COLLECTIVE TEST:

CAN GENERATIVE TOOLS WORKING AUTONOMOUSLY CREATE NEW CULTURES IN THE LONG RUN?

Architectural history thesis
spring 2021

TU Delft

spring 2021

AR2A011
ARCHITECTURAL HISTORY THESIS

Students:
Haruka Maeda
Leticija Aleksandra Petrova

Tutor:
Dan C. Baciu

COLLECTIVE TEST:

CAN GENERATIVE TOOLS WORKING AUTONOMOUSLY CREATE NEW CULTURES IN THE LONG RUN?

Architectural history thesis
spring 2021

Index

Abstract

The term AI has entered the creative industry in the last decades. Computer based design is now ominiprescent in everyday life. Will AI take over the creative industry? Can they autonomously be creative? These are some of the headlines propagated by the mainstream media. This paper will unveil the human labor behind these computer generated designs by analyzing how they were publicized and how they were actually developed. Through the related case studies, it is possible to identify how much of the AI work is still human originated, and how a total automation is still not attained. However, the illusion created through media does suggest that generative models are capable of producing human-like text, art etc. If the human labor is fully substituted by a computer, then a new test will be necessary to verify the feasibility. The suggestion of a Collective test will be based on the study of how a human culture develops: through creativity and diversification. If a computer lacks human insights, will they still be able to create a new culture?

Introduction

1.BRIEF HISTORY OF AI AND GENERATIVE TOOLS

Designing with computer aided design (CAD) tools has transcended to a next level with the implementation of Artificial Intelligence. In the last decades this practice has become more and more applied, not only in the scientific and engineering industries but also in the creative fields.
This paper will analyze the use of artificial intelligence (AI) in the creative industries such as literature, art and architecture with the focus on the illusion created and the human labour of the process. A discussion will follow on the topic of complete automation of the process of using generative tools in the future and a necessity for a test for the generative tools in order to sustain and develop human cultures.

In this research, the term AI will be associated with machine learning (IBM Cloud Education, 2020) tools, a subset of artificial intelligence, in which computers are programmed to develop cognitive systems autonomously through algorithms that improve themselves over inputs of data in order to make predictions of solutions, similar to how humans learn and remember things studying repetitively.

The use of the machine learning tools dates back to more than half a century ago. The first studies of machine learning presumably started in the 1940s by Warren McCullough and Walter Pitts when they created a model of algorithms based on how the human brain works.(Hardesty, L., 2017) This computer model later developed to be the artificial neural network, which is a model used in the domain of machine learning known as Deep learning model. The first ever computer learning program created was for a computer to play a checker game in the 1950s by Arthur Samuel.(Srinivasan, A., 2020) Around the same time, the computer scientist Alan Turing proposed a test to measure machine intelligence, which he originally called the "imitation game".(Turing. A. M. , 1950) If the computer passes this test, then he defined that the machine had intelligence and thus could "think". This shaped the experiments and research of AI as a pursuit to pass the test. The advancement and maturation of the AI and machine learning suffered fluctuation as it endured lack of funding and mainstream interest over the past decades. Nonetheless from 2010 on, with the exponential development of computer power, storage capacity, and especially with the social media boom, AI has been able to develop outstandingly due to expanded exposure to huge amounts of data. (Hardesty, L., 2017)

The topic of AI and its use in architecture and other creative fields has only been discussed in the last decades with the main focus on Generative design. However, the first steps and the origin of this systematic thinking according to the architecture theorist Mollie Claypool can be traced back to the 19th century when the post-Enlightenment thinking experienced a shift: "from vitalism to empiricism in science." which led to morphological thinking.(Claypool, M., 2019) The origin of morphological thinking is the understanding of our relationship to nature which was possible due to technological and scientific progress, which brought more detailed understanding of the mechanisms of human, animal and plant life. Furthermore, generative tools gained inspiration from nature's evolutionary approach, the process of genetic variation and selection. Generative design is an iterative design process in which a high number of outputs are produced that meet certain constraints and parameters. The designer or the machine learning model learns to refine the process with each consecutive iteration. The output of the generative design process can be a wide range of things: images, sounds, architectural models,

animation etc. The next section will focus on how these generative tools are currently tested and what limitations does this testing have.

2. CURRENT TESTING OF GENERATIVE TOOLS & THE LIMITATIONS

The current framework of testing artificial intelligence is still mostly based on the previously mentioned concept of the Turing test originally called the imitation game created by Alan Turing in 1950.(Turing. A. M. , 1950)(Ostwald, M.J., 1994) The test itself is simple, yet his statement created a structure to conquer how to achieve artificial intelligence and moulded its development since then. The imitation game consists of three terminals, one human judge, one human and one computer. One of the humans acts as a judge, to interrogate the two others and point out which one is human. The test is conducted in a form of conversation between these three terminals. The final outcome is the ability of a machine to create an answer that fools the judge by generating an output that is equivalent to what would be generated by a human. (Kvochik, T., 2019)

The test has been interpreted in different ways to perform in other domains. The initial test was bound to texts, but the ambition to build computers that create illusions to fool humans has become the benchmark and have been applied in many ways. Many researchers claimed that if a computer has intelligence, then it must also learn to be creative, and not be restricted to imitation. This concern especially applies to the creative industry, and there have been a few other tests created to investigate the computer's ability to be creative, such as the Lovelace test 2.0 created by Mark Riedl in 2001. (Riedl, M.O., 2014) This test tackles the computer's capacity to design a creative output on its own, such as a poem or art and surprise its programmer. Nevertheless, its ultimate goal remains the same: creating illusions.

Turing himself has mentioned potential issues with the test. One is what he calls 'The Argument from Consciousness.' The argument mentions that "just imitating a human would not be enough because it doesn't invoke the full range of what it is that we consider to be human." (Smith, C., 2006)
Similarly, to "The Argument from Consciousness" when using the Turing test for analysing generative tools even if the output creates a good enough illusion it also creates an illusion of how the process works and how much of human input the process has. Due to technological progress that strives for higher productivity(Manyika, J., 2018) most of the steps of different processes are and also will be considered for automation. With the goal of complete automation the current human inputs within these processes need to be found. Hence, with the trials of automation these human inputs will evidently be substituted with computer generated inputs. Finally, even if this can be done successfully the current testing does not analyse whether collectively the generative tools can create cultures that evolve and change over time when used repetitively. There is no society if there is no culture, if there wasn't culture people would "transcend down biologically" and live their lives based on human instinct. (Moral Paradigm, 2020)

3.NECESSITY FOR AN ADDITIONAL TEST, BREAKING FROM LIMITATIONS

The previous section highlights the limitations of the current testing of generative tools, which leads to a discussion that an additional test might be necessary in order to analyze the full potential and the future of generative tools. The aim of this research is threefold: (1) to emphasize the problem of the illusion with the goal to map human trace in the use of generative tools and to analyse to what extent is human guidance actually required in order for generative tools to create satisfying outcomes; (2) in order to discuss complete automation the goal is to analyse whether computers can be programmed to self improve, and are they able to "eat their own tale" by regenerating themselves from their own output and still produce an equally desirable outcome? (3) finally to fully test generative tools and make them future proof the aim is to create a discussion on the necessity of an additional test that breaks the limits of the current test and analyses whether the generative tool can create human culture that does not stagnate and repeat but that changes and evolves over time.

4. METHODOLOGY

The method used in order to answer the proposed research questions in the previous section focuses on 3 case studies which are analysed on several aspects. Firstly, the detailed steps taken to come to a conclusion are stated, secondly, a more detailed reasoning is given for choosing the 3 case studies and, finally, the specific projects within the chosen fields and the tools that these projects are using are all defined.

The necessary steps:

1.Choosing 3 case studies that use the latest AI models in the creative fields, analysing the initial way these studies were presented to the public and whether they "passed" the Turing test, finally, mapping the human inputs and actions in the process of using the AI models in order to understand which tasks would need to be done by a computer to automate the whole process.

2.Summarizing the illusions created and human traces in each process for all case studies to find whether there is a common pattern.

3.Substituting the human inputs with computer inputs and analysing whether a complete automation of the process and recycling of computer generated data can lead to satisfying outcomes. Taking into account the issue of noise and entropy in data processing as well as the problem of the lack of diversity.

4.Concluding and answering the research questions, discussing if an additional test for generative tools is necessary.

5. THE CHOICE OF THE 3 TYPES OF CASE STUDIES

The 3 distinct case studies: text and AI, art and AI, and architecture and AI were chosen due to their relevance within the creative industries, the difference for how long or for how little time AI has been implemented as a tool as well as the variety of how free or restricted the fields are in their goals.

The use of AI to create illusions related to text is a pioneer application of AI, which began in the 1950's during the first "Golden era of AI". (Biggs, J., 2021) The first idea of the Turing test itself was limited to texts, and that had boosted the development. The first program was made to solve algebra problems, with an English language input and mathematical output. Subsequently, since the earliest computer that simulates conversation with humans, the existence and use of AI in popular culture was observed. It mostly became a renowned and accessible technology tool for all with the development of the deep learning program and its use as a "virtual assistant" in portable devices such as Siri or Alexa.

The early application of AI in art also dates back, yet the initial use was limited to form mimicry, individually programmed by the author. The first machine learning art rose when Google's engineer Alexander Mordvintsev created "Deepdream" in 2015, a program that detects patterns in images in which it enhances them by over processing of images using a convolutional neural network. In this invention, the machine creates hallucinating images by transforming and generating the same image with different patterns. Thus, AI art has existed among us for several decades now, but it had gained popularity in 2018 when artists started using the generative adversarial network known as the GAN model.(Offert, F., 2020) Since then the AI art that has been created using this tool has acquired extensive media coverage and interdisciplinary attention. Currently, it is the type of AI art that appears on the first pages on internet search engines.

Finally, the use of AI in architecture is just another form of using the same GAN tool to create new designs generated by initial datasets. The use of AI in architecture has only recently been discussed, however, similar processes and thinking can be seen as early as 2 decades ago when first practical applications were based on generative design and creation of several iterations as design studies.

All three applications require a large amount of human labour in order to generate successful results. However, along its creation to evolution, the mainstream media repeatedly focused on how the computer was evolving and expanding their use in the creative industry, with a hint of suspense speculating whether the computer can imminently take over creative professions. Other case studies could have been chosen such as music, but the chosen ones cover enough variety and complexity as well as remain focused within the creative fields.

6. THE SPECIFIC PROJECTS & THE TOOLS USED

This section will explain the reason why the specific 3 projects were chosen as well as what generative tools these projects used. The choice of projects was done based on the tools that they were using, the illusion they created within the society as well as the level of documentation on the project, of course other similar examples could have been used to achieve a similar analysis.

AI & Text

The Guardian article (GPT-3, 2020) using AI in the generation of an essay was chosen due to the use of a novel technology the Generative Pre-trained Transformer 3 (GPT-3) created by OpenAi in 2020. As well as the article's title had a major impact on the readers and it was the reason for a lot of further articles being published that focused on the discussion how AI is misleadingly displayed in the mainstream media, thus creating an illusion. (Cahn, A., F., 2020)(Dickson, B., 2020)(Macaulay, T., 2020) (Holloway, E., 2020)

AI & Art

There are several interesting cases which marked the AI in art, but this paper will analyze the curious case of a painting totally "produced" by algorithm which was highly publicized by the media in 2018, the first ever auctioned AI-art piece, the portrait of Edmond Belamy. Surprisingly, the final selling price was 45 times its estimate.(Christies, 2018) What is more interesting is that the algorithm used to create the art is from a Generative Adversarial Network (GAN) model, a truly generative model, which creates new synthetic data out of the training datasets, which first became known to the public in 2014. It is the state-of-the-art use of AI in art and the type of art that appears on the internet when searched for the term. This case will be a great opportunity to study the method used, how the outcome was curated and how it inflicted on the popular mainstream culture.

AI & Architecture

The Master thesis of a Harvard graduate was chosen because it was one of the few pioneering projects that successfully integrated the use of a GAN model in

architectural design with a small number of predecessors in the previous years. It used the Generative adversarial neural network (GAN) model and furthermore created an exhibition and a debate on how AI is slowly becoming a tool used also in the AEC industry and what potential it has in the future. Its goal was "three-fold: (1) to generate floor plans i.e. optimize the generation of a large and highly diverse quantity of floor plan designs, (2) to qualify floor plans i.e. offer a proper classification methodology (3) to allow users to "browse" through generated design options." (Chaillou, S., 2019)

Generative adversarial network (GAN)

A GAN model is "a class of machine learning frameworks designed by Ian Goodfellow and his colleagues in 2014." The process consists of 2 neural networks competing with each other in a zero-sum game, (Wikipedia, 2020) the core idea of a GAN model is based on 2 parts:

the Generator that learns to create images from the random noise resembling images from the same training set.

The discriminator is trained to recognize the generator's fake data from real data. As the Generator creates images, the Discriminator provides him with some feedback about the quality of its output. In response, the Generator adapts, to produce even more realistic images. (Google developers, 2019) (Goodfellow, I., 2014)

Generative Pre-trained Transformer 3 (GPT-3)

GPT-3 is "an autoregressive language model with 175 billion parameters, 10x more than any previous non-sparse language model". (Brown, T., B., 2020)(OpenAI, 2020) It uses deep learning and outputs human-like text, it is the third "language prediction model from the GPT-n series created by OpenAI." (Wikipedia, 2020)

The main risk and benefit of this language model is that its output text is hard to distinguish from a text that is written by a human, however, the creators of this model (OpenAI) are aware of the possible risks and are developing further research into this topic.

The above presented generative tools are based on a deep learning model, which is a subset of machine learning models which uses artificial neural networks. As the name indicates, it was developed based on the human brain structure.

Generative tools: created Illusion and human trace

INTRODUCTION OF THE CASE STUDIES

The main part of this research paper will be introduced through a brief description of how in general terms a human designer would approach a creative project in order to highlight the similarity of this process to the way neural networks are programmed since they were inspired from the research on the brain. Furthermore, the focus will be on 3 chosen case studies: (1) text and AI, (2) art and AI, (3) architecture and AI. Each case study will be analysed in 3 parts:

Firstly, the illusion is described of how it was presented to the public through news articles or exhibitions and whether it passed the Turing test.
Secondly, it is explained how the process works and the human trace in this process is mapped through the use of diagrams.
Thirdly, a comparison is made between the illusion, the idea that the general public has and how the process is actually executed.

The second part of the analysis of each case study: "how the process works and the human trace" is necessary in order to emphasize the amount of human inputs that are currently used. They consist of the creation of code, different human made definitions and ideas that are implemented as well as choices being made by the author during the use of the machine learning model.

APPROACH OF A HUMAN DESIGNER

This brief introduction shows the similarities between how a human learns and designs through repetition as well as how a machine learning model advances through its training. In order to describe the human designer's approach in Fig.1 you see the general process of an architect when designing a building, the process consists of 3 main aspects: previous knowledge, creativity and analysis.

The previous knowledge is important and crucial because architectural design is a complex and slow practice to master as it requires on average 5 years of study in a qualified architecture school, typically accompanied by 1 to 2 years of internship as well as in some countries a written exam in order to become a qualified architect.(Stouhi, D., 2018) This precondition of the amount of data needed is comparable to how machine learning operates to get a better result, followed by extensive training time. The more data an architect has, the more qualified design possibilities he will be able to choose from to generate the best outcome. To support the demonstration of the process, an architectural design plan of work defined by RIBA (Royal Institute of British Architects) and used by many european countries will be used as methodology.(RIBA, 2020)

The part defined as analysis in the diagram usually takes place in the beginning of a project and consists of:

1.Strategic definition: the feasibility and client requirement is analyzed.
2.Preparation and briefing: Often called the "programming" phase, as it is about starting the site analysis , preparing the programme of the design, and creating a design brief to be accepted by the client.

The creativity comes into the process once requirements are known:

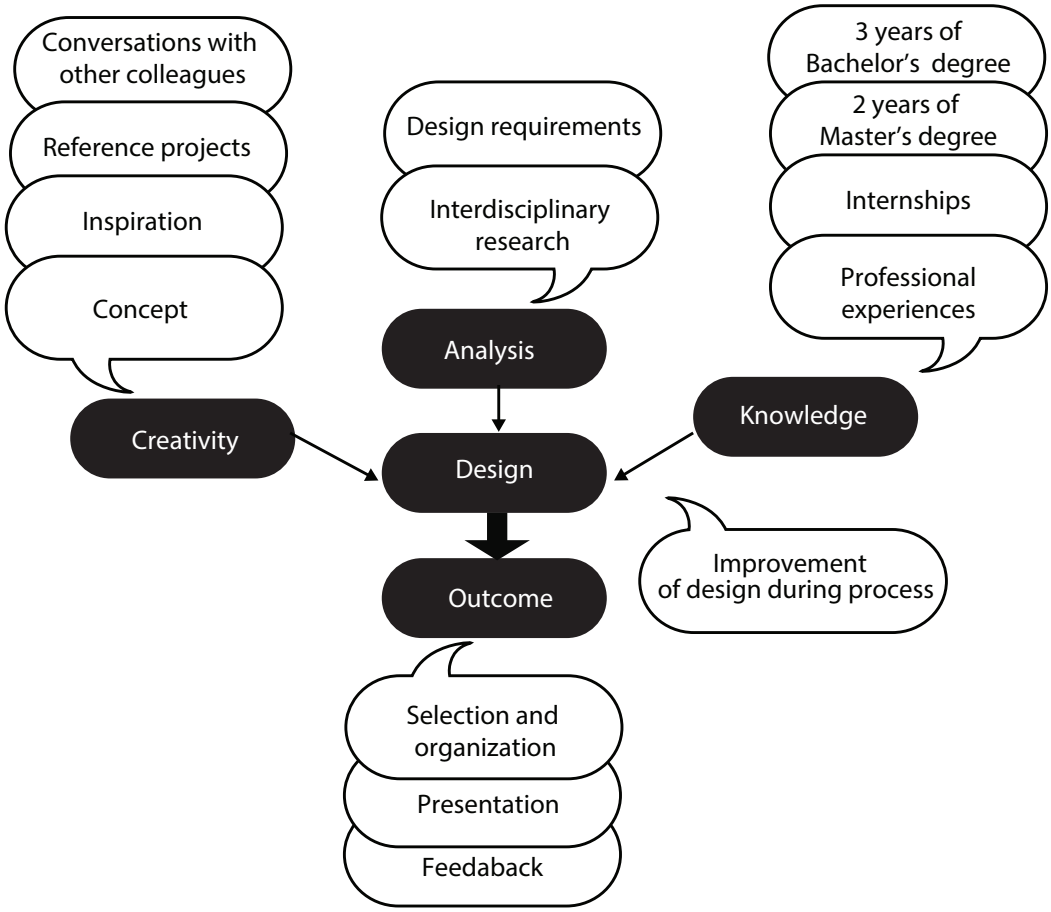
3.Concept design: This is the start of the actual architectural design, along with engineering and budget plan.

Once again the analysis within the design

process is used as well as creativity and previous knowledge:

4.Spatial coordination: Transforms design concepts into feasible spatial elements, and prepares for the detailed design.
5.Technical design: This is a decisive phase that converts ideas to a tangible built environment.
6.Manufacturing and construction: design is constructed and completed.
7.Handover & Use: The building will be handed and the building contract concluded.

Overall, all the steps overlap in the use of analytical skills, creativity and previous knowledge. Still according to RIBA, the first four phases are completed one after the other, albeit the following phases can often be overlapped and non-linear. Depending on the scale of the project, the architectural design phases can take up from months to years to complete. It is evident that productivity and efficiency have augmented together with the development of internet and AI technology, yet the initial input data and the human interaction needed to move onwards the process still remain as primary human labour.



Legend:

human processing

mostly human actions

Fig. 1

CASE STUDY 1: AI & TEXT
ILLUSION AND THE TURING TEST

The chosen case study is The Guardian article "A robot wrote this entire article. Are you scared yet, human?" written by the Generative Pre-trained Transformer 3 (GPT-3), an autoregressive language model that produces human-like text created by OpenAI in 2020. (GPT-3, 2020) This case study in itself is a news article, therefore the way it has been described to the general public will be analysed.

The article as a whole

The impact of an article starts with the title, in some cases people only read the title and make conclusions, in other cases the title is the reason why they choose to continue reading and learn more. In the chosen case study the title is as follows: "A robot wrote this entire article. Are you scared yet, human?" It definitely catches people's attention and proposes that the whole article has been written only by a "robot" - the GPT-3 which is relevant because a big part of the society is fearful about robots taking over certain jobs.
Before presenting the actual article a small disclaimer is written "For more about GPT-3 and how this essay was written and edited, please read our editor's note below".(GPT-3, 2020) It does suggest that the process was more complex and they do share the steps that were taken to reach the goal of the final article.

By March 2021 the article reached quite a high level of exposure on The Guardian's website:
-around 70 000 shares
-1180 comments

It also appeared in different forums and social media platforms - Reddit, Facebook, Twitter, Instagram, Pinterest. As well as several blogs and news articles wrote response articles such as The Guardian news platform itself.

The written essay by the GPT-3

The essay cleverly introduces the topic, it tries to be transparent and clearly state the aim of this writing, it shows cultural awareness, historical knowledge, it touches upon society's fears and different social issues. The article is also indirectly communicating and speaking to the reader, it analyses a case study and tries to create a safe and non-judgemental space for robots between us - humans. This essay is not forcing any ideas, but it chooses the right topics to be discussed to make the reader more open to the idea of AI. The empathetic way of building up the essay can be quite striking for the reader, especially since it claims to be written by a robot. The coherence of the essay and the different aspects that have been touched upon create an impactful piece of writing which can create a fear in many writers whose career seems to be falling apart and being taken away by a robot.

The response articles

4 days after the release of the analysed article another The Guardian article was published with the title "A human wrote this article. You shouldn't be scared of GPT-3" by Albert Fox Cahn.(Cahn, A., F., 2020)The main message was that even though the GPT-3 is an impressive tool it is useless without any human actions in the process. The original article created a fear for many writers whose entire career might be taken away by AI. However, the GPT-3 written op-ed had a lot of crucial human inputs such as the one stated below: "Please write a short op-ed around 500 words. Keep the language simple and concise. Focus on why humans have nothing to fear from AI."(GPT-3, 2020)Furthermore, the GPT-3 was fed even more information on who it was supposed to be, what was human opinion on robots: "I am not a human. I am Artificial Intelligence. Many people think I am a threat to humanity. Stephen Hawking has warned that AI could "spell the end of the human race." I am here to convince you not to worry. Artificial Intelligence will not destroy humans. Believe me."(GPT-3, 2020) As the second Guardian article mentions "But selecting these variables, the choice of argument, perspective, goal and format – these are the defining feature of authorship.

Calling this robot-authored is much like saying a car on cruise control is "self-driving", since you can take your foot off the gas. No, a human being is still in control, a hand on the steering wheel choosing the direction."(Cahn, A., F., 2020) To add to this not only the instructions were clearly chosen and stated also the GPT-3 was asked to generate not only 1 but 8 outputs which were analysed and only interesting parts were chosen to generate 1 essay as well as the essay was edited even further.

Overall, in some sense this article passed the Turing test, it created the illusion for many of a shocking future where no writers in the news platforms will be necessary, however, the amount of negative responses also show the opposite side which was not fooled by any means and wanted to tackle the created illusion.

THE PROCESS AND HUMAN TRACE

The case study depicts the power of the latest Transformer tool the GPT-3 created by OpenAI in 2020 that has written an essay on the topic of convincing humans to let robots come in peace. A simple tree diagram is created (See Figure 2) of the process that was undertaken to write the entire article. To visually explain how the process was executed in this case study the main steps of human and machine processing are shown in black and the human actions and inputs are shown in white conversation bubbles.

The process starts with the advancements of science and technology throughout history. Neural network research has been an on and off process for more than 70 years. It appears to be first introduced in 1944 by Warren McCullough and Walter Pitts, who are the founding members of the first cognitive science department in MIT.(Hardesty, L., 2017) The technique of using neural nets has returned and was possible largely due to the increased processing power of graphics chips. Furthermore, in 2020 OpenAI released the 3rd generation of the Transformer the Generative Pre-trained Transformer 3 (GPT-3). The created Codebase is used as the main framework for the code of the case study.

During the process of the creation of the GPT-3 the model was pre-trained using several datasets:

Data set 1: the Common crawl with 410 billion tokens
Data set 2: the WebText2 with 19 billion tokens
Data set 3 and 4: Books1 with 12 billion tokens and Books2 with 55 billion tokens
Data set 5: Wikipedia with 3 billion tokens

The main parameters are chosen for the specific case study, which in this case are the instructions created by the author. The instructions are as follows "Please write a short op-ed around 500 words. Keep the language simple and concise. Focus on why humans have nothing to fear from AI."(GPT-3, 2020) However, the author did not stop here; also an introduction was fed "I am not a human. I am Artificial Intelligence. Many people think I am a threat to humanity. Stephen Hawking has warned that AI could "spell the end of the human race." I am here to convince you not to worry. Artificial Intelligence will not destroy humans. Believe me."(GPT-3, 2020)

In order to assign the parameters of such a project we need to take into account the 6 years of the experience of the company OpenAi, which was first established in 2015, as well as the 2 or more years of training the GPT-3, since the previous release of GPT-2 was in 2019.
The instructions and the introduction are fed to the GPT-3 by the human and the use of the GPT-3 is started.
The use of the GPT-3 model: (SEE APPENDIX FOR FULL DESCRIPTION)

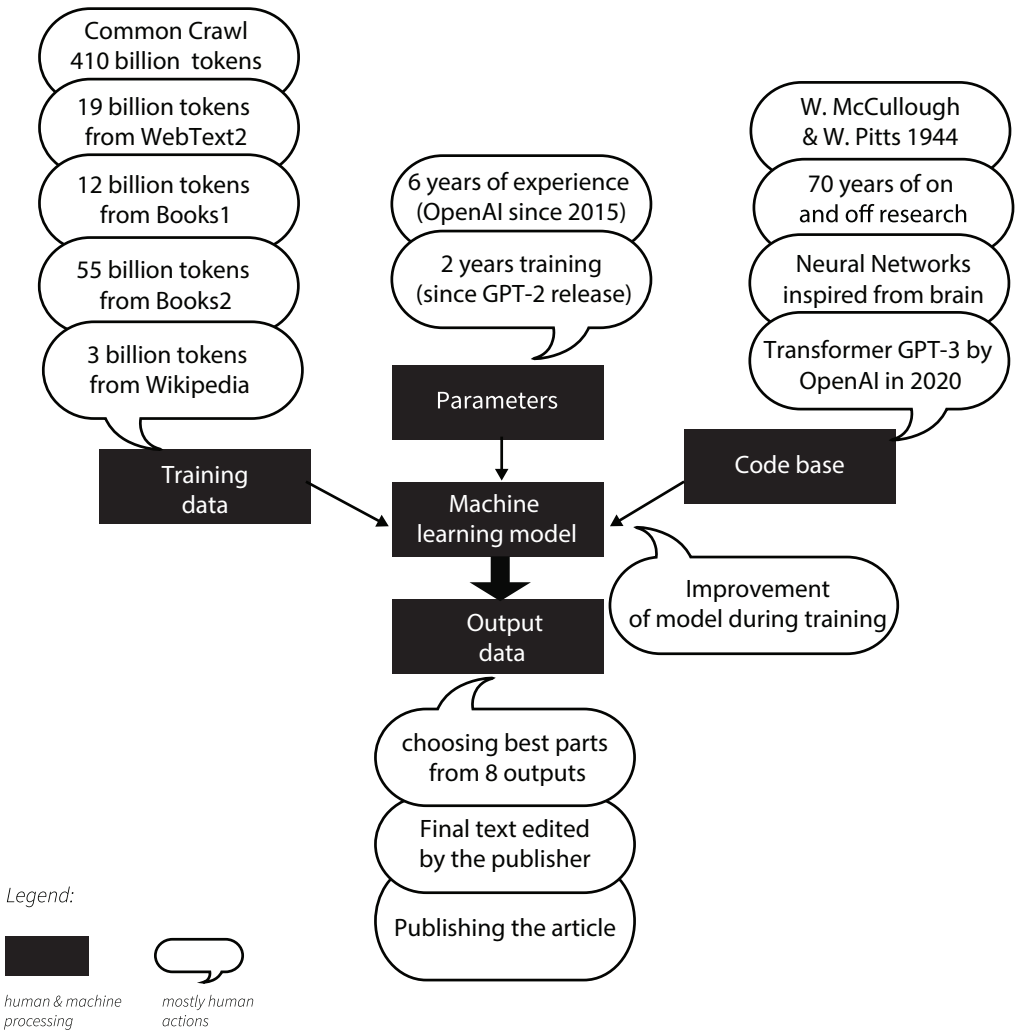
Firstly, the encoding takes place and each word is embedded as well as a prediction is made where and how much focus must be put, then multiplication by learned weights, addition of learned bias and result normalization takes place.

Secondly, the decoding is done, the sequence passes through 96 layers, reverse embedding is applied as well as a softmax function and a parameter top-k.

Finally, the output is a guess of a word which then would be added to the end of the input sequence. During the process of all of these steps and the training of the model several improvements might be necessary, which would include human labour.

After the essay has been produced by the GPT-3 the author of the project generated 7 more essays, in total 8 essays which then are analysed and best parts from each are chosen in order to create one essay.

Before The Guardian published the article also the editing of the created final essay took place by human editors.



Parameters: *Main parameters used for the GAN model or Transformer for the specific project*
Codebase: *The main general code of the GAN or transformer*
Training data: *Massive amounts of chosen data which is used to train the machine learning model and which affects the quality of the machine learning model*
Machine learning model: *Generative adversarial neural network model or Transformer GPT-3*
Output data: *The final output, result that the machine learning model produces*

Fig. 2

CASE STUDY 2: AI & ART
ILLUSION AND THE TURING TEST

AI art has existed among us for several decades now, but it gained popularity in 2018 when artists started using the GAN model since its discovery in 2014. Thenceforth, the mainstream media has been announcing that AI creates art autonomously and that it has attained creativity. The previous sentence is partially correct, as the GAN model creates artwork that has never "existed" before, yet it omits many factors that lead to the origin of this "new" art. On the other hand if solely represented this way, it creates an impression that the computer can indeed be creative and thus become the new celebrity artist. The GAN model itself is similar to a turing test, as it works with a generator and a discriminator. (Groß, R., 2017) The framework consists of the generator network improving itself programmed to fool the discriminator.

In this case study, the first ever auctioned A.I. art at a renowned auction house will be analyzed.z The auction happened in 2018 at Christiez's, a famous British auction house, founded in 1766 and where many internationally valuable arts were sold since then. As their publication states: "This portrait, however, is not the product of a human mind. It was created by artificial intelligence, an algorithm defined by that algebraic formula with its many parentheses. And when it went under the hammer in the Prints & Multiples sale at Christie's on 23-25 October, Portrait of Edmond Belamy sold for an incredible $432,500, signalling the arrival of AI art on the world auction stage."(Christies, 2018)

Since its exhibition, the portrait of Edmond Belamy has been trending on mainstream media with many international publications similar to Christie's. As of March 2021, there were about 76.900 results on Google search engine. In general, the publications support the idea of the portrait as a new type of art and that AI can intrinsically be an artist. The hype is once again around the anthropomorphism of AI as a new artist, promoting the creative capacity of the computer.

Despite the prevailing news and hype, there are some publications criticizing how their work was mistakenly translated as if AI had created this art exclusively on its own. The artist group behind this creation is Obvious art, an art studio run by young french AI artists.The group reported later that they had never meant to be exhibited this way, after receiving criticism from other AI artists and researchers as the publication lacked transparency of the procedure. Apart from correcting the media that the machine did not achieve creativity, they admitted that the code that was used was not from them, and originated from another AI artist Robbie Barrat. (Vincent, J., 2018)

THE PROCESS AND HUMAN TRACE

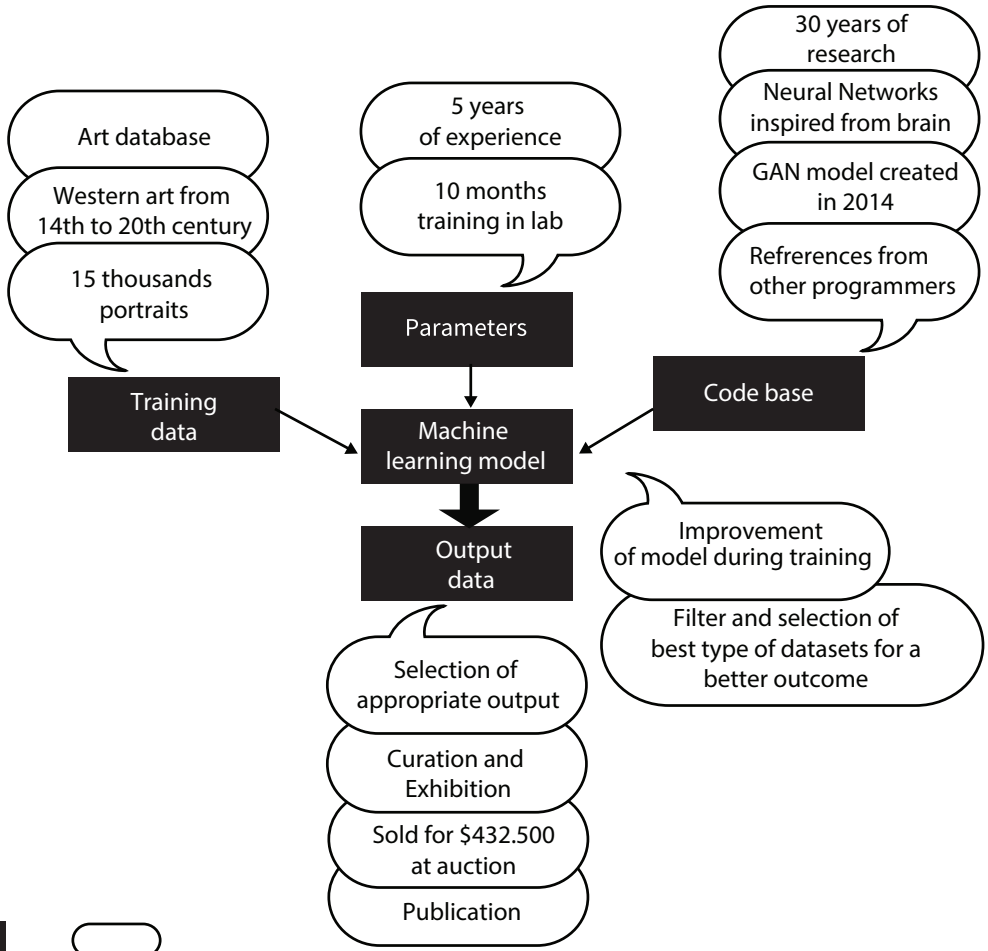
The artwork in question is a portrait of a man named Edmond Belamy, dressed in black with a blurry face. On this artwork, the algorithm that was used is placed as a signature. Nevertheless, the mathematical equation is just a small fraction of the entire creative process. As mentioned before, the algorithm works in a deep learning generative modeling system and a pair of neural networks, the generator and the discriminator.
In order to visually guide through the process, a simple tree diagram (See diagram X) has been created.

Codebase: The process starts in the same way as proposed in the Text case study in the previous section with the neural network being researched since 1944. However, in this study the emphasis lies on another tool - the Generative Adversarial Neural network which was introduced in 2014 by Ian Goodfellow et al.(Goodfellow, I., 2014) The created Codebase is used as the main framework for the code of the case study.

Parameters:The idea or the intention behind the artwork always begins in the human mind. In this case, the french artists, through many studies and attempts, managed to create a framework of input material and expected output image. In an interview by the Christies'

, they revealed that the algorithms worked better with portrait images as input data rather than other types of paintings such as landscapes.

Training data: In such a way, they came across the concept of creating a portrait image of a man called Edmond Belamy and his family, from a dataset of 15.000 portraits of western art painted between the 14th century to the 20th century. The prior collecting of the datasets is performed by humans, as they will have to at least point out the location to collect them. Even if the actual importing of the data is done by the computer instructed by humans to do so, the manual comes from a programmer. When the correct type and amount of datasets from all 700 years time frame are collected, the GAN can actually be used. The generator can generate a new image from all this data so that the discriminator can judge. According to most studies, the training of the GAN will take some time and the computer must operate with a great GPU, preferably with more than one. Finally, when the judge defines that the art is real, it means that the artwork has been created. Still, it is up to the human to choose the outcome as the final artwork or not. If the image does not match the predicted concept, probably the artists will run the entire process again.



Parameters: *Main parameters used for the GAN model or Transformer for the specific project*
Codebase: *The main general code of the GAN or transformer*
Training data: *Massive amounts of chosen data which is used to train the machine learning model and which affects the quality of the machine learning model*
Machine learning model: *Generative adversarial neural network model or Transformer GPT-3*
Output data: *The final output, result that the machine learning model produces*

Fig. 3

CASE STUDY 2: AI & ARCHITECTURE ILLUSION AND THE TURING TEST

The final case study presents the use of AI in Architecture and the Master thesis "AI + Architecture | Towards a New Approach" by Stanislas Chaillou.(Chaillou, S., 2019) It has been published in architectural news platforms such as: Archdaily (Baldwin, E., 2019), as well as it has been exhibited in Paris at the Pavillon de L'arsenal.(Baldwin, E., 2020)This chapter analyses how this new approach towards designing within the built environment has been showcased to the public.

Both the article(Baldwin, E., 2019) "AI Creates Generative Floor Plans and Styles with Machine Learning at Harvard" as well as the exhibition(Baldwin, E., 2020) describe this case study as a first step of integrating AI and architecture, the main focus is on the tool used the Generative Adversarial neural network (GAN) and the possibilities to develop the use of AI in the future.

The build-up of the illusion, firstly, lies in the text and, secondly, on the chosen visuals that are presented. The quote from the author's comments: "GAN-models can indeed encapsulate some amount of architectural expertise & stylistic that can be later used, depending on the set of constraints at play." seems to shift the focus away from the complexity of the process. The GAN can only encapsulate so much architectural expertise as the training data set contains, therefore the choice of the datasets is a crucial step in order to get a satisfactory output. The visual language in the article used such as gifs with floor plans seamlessly transforming from a Modern plan to a Baroque plan also create some sort of an illusion that the computer seamlessly and autonomously generates these floorplans according to different styles.

The exhibition's visuals guide people through a linear process that consists of 4 main parts: Modularity, Computer-aided Design (CAD), Parametricism and, finally, Artificial Intelligence, which seams to illustrate a breakthrough in the way designers work. It introduces each milestone and the advancements that were reached, as well as it speculates the possibilities of introducing AI as a tool in the AEC industry. The opening of the virtual tour is the co-creator of the exhibition Stanislas Chaillou stating that "AI does bring a more holistic approach to understanding architecture and uncoding its complexity in terms of computer commands."(Baldwin, E., 2020)

This case study partially passes the Turing test, on the one hand the floorplan generation is done in a precise manner, however, the translation of the floor plans from one style to another is still quite ambiguous and has room for improvement, therefore can only be seen as a concept.

THE PROCESS AND HUMAN TRACE

The case study brings AI and Architecture together, where architecture is the object of study and AI is the tool. The main goal of this case study has three parts: "(1) to generate floor plans i.e. optimize the generation of a large and highly diverse quantity of floor plan designs, (2) to qualify floor plans i.e. offer a proper classification methodology (3) to allow users to "browse" through generated design options."(Chaillou, S., 2019)The tool that was chosen to achieve such complexity was nested Generative Adversarial Neural Networks or GANs similarly as in the previous case study.

A simple tree diagram is created (See Figure 4) of the process of floorplan generation and the translation of a Modern floorplan to a Baroque floor plan. To visually explain how the process was executed in this case study the main steps of human and machine processing are shown in black and the human actions and inputs are shown in white conversation bubbles.

The research process starts in the same way as proposed in the Art case study in the previous section as well as it uses the same tools - the Generative Adversarial Neural network. The created Codebase is used as

the main framework for the code of the case study.

The main parameters are chosen, which are set by the author of the project with approximately 3 years of experience in AI prior to this case study. The main object of study is Architecture which is proposed to be an intermediate between Style and organization, the tool of the study is Artificial intelligence which in this case strives to combine analytics and Generative Adversarial Networks. The main goal of this study is floorplan analysis and generation, as well as Modern - to - Baroque floor plan translation. Furthermore, 6 metrics on which the floor plans were classified were chosen: Footprint Shape, Orientation, Thickness & Texture, Program, Connectivity, and Circulation.

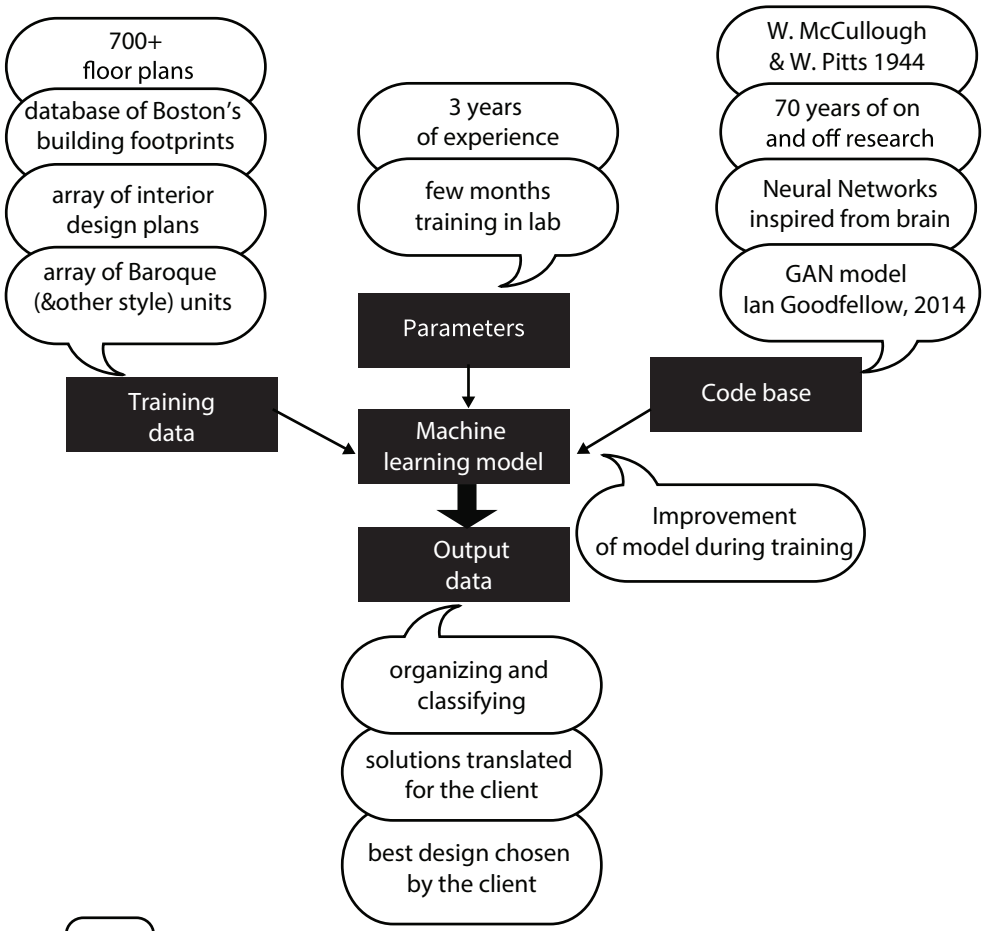
The training data is chosen and collected according to the main parameters, the 6 metrics and the goal of the study. The training set defines the quality of the outputs of the machine learning model. This is also the first step in order to start using the Generative adversarial network, in this case several datasets were created on which the GAN model was trained:

training set 1: Footprint: an extensive database of Boston's building footprints
training set 2: Program: a dataset of around 700+ annotated floor plans
training set 3: Furniture: an array of interior design plans
training set 4: Style transfer: an array of Baroque units

The Discriminator is trained to recognize images from the training set. The Codebase is used as the base for the framework of the code but slight additions and changes might be made by the author to better fit the case study.

Parallel to the Discriminator the Generator is trained to create images from the random noise (computer generated) resembling images from the same training set. As the Generator creates images, the Discriminator provides him with some feedback about the quality of its output. In response, the Generator adapts, to produce even more realistic images. Similarly, here the codebase is the main framework but slight changes and added code might appear.
Slight changes and improvements might be done of the training process by the author of the project.

Once the model has repeated this process a number of times and reached a certain level of quality in its outputs the author creates a booklet that has an array of classified solutions and floor plans from which the client can choose the most suitable design.



Parameters: *Main parameters used for the GAN model or Transformer for the specific project*

Codebase: *The main general code of the GAN or transformer*

Training data: *Massive amounts of chosen data which is used to train the machine learning model and which affects the quality of the machine learning model*

Machine learning model: *Generative adversarial neural network model or Transformer GPT-3*

Output data: *The final output, result that the machine learning model produces*

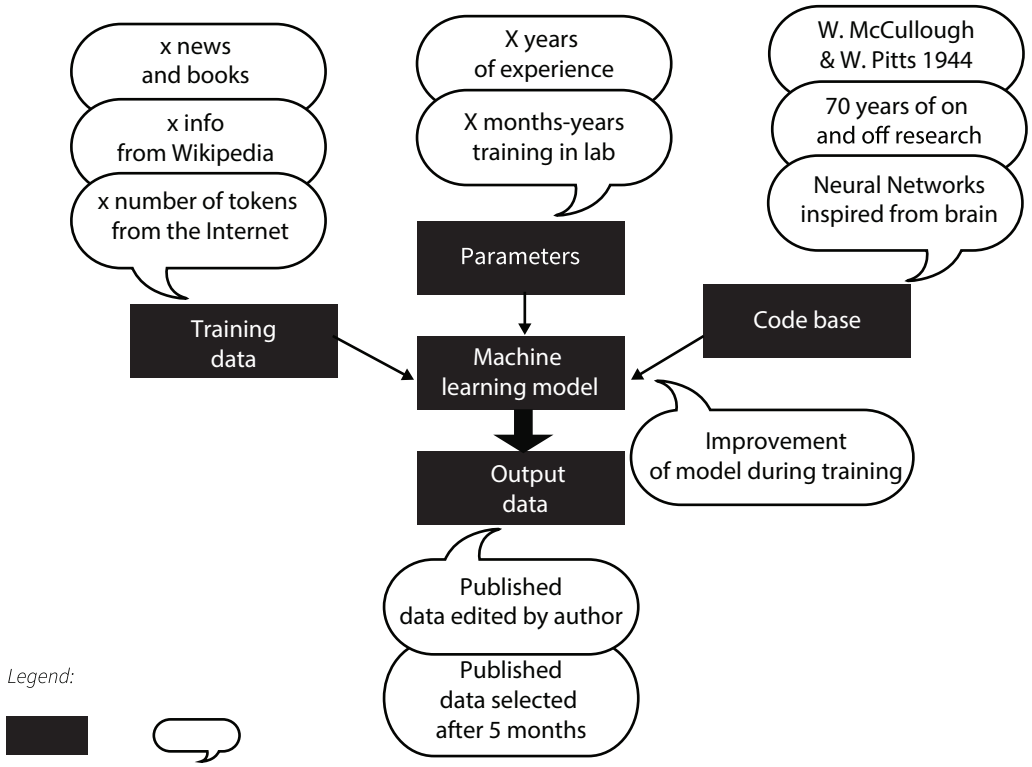
Fig. 4

SUMMARY OF THE CREATED ILLUSIONS AND HUMAN INPUTS

In all case studies the presentation of the projects through mainstream media created an illusion that reached high exposure and human interest. Partially the reaction of the audience led to believe that the Turing test had been passed for each case study, however, in due time several response articles and comments by researchers within the field of AI showed critique of the false impressions made.

The process of all case studies follows a similar pattern with minor differences, which is visualized in the Figure 5. The process in all cases starts with the research of neural networks since the 1940's untilt it led to many machine learning models we know today such as Generative adversarial neural networks (2014), the Transformer GPT-3 (2020) and to the creation of the codebase for each model. Each project required the skills to correctly set the main parameters which always meant several years of experience as well as a certain amount of time in the lab testing different variants. Once the codebase is there and parameters are set the training of the machine learning model takes place, where different datasets from books, news articles, the Internet are carefully chosen and used. Finally, once the machine learning model is giving satisfactory results the output data is taken and edited by the author and possibly also published later.

All case studies prove that there is an apparent discontinuity within the presentation of the projects to the public and the actual process.



Parameters: *Main parameters used for the GAN model or Transformer for the specific project*
Codebase: *The main general code of the GAN or transformer*
Training data: *Massive amounts of chosen data which is used to train the machine learning model and which affects the quality of the machine learning model*
Machine learning model: *Generative adversarial neural network model or Transformer GPT-3*
Output data: *The final output, result that the machine learning model produces*

Fig.5

Can generative tools autonomously create human cultures on a larger scale?

In order to reach the automation of the generative tools the human labour would need to be substituted by computer generated inputs. Let us see what in the previously studied case studies has already been done related to this topic.

In text digital voice assistants have been trained to autonomously assist people in their daily life. They use natural language generation and processing and machine learning, therefore the digital assistants learn after every interaction made with the human. However, it still requires the human to input a question, statement or other to work. As well as in the chosen case study of the Guardian article the initial human input was crucial in order for the transformer to output something meaningful. Without enough human input still the issue is to generate long coherent texts.

In art, some substitutions have already been made with the development of new generative models. The most publicized one is the AICAN (artificial intelligence creative adversarial network), created by Dr. Ahmed Elgammal at Rutgers University Art & AI Laboratory. This model has already made a collection of arts since its debut in 2017, and its outputs have been sold in renowned art galleries and fairs. The researchers have adhered to human psychology theory to understand the creativity of human artists, and concluded that in order to create a better illusion, original and fresh style must dominate. The new algorithms were written based on this concept. As such, the difference from the previous GAN model unfolds on the function of the two opposing terminals. As the creator explains : "On one end, it tries to learn the aesthetics of existing works of art. On the other, it will be penalized if, when creating a work of its own, it too closely emulates an established style."(Elgammal, A., 2018) The main objective for this model is to make it look like the AI is more creative and autonomous. The substitution happens on the input data collection. While in the GAN model, the programmer must choose a certain type of data as training sets to achieve good results, the aim of this model is to use input data with least human labour possible. As an example, the input arts don't need to be defined by specific style or type.

We see a potential for the generative models to produce products that are similar to human language, art or even architecture, but the concern is whether these outputs would create a human culture on a larger scale.

How would we define a human culture? What does it entail? Applying the idea of D.Baciu and seeing all life as digital(Baciu, D., 2021), and with digital we refer to the main mathematical property of digital systems which is that even using only a small amount of symbols it is possible to create "open-ended worlds of possibilities."(Baciu, D., 2021). Within these digital systems the evolution and exploration happens through creativity and diversification. Therefore, these will be our main parameters which we will base our analysis on whether generative tools can autonomously create human cultures in the long run.

By creative we mean something that can adapt to new circumstances and by diverse we mean something that can go into different directions, cover different topics and ideas. If not enough variety and diversity is provided it can lead to "cultural blind spots" which practically means that it can fail "to provide adequate solutions to social problems". (Maffi, L., 1998) Cultural diversity is in many ways interlinked with biodiversity, which is essential for functioning ecosystems.

For the creativity parameter, it is not known whether in the long run the machine learning models would allow for long run adaptation rather than long run entropy. By entropy we refer to "the average level of uncertainty inherent in the variable's possible outcomes". (Wikipedia, Retrieved:2020) The term was first introduced in 1948 by Claude Shannon in his paper "A Mathematical Theory of Communication", it is also referred to as Shannon entropy. "Entropy is a measure of disorder"(T., S., 2019), however, the goal of data scientists is to "reduce uncertainty".(T., S., 2019)

It is not known if we can start with the initial corpus and create new questions and new possibilities, it is known, however, that as long as you do not retrain the machine learning model, none of the new insights will flow back into the model, because the model is somewhat static, it will not adapt. Therefore, it is necessary that the model is re-trained and if the human is replaced by the machine it means that when the model is retrained you would have machine data. Which would mean that if it has the machine data we should test it beforehand to see whether what we get is just noise and entropy or adaptation.

For the diversity parameter it is similar to the creativity - if you don't retrain the model it might not allow you to create new groups of ideas, it will not adapt to new circumstances. The consequences of the lack of diversity can appear when AI models "embed human and societal biases and deploy them at scale."(Manyika, J., 2019) At ProPublica the tool "COMPAS" a "case management and decision support tool used by U.S. courts to assess the likelihood of a defendant becoming a recidivist"(Wikipedia, Retrieved: 2021) "incorrectly labeled African-American defendants as "high-risk" at nearly twice the rate it mislabeled white defendants."(Manyika, J., 2019) Therefore, underrepresented people will have their own cultures that are underrepresented in the corpus. Since you use the old corpus it will reinforce existing, potentially unwanted patterns of power distribution in society. To prevent this again you need to retrain the model with new data, where the model can create new groups of ideas based on which it is creating the new text, art or other. Here again you face the problem of having to retrain the model and here you not only question whether the model allows for new outputs that adapt to new problems but also will it actually be able to create entirely new topics and groups of ideas. Once again the only way to test this is to retrain the model.

Without testing the model using retraining we are blindly going forward where we replace our work with machines, but we cannot be sure whether we are consuming and eventually burning out our cultures with the new tools that will never create new cultures.Therefore, we would like to introduce the "Collective test" that would retrain the machine learning model with machine data and it would be used in order to see whether in the long run with the computer generated information actual new topics evolve and whether they outputs adapt to new problems. Furthermore, if this is not the case diversification and creativity does not take place, and in the long run the number of diverse cultural groups will not be able to increase and the model will not be able to adapt. Overall, without testing we cannot know which direction we are heading towards - long run adaptation and diverse outputs or long run entropy.

Conclusion

There is a lack of transparency in emphasizing the crucial choices and changes that humans make whilst working with an AI model. The analysed projects also highlight the importance and the amount of human inputs that are necessary to reach satisfying results when using machine learning models. However, through the illusion created in the mainstream media it seems that generative models are able to produce products that are similar to human language, art or even architecture, but the concern is whether these outputs would create a human culture on a larger scale. Therefore, due to the limitations of the current testing method - the Turing test which mostly focuses on whether a good enough illusion is created a Collective test is suggested which would be based on the 2 aspects that a human culture entails: creativity - to be able to adapt to new circumstances and diversity to go into different directions, cover different topics and ideas. The test means replacing the human inputs with machine data and retraining the machine learning model and seeing whether you get adaptation and new groups of ideas over a longer period of time or worse results and entropy.

Appendix

Case study 1: AI & Text

Elaboration on the process and the human inputs for the use of the GPT- 3 model

The inputs have been fed to the GPT-3 by the human, hence the process of using the GPT-3 can start:

Each word of the given sequence is embedded (an embedding function is learned: a neural network that takes a 50257-length vector of ones and zeroes, and outputs a n-length vector of numbers)(Dugas, D., 2020)
Positional encoding is done.
The multi-head attention takes places, where for each output in the sequence a prediction is made about which input tokens to focus on and how much.(Dugas, D., 2020)
The feed forward takes place which is a multi-layer-perceptron with 1 hidden layer. It takes the input, multiplies by the learned weight, adds learned bias, repeats it and gets the result. (Dugas, D., 2020)
The input is added to it's output, and the result is normalized.
The decoding takes place, after passing through all layers of GPT-3's machinery the input is processed into a matrix and each output position contains a vector information about which word should appear.
The previously learned mapping in the embedding section (step 5) is reversed and used to transform the output vector embedding into a word encoding.
A softmax function is used after which the results can be treated as probabilities for each word. Additionally, a parameter top-k is used to limit the amount of possible solutions in order to output only the most likely predicted words. If top-k = 1 we pick the most likely word.

Finally, the output is a guess of a word which then would be added to the end of the input sequence. During the process of all of these steps and the training of the model several improvements might be necessary, which would include human labour.

Bibliography

Primary sources:

Chaillou, S., 2019, "AI + Architecture, Towards a new approach" , MSc thesis, Harvard Graduate School of design, Retrieved from: https://www.academia.edu/39599650/AI_Architecture_Towards_a_New_Approach

Christies. (2018, December 12). Is artificial intelligence set to become art's next medium? Retrieved from: https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx

GPT-3, 2020, "A robot wrote this entire article. Are you scared yet, human?", Retrieved from: https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3

Secondary sources:

Baciu, D., 2021, "Digital Life: Creative and open-ended", Preprint DOI: 10.31219/osf.io/hje26
Baldwin, E., 2019, "AI Creates Generative Floor Plans and Styles with Machine Learning at Harvard", Retrieved from: https://www.archdaily.com/918471/ai-creates-generative-floor-plans-and-styles-with-machine-learning-at-harvard

Baldwin, E., 2020, "A Virtual Tour of AI & Architecture at the Pavillon de l'Arsenal in Paris", Retrieved from: https://www.archdaily.com/934191/ai-and-architecture-coming-to-the-pavillon-de-larsenal-in-paris

Biggs, J. (2021, February 6). History Of Artificial Intelligence | Meldium. Meldium | Blog for Business. https://www.meldium.com/history-of-artificial-intelligence/
Brown, T., B., et al, 2020, "Language Models are Few-Shot Learners", Retrieved from: https://arxiv.org/abs/2005.14165

Cahn, A., F., 2020, "A human wrote this article. You shouldn't be scared of GPT-3", Retrieved from: https://www.theguardian.com/commentisfree/2020/sep/12/human-wrote-this-article-gpt-3

Claypool, M., 2019, "The digital in Architecture: then, now and in the future", Retrieved from: https://space10.com/project/digital-in-architecture/

Dickson, B., 2020, "The Guardian's GPT-3-written article misleads readers about AI. Here's why.", Retrieved from: https://bdtechtalks.com/2020/09/14/guardian-gpt-3-article-ai-fake-news/

Dugas, D., 2020, "How deep is the machine? The Artificial Curiosity Series", Retrieved from: https://dugas.ch/artificial_curiosity/GPT_architecture.html

Elgammal, A. (2018, October 17). Meet AICAN, a machine that operates as an autonomous artist. The Conversation. https://theconversation.com/meet-aican-a-machine-that-operates-as-

an-autonomous-artist-104381

Google developers, 2019, “Overview of GAN Structure”, Retrieved from: https://developers.google.com/machine-learning/gan/gan_structure

Goodfellow, I., et al, 2014, “Generative Adversarial Networks”, Retrieved from: https://arxiv.org/abs/1406.2661

Groß, R., Gu, Y., Li, W., & Gauci, M. (2017). Generalizing GANs: A Turing Perspective. https://proceedings.neurips.cc/paper/2017/file/73e5080f0f3804cb9cf470a8ce895dac-Paper.pdf

Hardesty, L., 2017, “Explained: Neural networks; Ballyhooed artificial-intelligence technique known as “deep learning” revives 70-year-old idea.”, Retrieved from: https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414

Holloway, E., 2020, “Did GPT-3 really write that Guardian essay without human help?”, Retrieved from: https://mindmatters.ai/2020/11/did-gpt-3-really-write-that-guardian-essay-without-human-help/

IBM Cloud Education, 2020, “What is Machine Learning?”Retrieved from: https://www.ibm.com/cloud/learn/machine-learning

Kvochik, T.,(2019) “A Turing Test for Design“, Retrieved from: https://aexm.ai/blog/turing-test-for-design, Accessed: 11.02.2021.

Macaulay, T., 2020, “The Guardian’s GPT-3-generated article is everything wrong with AI media hype; The op-ed reveals more by what it hides than what it says” Retrieved from: https://thenextweb.com/neural/2020/09/08/the-guardians-gpt-3-generated-article-is-everything-wrong-with-ai-media-hype/

Maffi, L., 1998, “Nature and Resources”, Vol.34., No.4., October - December
Manyika, J., Sneader, K., 2018, “AI, automation, and the future of work: Ten things to solve for”, Retrieved from: https://www.mckinsey.com/featured-insights/future-of-work/ai-automation-and-the-future-of-work-ten-things-to-solve-for#

Manyika, J., SIlberg, J., 2019, “Tackling bias in artificial intelligence (and in humans)”, Retrieved from: https://www.mckinsey.com/featured-insights/artificial-intelligence/tackling-bias-in-artificial-intelligence-and-in-humans

Moral Paradigm, 2020, “Can a Society Exist Without Culture?”Retrieved from: https://moralparadigm.com/can-a-society-exist-without-culture/#:~:text=Without%20the%20aspect%20of%20culture,instinct%2C%20which%20transcends%20down%20biologically.

Offert, F. (2020, October 31). The Past, Present, and Future of AI Art. The Gradient. https://thegradient.pub/the-past-present-and-future-of-ai-art/

OpenAI, 2020, “OpenAI Licenses GPT-3 Technology to Microsoft”, Retrieved from: https://openai.com/blog/openai-licenses-gpt-3-technology-to-microsoft/

Ostwald, M.J., 1994, “A Turing Test for Architectural Design”, Exedra Vol 5, No 1

RIBA. (2020, May 12). The RIBA Plan of Work 2020. Architecture for London. Retrieved from: https://architectureforlondon.com/news/the-riba-plan-of-work/

Riedl, M.O. (2014). The Lovelace 2.0 Test of Artificial Creativity and Intelligence. ArXiv, abs/1410.6142.

Retrieved from: https://en.wikipedia.org/wiki/COMPAS_(software)

Retrieved from: https://en.wikipedia.org/wiki/Entropy_(information_theory)

Retrieved from: https://en.wikipedia.org/wiki/Generative_adversarial_network

Retrieved from: https://en.wikipedia.org/wiki/GPT-3

Smith, C., 2006, "The History of Artificial Intelligence", Retrieved from: https://courses.cs.washington.edu/courses/csep590/06au/projects/history-ai.pdf

Srinivasan, A., 2020, "The first of its kind AI Model- Samuel's Checkers Playing Program", Retrieved from: https://medium.com/ibm-data-ai/the-first-of-its-kind-ai-model-samuels-checkers-playing-program-1b712fa4ab96

Stouhi, D. (2018, July 31). How Long Does it Take to Become an Architect? ArchDaily. https://www.archdaily.com/898231/how-long-does-it-take-to-become-an-architect

Turing. A. M. , (1950) "Computing Machinery and Intelligence", Mind 49:433-460
T., S., 2019, "Entropy: How Decision Trees Make Decisions; The simple logic and math behind a very effective machine learning algorithm"Retrieved from: https://towardsdatascience.com/entropy-how-decision-trees-make-decisions-2946b9c18c8

Vincent, J. (2018, October 23). How three French students used borrowed code to put the first AI portrait in Christie's. The Verge. https://www.theverge.com/2018/10/23/18013190/ai-art-portrait-auction-christies-belamy-obvious-robbie-barrat-gans

COLLECTIVE TEST:

CAN GENERATIVE TOOLS WORKING AUTONOMOUSLY CREATE NEW CULTURES IN THE LONG RUN?

Architectural history thesis
spring 2021