a copy of the question(s) and computer generated content
- record a noneditable format (screenshot) of the question and output
- write a brief explanation of how the GAI has been used

These requirements can be recorded in a text document such as word, and referenced to in your project activity log.

Here are some key questions/milestones to consider when documenting your use of GAI in your project activity log:

- Which GAI(s) are you going to critically review? (Ensure that you reference and evidence your use in a document)
- How did your use of the GAI(s) evolve during your project, and did this inform your own analysis in any way? (Make sure to document the iteration of your written inputs into the GAI each time)
- What benefit did you personally gain from critically reviewing the GAI outputs alongside your own analysis and investigation work, and what implications might this have for associated industries? (Remember to imagine yourself as a researcher, and consider what is being improved by the availability of a GAI)
- How will you ensure that your own biases do not influence how you instruct the GAI in relation to answering your hypothesis? (What approaches will you take to ensure that you are not simply instructing a GAI to show you want you want to see?)
- From your project experience of investigation and field study related AI generated outputs, how will the pace of technological advancement in these GAI impact the sector now and in the future? (e.g. will humans retain a role, and will bigger data capture require GAI to advance to meet future needs?)

GAI
The following non exhaustive list of GAI are capable of generating outcomes associated with a Unit 2 Field study / Investigation EPQ.

Data analysis
- ChatGPT
- Akkio
- Polymer
- MonkeyLearn
- Looker (Google Suite)
- Knime
- Rapidminer
- PyTorch
- DataRobot
- H20.ai

GAI issues to consider
The use of GAI is intrinsically linked to potential issues that are a challenge for this evolving technology. Here are some specific field study / investigation related issues that you may wish to consider within your EPQ.

Ethics
A significant issue associated with all GAI use for generating content, is around ethics. Ethical use of any tool particularly to replicate or replace a human, honestly or deceptively, has a wide range of associated and unanticipated issues, that may be of interest to you in your EPQ, specifically relating to how GAI interprets data unemotionally, without empathy, and with the capacity to present factual translation of data irrespective of its impact on the audience.

Instructor bias
How instructors task a GAI to carry out a task, the language they use in their instructions or "prompts", will directly impact the output from the GAI, which as part of the EPQ you need to remain mindful of. If your instructions include bias such as gender, racial, ethnic or sociopolitical bias, and you are unaware of this, the responses you receive will lack the rigour to provide neutral and appropriate academic critique. Whilst it is almost impossible to remove bias in the application of GAI, with a considered level of instruction applied by you, a human, it is possible to:
1. Apply user awareness that accounts for bias in the first number of outputs from the GAI
2. Train the GAI through a series of prompts that account for potential bias, such as requests to include diverse datasets or a wide range of sources of information.
3. Carry out sufficient human led research on the topics being prompted by a GAI, to mitigate bias within a GAI conversation.

Evolving data sources
Though some GAI have limited access to updating or changing data sets, some tools are actively expanding the data from which it is capable of sourcing answers to instructor prompts. Because of this, it is important to acknowledge that some GAI will generate different solutions at different times during the academic year that you are completing your EPQ during. As some GAI are continually learning, increased data sources or inputs on any topic will change the outputs, and it is therefore important to remain mindful that outputs you reference will reflect an insight at a point in time, with a specifically available data set. Where data sets do not evolve or

change (such as drawing from a closed data pool and therefore not connected to the internet for example), the GAI might change their outputs in response to increased instruction, activity by humans, or updates.

<u>Unintended findings</u>

By using a GAI to carry out specific analysis tasks using research or data, it is likely to complete that task successfully with the appropriate instruction. However, it is important to account for unintended outputs from your instruction of a GAI, and the potential impact this could have on humans. For example, if data analysis by a GAI identifies a particular demographic as being more susceptible to types of high mortality diseases such as cancer, this information may be upsetting to those that identify in that demographic. Whilst intentions for GAI use for your EPQ will be focused on arriving at an answer to your hypothesis, you may in the process of asking, uncover insights and information that is not appropriate for your audience who will read and review your project write up.

An interesting research piece in this field for you to consider

Speeding up to keep up: exploring the use of AI in the research process, is an interesting paper that explores the increasing use of AI within academic research, in order to carry out timely and effective information gathering, analysing large unstructured datasets, increasing the amount of data based research that any individual can accomplish, and arriving at insights, at the expense of losing the creativity of the researcher (a human). This is justified as a need driven to keep up with advances in technology and demand for high quality research, and questions whether AI has a positive role to play.

Potential writing approaches for GAI

In order to use a GAI to analyse any form of data, small or large, you will need to be clear in your instruction of it through the prompts you choose.

As an example, if you were to use a GAI to analyse a body of data you have captured from a series surveys, you will need to instruct the GAI using the following prompts:

1. **Define your objectives** - Clearly define the objectives and goals for your analysis. What are you trying to learn from this dataset? What insights or answers are you seeking?
2. **Data collection and pre-processing** - Collect and clean-up your data. Ensure that it is well-structured and free from any inconsistencies and errors. This step is crucial to avoid introducing bias from the data itself.
3. **Select relevant features** - Choose the features (known as variables) from the dataset that are relevant to your analysis. This step can help you to focus on the most critical insights.
4. **Analyse data distribution** - Conduct exploratory data analysis (EDA) to understand the distribution of data and identify any initial patterns or trends.
5. **Formulate your research questions** - Define the research questions you want to answer in relation to your hypothesis, from this dataset. These questions will guide the analysis and help you to stay on track.
6. **Model training and fine-tuning** - You may need to fine-tune your prompt to a specific dataset related to your problem. Fine-tuning allows the GAI to provide more accurate and relevant responses.
7. **Bias mitigation** through these three steps;
    a. Carefully prepare and pre-process your data to remove biases.
    b. Evaluate the GAI outputs for any biased responses and iteratively improve through repeated testing.
    c. Implement actions to identify and rectify any bias.

8. **Data sampling** - Depending on the dataset size, consider sampling a representative subset of your data for initial analysis as a test.
9. **Data analysis** - Use the GAI to assist you with data analysis by asking it questions, generating you a report, or providing explanations of any patterns and anomalies.
10. **Visualization** - Create visual of the outputs to help you better understand the data and communicate insights effectively in your assignment.
11. **Iterative analysis** - The analysis process should be iterative. Continuously refine your questions and analysis techniques as you uncover insights.
12. **Interpretation and validation** - Interpret the results and validate your findings using your own research and knowledge of the data, and other comparable sources.
13. **Documentation** - Document your entire analysis process, including the data used to support your hypothesis and as part of the assignment. Ideally, you will provide sufficient evidence for someone else (like your assessor) to repeat the process themselves.
14. **Reporting and communication** – Present a summary of your findings for your assignment, including the GAI findings and insights. Did you find any limitations in the GAI?
15. **Peer review** - If possible, ask an expert in your field or sector of interest to quality assure and validate the insights from the GAI.
16. **Ethical considerations** - Be mindful of ethical considerations, including privacy, confidentiality, and responsible data handling.

Referencing

As a requirement for your EPQ assignment, you must create a bibliography of the research and academic sources of information you read, that support any written or performed output. As covered earlier, due to the advancing nature of GAI, you are also required to create evidence of AI use with the following details:
- retain a copy of the question(s) and computer generated content
- record a noneditable format (screenshot) of the question and output
- write a brief explanation of how the GAI has been used

Here is a suitable reference for the paper *"Speeding up to keep up: exploring the use of AI in the research process"*:

Title: Chubb, J., Cowling, P. & Reed, D. Speeding up to keep up: exploring the use of AI in the research process. AI & Soc 37, 1439–1457 (2022). https://doi.org/10.1007/s00146-021-01259-0URL: https://link.springer.com/article/10.1007/s00146-021-01259-0

Date of access: 18th October 2023

JCQ Guidance on Coursework Assessment
Please review the rules and guidance relating to the use of AI within coursework and assessments here:
AI Use in Assessments: Protecting the Integrity of Qualifications
Information for candidates Coursework assessments