**INTRODUCTION AND HYPOTHESIS**

The question “What is Artificial Intelligence?” is a hard one to answer. Truthfully, scientists and researchers still disagree to this day about what it is, what it can do and what it should do. Most researchers in the field commonly agree that “people in other fields cannot agree on either the relevance or the place of AI,” **1**and that “Artificial Intelligence is a field that is not entirely sure what it is about.”**1**To the average person the concept of this may seem ridiculous - people dedicating their lives to something that has no solid definition or application – but the truth is that AI, while being a complete mystery, is the future. Even today Artificial Intelligence is something that is not easily classified due to it being incorporated into so many aspects of society –finance, transport, medicine, law, education and gaming to name a few. Some say that “AI is the study of knowledge”**1** and others say that it is “the characterisation of the nature, form and scope of human knowledge.”**1** However, nobody seems to be able to agree on a standard definition.

Artificial Intelligence is a rapidly changing subject. In the 1970s, the cutting edge AI technology was industrial automation, and they dreamed of machines that could learn from experience and think like a human. Nowadays, Artificial Intelligence is used everywhere – we have self-driving cars, machines with a human-like ability to think, and are developing AI weapons, lawyers and even nurses. Even in the past the fast paced evolution of AI was recognised – an AI researcher, Groover (1986) stated that “The state of the art in this field is changing rapidly,”**2** and even today the extent to which the state of the art in AI will evolve is unknown.

Because of this, AI tends to attract a lot of negative opinions. AI has been named “the bad child of the sciences, the half breed of academic research,”**3** as well as “dehumanizing and dangerous”**3** simply because people do not understand it, and so they fear it. In the late 1900s it even went so far as some US states having “virtually outlawed AI research”**3** due to the fact that they felt threatened by it, and while this is no longer the case, the strong negative attitudes still remain.

I plan to deeply research the reasons behind the negative attitudes towards Artificial Intelligence, both from the perspective of AI researchers and the general public. I want to understand the common reasoning behind people’s rejection of the technology – is it influenced by the world around us? Or is it simply the fear of the unknown and the fear of change that trigger an automatic dislike? Personally, I have a very positive opinion of AI as I know all of the good it can do as well as the role it will play in the future. I’m intrigued to see whether negative attitudes correlate with knowledge on the subject, because while I am by no means an expert on AI and how it works, I have a vision of the future the technology could create. Researchers are already predicting unbelievable applications, for example using nanotechnology to cure almost any human ailment, implants to improve hearing or vision, or even curing the effects of ageing. These things may seem very far off in the future, or something out of a sci-fi series, but with Artificial Intelligence technology they may be a reality sooner than you would think.

I think I will find that the fear of the unknown is the most prominent of the reasons behind negative opinions. Humans have a habit of rejecting something they don’t understand, and AI is definitely one of those things. I also think that media portrayal – in movies like Terminator, for example – may have an effect, especially on the younger generation who have grown up surrounded by these extreme representations of AI and may have shaped their opinions around them. I expect there to be many other reasons that I am not currently aware of before starting the project, and these are the ones I am most interested to learn about and understand. Gaining an understanding of the other side of AI, the side people dislike, will give me a wider understanding of the topic as a whole and help me when I come to study it further at University.

**PROJECT AIM**

The aim of this project is to understand the reasons behind negative attitudes towards AI, and conclude as to which ones are the most common and important. I have a very positive attitude towards the topic, and want to develop my understanding of both sides of the argument so that I can see the whole picture. Furthering my understanding in this will help me when I go on to study AI at University, as I will have a broader understanding of the subject, and may also help me further on if I end up working in the field as I will understand the ethics behind it.

I am going to gather a large amount of secondary research from different media, including books, scientific journals and web articles, so that I can gather a wide range of data from a range of perspectives. I will also construct a questionnaire, which I will give to people from various different backgrounds, for example engineers, nurses and students. I will analyse the results I get from my questionnaire, and write an analysis of any patterns and conclusions I can see from the data. After I have researched and analysed all of my data, and compiled it all into a dissertation, I will form a conclusion where I will tie up everything I have found out and hopefully answer my question. I will evaluate each reason for negative attitudes and decide which is the main reason, as well as which reasons are insignificant and uncommon.

Once I have finished my project I will evaluate it, and outline what I could have improved. This will help me in University when I am doing projects for my degree as I will have learned from the mistakes I will make during my EPQ.

**PROJECT OBJECTIVES**

* To further my own knowledge in Artificial Intelligence as it’s a subject I am very interested in and want to study further.
* To understand the reasons people have for rejecting or fearing Artificial Intelligence, and why they may have opposing views to me.
* To gain research and project skills that will be valuable to me in University, as well as time management skills and prioritisation skills that I will no doubt learn during this project.
* To gather a wide variety of sources, and in the process learn about referencing and read books on Artificial Intelligence that otherwise wouldn’t be available to me.
* To construct and use a questionnaire to support my secondary research, and in the process find out more about the opinions of those around me.
* To write my dissertation to the highest possible standard, and include lots of varied research and information while answering my chosen question.
* To reflect on my project and find out what I can improve when I am asked to do similar projects at University.
* To complete my project comfortably within my time frame and have enough time to review and improve anything that I need to.

**ETHICAL GUIDELINES**

My EPQ topic does not really have the potential to offend anyone or cause any ethical issues during the making of this project. While collecting data for my primary research questionnaire, I will ensure that everyone’s results remain anonymous and that there is no bias in the nature of the results. I will do this by asking equal numbers of students, engineers and nurses, so that the results are not skewed a certain way, whether intentionally or unintentionally.

Mentions of religion in my primary research may cause offense, as potentially dehumanizing technology and religion will never agree and this may cause the discomfort of some people. To ensure this doesn’t happen I will ensure that nobody is pressurised into filling out my questionnaire, and make sure to be brief whenever mentioning beliefs. Another topic that may cause offense is employment, and the prospect of workers being replaced by machines. This is particularly relevant for the engineers that I will be asking to complete the questionnaire – engineers and factory workers are among those most at risk. I will make sure the questionnaire is based on opinions rather than providing potentially concerning information, in order to avoid any problems.

Artificial Intelligence as a whole has its own ethics and moral guidelines so as not to overstep any boundaries. However, whether these guidelines are being followed by everyone is something I will revisit later in the project.

**METHODOLOGY**

To ensure that I have a wide variety of data from a wide range of sources, I will include both primary and secondary research in this project.

**Primary Research**

* To gain unpublished data I will form a questionnaire and get people from various backgrounds to share their opinions and the reasons behind them.
* I will ensure I have mixed data by compiling data from people from different occupational backgrounds, ages and beliefs.
* Once I have my data, I will analyse any patterns I find and compile a small conclusion that links to my question and sums up anything interesting that I find.

**Secondary Research**

* I will use a range of books and scientific journals from Newcastle University Library, as well as online articles from various opinionated sources to get a full scope of perspectives.
* From this research I want to find the scientifically accepted reasons to reject or fear AI as well as any negative opinions of actual researchers.
* Before writing my dissertation I will need to split my secondary research into subtitles, in order to split up my writing and avoid repeating myself.
* I will also need to keep a detailed bibliography with comments on both relevance and reliability of the sources I use.

**EVOLUTION OF THE QUESTION**

When I first started this project, I knew I wanted to focus on Artificial Intelligence, but didn’t know exactly what I wanted to get out of the experience. Since then, my project title has developed immensely.

**First Question Idea: To what extent could Artificial Intelligence replace the human race?**

Initially, I was curious about the answer to this question, so it served as a starting point to my project. However, I very quickly realised that it was a very broad question that I could not possibly answer within a realistic time frame or word count, so I had to change it.

**Second Question Idea: To what extent could Artificial Intelligence replace humanity in the workplace?**

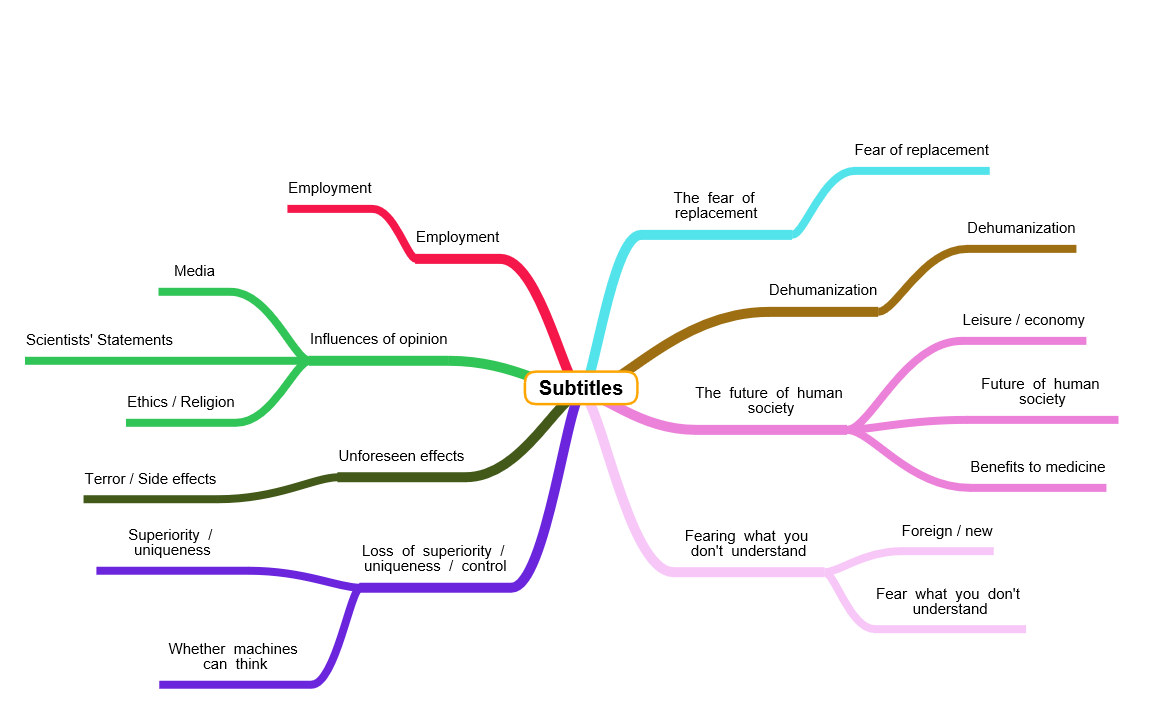
This question was a lot more realistic than the previous version as it was specific to employment and the workplace, and could easily be done in a decent amount of time with a relevant conclusion. At this point I was starting to look into research, and I quickly realised that I had a lot more information available to me than I had previously thought. I decided to change the focus of my project slightly, focusing instead on attitudes towards AI, which had been one of my initial ideas in the first planning session I had.

**Third Question Idea: Analyse and evaluate common reasons behind positive and negative attitudes to Artificial Intelligence.**

After going to the Newcastle University Library and gathering most of my secondary research, I realised that I would need to cut down my question to be able to write a relevant conclusion. If I had continued on with both positive and negative attitudes, I would not have been able to write a conclusion and would instead have just rambled about the information I had gathered. Since I personally have a positive attitude and am very curious about the negative side, I decided to focus my topic on the negative attitudes, which brought me to my final question.

**Final Question Idea: Analyse the reasons behind common negative attitudes towards Artificial Intelligence.**

**INITIAL DISSERTATION PLANNING**

****I initially split all of my secondary research into 8 subtitles, with a brief summary of what I would include within each section.

**Influences of opinion -** The origin of all opinions / ideas that people have about AI – media, religion etc.

**The fear of what you don’t understand -** People uneducated, unaware of uses, current applications, researchers don’t even have the answers so the public are wary of it.

**The future of human society -** Fears of what it will do to our society, danger and paranoia vs improvements to medical and other fields, also fear of change – change of human attitudes to creativity, revaluate what we value etc.

**Dehumanization -** Some fear robotising of humanity, loss of what it means to be human, vs this can be good – medical enhancements.

**The fear of replacement -** Fear that AI evolution will surpass humanity and we will be replaced / wiped out by robots, vs those who think that we should, or those who think that just because it could happen doesn’t mean it will.

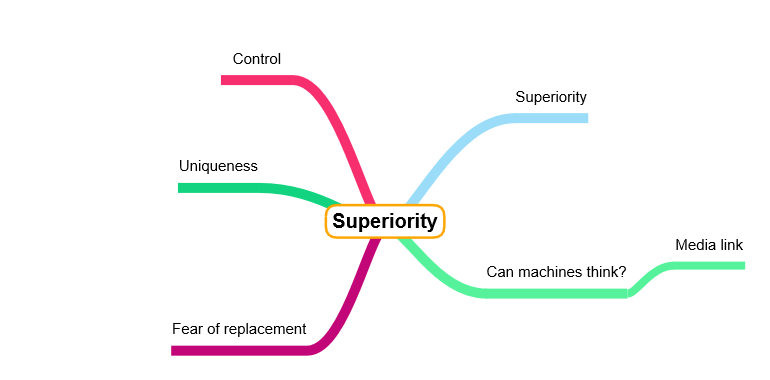
**General fears and paranoia / Unforeseen effects / Dangers -** Predictions of what AI will mean for us, generally pessimistic and involving a revolution and the death of humanity – fuelled by media, technology in wrong hands, warfare etc.

**Man’s loss of superiority and control -** As AI comes closer to actually being able to ‘think’, humans lose their uniqueness and superiority, perceived as a threat, fear of loss of control – debate on whether machines can think.

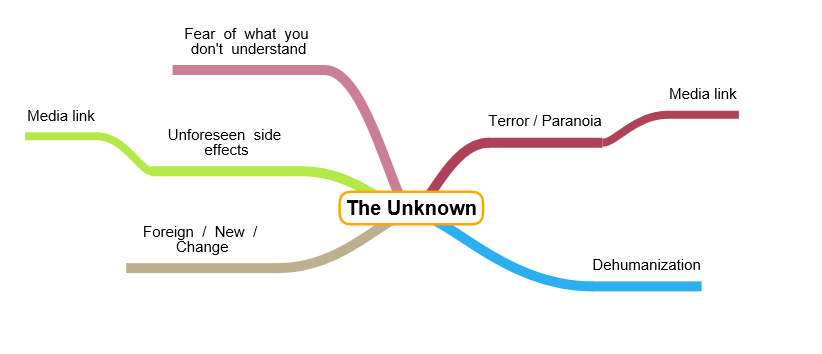
**Employment -** AI is reducing the amount of jobs and causing more inequality vs AI is actually creating more, higher paying jobs – opinions depend on perspectives and wealth.

**REFINED DISSERTATION PLAN**

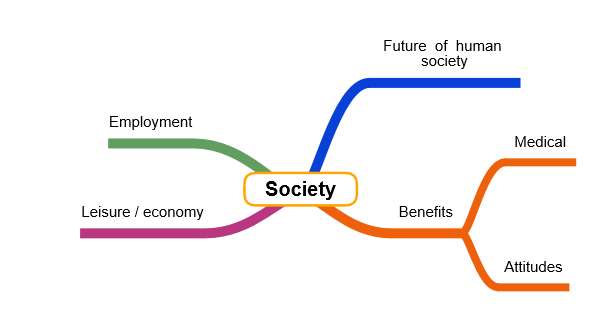
Instead of having 8 subtitles, I decided to limit myself to 3 as most of the topics linked to each other and had the same idea behind them, and could therefore be grouped together. This will make it easier to draw a relevant conclusion, as well as prevent me from repeating myself.

**Subtitle 1 – The Superiority Complex**

**Subtitle 2 – Fear of the Unknown**

****

**Subtitle 3 – Societal Impacts**

****

At this stage I think I am ready to start writing up my secondary research in the form of a dissertation laid out in the above format. I will use the Harvard referencing system to reference all of my secondary data, and once it is completed, will complete my primary research analysis.

**PRIMARY RESEARCH PLANNING**

**Type of questions I want to ask:**

* Perspective – ask for job title eg. student, nurse etc, ask for age range
* Effect on workplace personally / fear of overtaking
* Would you consider yourself to be knowledgeable on the subject?
* Should there be limits, can/should machines think
* Aware of applications, of what technology could do for us in the future
* Does media /religion/morals have an effect on opinions
* Biggest fears – eg. dehumanization, overtaking, unemployment
* Benefits >< Drawbacks
* Whether robotization should happen
* Personally feel threatened by the technology
* Do you believe that there will be unforeseen negative side effects

**People I want to ask:**

* Students
* Family – different jobs, different perspectives, jobs and ages – mainly nurses/carers
* Engineers – from work experience I am already doing

**What I want to get out of doing this:**

I want to understand the opinions of those around me and use the data I receive to support my dissertation. I will analyse the data I receive with graphs and pie charts, and comment on any patterns I find, and any conclusions I can gather. I have tried to include questions from all subtitles that I plan to use in my dissertation, so my secondary research can always be backed up or contradicted by the data I will receive from doing this questionnaire. The focus of this questionnaire is exploring attitudes, and I have tried to tailor my questions to reflect this.

I will explain what the data will be used for as well as a small introduction at the start of my questionnaire, to ensure that the people taking the questionnaire know what it is about and know their anonymity is safe. I will keep my questions simple and easy to understand, and as short as possible to reduce the time the questionnaire takes to complete.