Faculdade de Design, Tecnologia e Comunicação
Universidade Europeia

2024

William Albuquerque Terra

Artificial Comedy: Assessing the potential of generative AI in Visual Humour

MOD-195.IADEV02; 07-07-2023

Faculdade de Design, Tecnologia e Comunicação
Universidade Europeia

2024

William Albuquerque Terra

Artificial Comedy: Assessing the potential of generative AI in Visual Humour

Dissertação apresentada ao IADE - Faculdade de Design, Tecnologia e Comunicação da Universidade Europeia, para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Design e Cultura Visual realizada sob a orientação científica do Doutor Eduardo Corte-Real, Professor Catedrático do IADE - Universidade Europeia

MOD-195.IADEV02; 07-07-2023

Dedico este trabalho à todos aqueles que têm sede de conhecimento e ostentam a curiosidade de melhor entender o mundo.

This work is dedicated to all those who are thirsty for knowledge and have the curiosity to better understand the world.

This dissertation is written in English (UK). However, it might be possible to find some mixing in spelling and writing styles between American and UK English.

Agradecimentos Agradeço aos meus amigos e familiares e também aos inúmeros professores que tive ao longo da minha jornada até aqui.

Palavras-chave Cultura Visual, Inteligência Artificial, Humor e Estética.

Resumo

Esta dissertação investiga o uso das recentes tecnologias de inteligência artificial generativa no campo do humor visual, explorando o potencial dessas tecnologias na criação de imagens humorísticas popularmente disseminadas pela internet na forma de memes. Espera-se que as novas tecnologias de inteligência artificial revolucionem diferentes aspectos da sociedade humana. Durante algum tempo, pensou-se que o campo criativo estava excluído da interferência de tais ferramentas. No entanto, a popularização da IA generativa demonstrou uma grande capacidade dessas ferramentas em criar imagens de alta qualidade. Através de uma revisão abrangente da literatura que abrange os campos da história da arte, estudos de humor, estética, cultura visual e inteligência artificial, este estudo lança luz sobre a dinâmica complexa da criação e recepção do humor visual, especialmente no contexto da cultura da internet. Para obter uma análise adequada, alguns métodos de análise visual foram aplicados para complementar a revisão da literatura. Os métodos incluíram a avaliação de memes da internet criados por diferentes criadores de conteúdo humorísticos e a tentativa de reproduzir memes famosos usando ferramentas de IA generativa. A pesquisa revela que o humor visual opera como uma linguagem cultural sutil, altamente dependente de referências culturais específicas para a sua eficácia. Além disso, a investigação explora as limitações inerentes das ferramentas de IA na reprodução dessas nuances culturais visualmente, limitando assim sua capacidade de se tornar uma ferramenta eficaz para criar humor visual. Além disso, o estudo destaca o desafio de delinear a fórmula do humor, uma tarefa altamente complexa para os humanos e extremamente limitada para ser traduzida em dados matemáticos, que é a forma como os computadores operam. Também existem dificuldades encontradas na tradução de comandos escritos para elementos visuais. Essas limitações destacam a necessidade de uma compreensão mais profunda das complexidades culturais e contextuais do humor e das limitações atuais da inteligência artificial.

Keywords Visual Culture; Artificial Intelligence, Humour and Aesthetics.

Abstract

This dissertation investigates the use of the recent technologies of generative artificial intelligence in the field of visual humour, exploring the potential of these technologies in the creation of humorous images popularly spread through the internet in the form of memes. The new technologies of artificial intelligence are expected to revolutionise different aspects of human society. For some time it was thought that the creative field was excluded from the interference of such tools. However, the popularisation of generative AI has shown a great capacity for these tools to create high-quality images. Through a comprehensive literature review that englobes the fields of art history, humour studies, aesthetics, visual culture, and artificial intelligence, this study sheds light on the complex dynamics of the creation and reception of visual humour, particularly in the context of internet culture. To obtain a proper analysis, some methods of visual analysis were applied to complement the literature review. The methods included the assessment of internet memes created by different meme creators and the attempt to reproduce successful memes using generative AI tools. The research reveals that visual humour operates as a nuanced cultural language, heavily reliant on specific cultural references for its efficacy. Furthermore, it uncovers the inherent limitations of AI tools in reproducing these cultural nuances visually, hence limiting their ability to become an effective tool for creating visual humour. Moreover, the study highlights the challenge of delineating the formula of humour, a task that is highly complex for humans and extremely limited to be translated to mathematical data which is the way computers operate. There are also difficulties encountered in the translation of written words into visuals. These limitations underscore the need for a deeper understanding of the cultural and contextual intricacies of humour and the current limitations of artificial intelligence.

1. Introduction

2. The Complexity of Humour

3. Memes and The Online World

4. AI and Computer Visuality

5. Visual Humour Analysis I: Quantitative

6. Visual Humour Analysis II: Qualitative

Index of Images

Index of Graphs

Index of Diagrams

1.0 - Introduction

In an age of fervent technological development, artificial intelligence is now considered the breaking point of all human inventions; for the first time, humans are developing a technology that might be capable of thinking for us (Harari, 2023). In many areas of society, including the field of visual culture, the impact of AI is causing great concern. New tools with artificial intelligence are already being used. Image-generating technologies such as DALL·E 2, Stable Diffusion, and Midjourney are easily accessible to internet users who can generate impressive images quickly, requiring no artistic skills, only using words as prompts. This new type of technology is called generative artificial intelligence. Anyone with access to high-speed internet can generate images without the need to have advanced artistic skills.

The presence of artificial intelligence in visual culture has enormously increased in the past few years. Many designers and artists are now incorporating AI tools in their creative process. Artificial Intelligence is now present in different parts of visual culture, such as advertisement, illustration, animation, film and graphic design (Kaput, 2022). Big tech companies such as Alphabet, Microsoft, and Meta are investing billions of dollars in developing AI tools (Leswing, 2023). Adobe, one of the most substantial companies in the visual culture industry, has already implemented its own prompt image-generating tools (Still, 2023). Despite the widespread use of generative AI, one specific part of visual culture seems to be less affected by these new tools: The area of visual humour.

Visual humour is probably the most consumed content on the internet, and it has been popularised in the last decade, mainly in the form of memes. According to the Oxford Dictionary (n.d), a meme can be defined as “an image, video, piece of text, etc., typically humorous in nature, that is copied and spread rapidly by Internet users, often with slight variations” The most popular form of internet memes is the

‘humorous memes’; this category is by far the most widespread form of content (Shifman, 2015).

Nowadays, millions of images are created and shared on social media by numerous users worldwide. With the new technologies of smartphones, the sharing mechanisms of social media significantly increased the reach of communication. A common person can create an image, share it and reach large audiences outside their local community.

Among those images, there are thousands of memes. Humorous memes take form in images and videos, usually with minimal editing needed. This form of visual content is extremely popular and requires little effort to be created and distributed (Shifman, 2015). Curiously, the AI tools seem to have had minimal interference in this type of visual production. Despite image-generating AI tools' benefits, such as speed and minimum technical skills, these new tools seem to be little appealing to meme creators. According to the data acquired in this research, only a tiny number of memes utilise AI tools.

Due to the theme's importance, this investigation proposes to explore the potential of generative artificial intelligence in visual humour by assessing computer vision capabilities and researching the complex factors that influence humour. This study aims to clarify the following questions: Is visual humour the last frontier of human creativity against AI? Why is the presence of artificial intelligence in the creation of visual humour still minimal? What is the potential of generative AI in the creation of visual humour?

To explore these questions, this dissertation is divided into four main parts. The first section is focused on a literature review of humour, memes and artificial intelligence. The second part is dedicated to analysing existing visual humour data. The third section has a pragmatic approach and will test the capacity of AI to recreate successful memes. The last chapters focus on discussing the outcomes of the experiments sustained by the theoretical arguments.

1.1 - Methodology

This investigation encompasses mostly the scientific areas of humanities, focusing on visual culture and humour. To approach these questions and develop a well-defined hypothesis, this dissertation analysed the visual humour production of 6 different humorist pages on Instagram's social media platform (3 in English and 3 in Portuguese). The criterion for the selection was the high number of followers, all content creators have more than one million followers which indicates a strong memetic and humorous power. Due to the complexity of the theme, the data analysed focused primarily on static memes, although some videos were included. The comic images were analysed using two methodologies proposed in the book "*The Handbook of Visual Analysis*"(2001) by Theo van Leeuwen and Carey Jewitt. The first method is the Visual Content analysis, which is focused on quantitative analysis and provides overall information about the sample, unveiling the main topics, themes and characters in the humorous images. This dissertation selected 20 images per content creator, 120 images in total. The images were chosen according to the criteria of time and engagement; they were not older than 12 months before the writing process of this document. In addition, images with high engagement were prioritised.

The second method is the Semiotic and Iconography Analysis; this method provides a qualitative approach to the data, exploring the deeper meanings and the cultural aspects. For this analysis, six images were selected. This sample included AI-generated images. The Semiotic and Iconography Analysis method contributes to a deeper study of the cultural aspects, icons, symbols, colours, hierarchies, context and other attributes of the images.

The third method used in this research was an empirical analysis of the AI tools. As a pragmatic exercise, three famous memes were selected and reproduced using the three main image-generating AI tools: DALL·E 2, Midjourney, and Stable Diffusion.

The researcher attempted to recreate these three famous memes using only AI image-generating tools. Midjourney was the AI tool primarily used in this experiment. The images generated were analysed in comparison to the original memes using semiotic and iconographic analysis.

In addition, this dissertation used a literature review of books, journals, and academic papers in multidisciplinary areas, including Media Studies, Memes, Humour, Art History, Visual Culture, Aesthetics, and Artificial Intelligence. The literature review is an essential part of the study and the foundation of the information discussed in this paper. This study explores the potential of the new AI image-generating technologies and the complexity of visual culture and humour.

What will be the impact of these tools in the creation of visual humour? Is there already enough information to evaluate the potential of these tools? To better understand this topic, it is essential to explore what creates humour and how artificial intelligence works.

2.0 - The Complexity of Humour

It is not for the lack of theories that the subject of humour is still a mystery. Over the centuries, many theories have been developed to explain why we find some things funny and others not. Why do we laugh? What is the secret of humour? The world must still agree on the right answer (McGraw & Warner, 2015).

Humour plays a vital role in our lives. People often mention humour as essential, and many couples rank the sense of humour as one of the best traits to look for in a partner (Morreall, 2020). Humour is also one of the key attributes of memes, which heavily rely on the mechanisms of humour to become viral.

The term "humour" is relatively recent, gaining popularity only in the 18th century; before that, the terms "laughter" and "comedy" were predominantly used interchangeably. Much of the academic discourse on humour before the 20th century concentrated on scornful and mocking humour rather than wit and jokes. The Greek philosopher Plato criticised laughter, viewing it as an emotion that undermines rational self-control, and he argued that laughter was often malicious in nature.

As described by Morreall (2020) in the online article "*Philosophy of Humor*", "Another of Plato's objections to laughter is that it is malicious. In *Philebus* (48–50), he analyses the enjoyment of comedy as a form of scorn. "Taken generally," he says, "the ridiculous is a certain kind of evil, specifically a vice." That vice is self-ignorance: the people we laugh at imagine themselves to be wealthier, better looking, or more virtuous than they really are. In laughing at them, we take delight in something evil—their self-ignorance—and that malice is morally objectionable"

Aristotle, another Greek philosopher after Plato, also expressed opposing views on the subject. According to Aristotle, laughter was a sign of scorn, but wit was still a valuable part of the conversation. The Stoics agreed with Plato's ideas of laughter,

describing it as an uncontrolled emotion, often malicious and a sign of lack of self-control. Early Christians later supported this negative view on the topic. The idea of humour as malicious and hostile is encountered in some Bible verses. The Bible mentions God's laughter with a hostile tone:

"The kings of the earth stand ready, and the rulers conspire together against the Lord and his anointed king…. The Lord who sits enthroned in heaven laughs them to scorn; then he rebukes them in anger, he threatens them in his wrath (Psalm 2:2–5)."

The depiction of laughter as a negative sign used for scorn survived many centuries among Christian leaders. The Christian monasteries condemned laughter and advised monks to avoid jokes and foolish chatter. In some cases, laughter could even cause physical punishment.

"In Benedict's Ladder of Humility, Step Ten is a restraint against laughter, and Step Eleven is a warning against joking (Gilhus 1997, 65). The monastery of St. Columbanus Hibernus had these punishments: "He who smiles in the service … six strokes; if he breaks out in the noise of laughter, a special fast unless it has happened pardonably" (Resnick 1987, 95) (Morreall, 2020).

The rejection of laughter continued in most Christian societies, including the protestant movements of Reformation. The Puritans continued to condemn comedy and advised people to live sober lives and not give space for lascivious vanities. Consequently, comedy was outlawed in mid-17th-century England (Morreall, 2020). During the same period, Rene Descartes and Thomas Hobbes described laughter as a sign of our competitiveness and individualistic behaviour associated with winning and losing. Following this reasoning, humans laugh mostly because of feelings of superiority and mockery (Morreall, 2020).

2.1 - The Superiority Theory

The predominance of the theory of comedy as a way to scorn and express superiority led to the "Superiority Theory". According to this view, sustained until the XX century, laughter expresses our feelings of superiority over others and former states of our being. Roger Scruton, a defender of this theory, makes his point when saying if people dislike being laughed at, it must be because laughter somehow devalues them. On the other hand, Francis Hutcheson criticised the Superiority Theory, claiming that laughter was not limited to comparison and feelings of superiority. Humans might feel superior to other animals and beings and still not produce laughter; the case of piety is a good example. Some people have feelings of sorrow when seeing people in inferior or vulnerable situations. Also, some people laugh not at the previous version of themselves but at the current self, for instance, when a surprise happens. The superiority theory fails to explain compassion and sudden joy (Morreall, 2020).

2.2 - The Relief Theory

There is another theory for humour called *"The Relief Theory"*. The Relief Theory has nothing to do with feelings of superiority but is related to the anatomic structures of our bodies. The theory was first formulated by Lord Shaftesbury in 1709. It was one of the first times the term humour was used to describe funniness. Lord Shaftesbury explained that humour was caused by the release of pressure in the nerves. Centuries later, other thinkers such as Herbert Spencer and Sigmund Freud reformulated the theory but kept the idea that laughter is a release of energy in the brain nerves. Spencer's theory is based on the idea that emotions take a physical form of energy that runs through the nerves. Hence, laughter releases energy with no specific goal, unlike other body symptoms such as anger that may cause us to

clench our fists. In the Relief Theory, laughter is just the mere release of energy accumulated from inappropriate emotions (Morreall, 2020).

Sigmund Freud also had a version of the Relief Theory. In his work "*Jokes and their Relation to the Unconscious* (1905)," Freud analyses laughter in different situations: jokes, the comic and humour. In all three situations, laughter is the release of energy accumulated for a specific psychological task that later became superfluous. In the case of jokes, energy is used to repress feelings. In the case of the comic, it is the energy used to think, and in humour, it is the energy of feeling emotions.

The term used by Freud to describe jokes is *Der Witz*. The term includes telling spontaneous comments and preparing jokes. According to this theory, the psychic energy used to repress emotions escapes through laughter. The feelings which are mostly repressed are sexual desires and hostility, so most jokes have sexual or hostile connotations. When people listen to a joke that puts down an individual or group they dislike, they let out the energy for the hostility that is usually repressed; this energy release assumes the form of laughter (Morreall, 2020).

According to Freud's theory, the comic also involves a release of energy, but it is the energy that is found unnecessary and is normally devoted to thinking. An example is laughter in clumsy situations. As people watch a clown do clumsy things, the energy used in the expectation of watching an action being performed is normally released in the form of laughter. Freud's analysis of humour is similar to Spencer's; it is the release of energy accumulated by expecting the performing of an action that did not occur.

Nowadays, few contemporary thinkers support Spencer and Freud's ideas. Of course, energy is spent when laughter occurs, involving movements and muscles of different parts of the body. Nevertheless, the idea that this energy was accumulated from repressed emotions seems incorrect because not all situations provoke strong emotions and release unnecessary energy in the form of laughter.

“If Freud is right that the energy released in laughing at a joke is the energy normally used to repress hostile and sexual feelings, then it seems that those who laugh hardest at aggressive and sexual jokes should be people who usually repress such feelings. But studies about joke preferences by Hans Jürgen Eysenck (1972, xvi) have shown that the people who enjoy aggressive and sexual humour the most are not those who usually repress hostile and sexual feelings, but those who express them.” (Morreall, 2020).

The concept of energy release does not seem to explain why people laugh harder at different things and why some people accumulate more energy and others do not. All these flaws in the Relief Theory contribute to a high rejection among philosophers and psychologists studying humour today (Morreall, 2020).

2.3 - The Incongruity Theory

The third alternative to explain humour is the “*Incongruity Theory*”. The theory says the perception of something incongruous provokes humour. This theory was supported by Arthur Schopenhauer, Søren Kierkegaard, Immanuel Kant and many later philosophers and psychologists. This approach is now the most acceptable among the academic community. This idea is, therefore, nothing new; philosophers like Aristotle and Cicero proposed the idea of a break of expectation.

“Cicero, in *On the Orator* (ch. 63), says, “The most common kind of joke is that in which we expect one thing and another is said; here our disappointed expectation makes us laugh”(Morreall, 2020).

The break of expectation is a common technique used by many standup comedians today. The first part of the joke explains the context and creates expectations. The punchline is the break of expectation, the incongruity in the story, the part that makes us laugh.

"A joke amuses us by evoking, shifting, and dissipating our thoughts, but we do not learn anything through these mental gymnastics. In humour generally, according to Kant, our reason finds nothing of worth. The jostling of ideas, however, produces a physical jostling of our internal organs, and we enjoy that physical stimulation" (Morreall, 2020).

Kant compares the enjoyment of jokes with games of luck and the enjoyment of music. In all these three experiences, there is a change and a play of sensations caused by shifting ideas in the mind. This shifting of ideas, sounds and sensations brings pleasure to jokes, music and games of fortune (Morreall, 2020).

Like Kant, Kierkegaard also defended the incongruity theory, calling the comical as a disparity between what is expected and what is experienced; this contradiction is what causes humour. The contraction between expectation and reality is what makes errors comic. The violation of expectation is at the heart of both tragedy and comedy. Kant used a concept from Aristotle's definition of the Comic: "The ridiculous is a mistake or unseemliness that is not painful or destructive." The contradiction that does not produce pain is comic; the break of expectation that creates suffering is a tragedy.

The incongruity cannot always cause enjoyment and pleasure, there are several occasions in which a contradiction occurs and the results are anger and frustration, instead of the pleasurable laughs produced by humour. The Incongruity Theory is now refined in the version of the "*Benign Violation Theory*". The benign violation concept complements the incongruity with the "benign" factor. The contradiction must fit the requirements of something to be considered benign to occur; if the receiver of the joke judges the incongruity to cause "something evil", the humour will not happen, and the response will be instead anger and frustration.

2.4 - Problems with the Incongruity Theory

Despite the Incongruity Theory being the most acceptable theory for humour, some unresolved questions remain. If humour comes from the enjoyment of violations of patterns and expectations, then it is irrational. This irrationality in humour does not require intellectual enjoyment, which is still questionable; there are numerous examples of humour as the promoter of intellectual benefits. Remarkably, only a few philosophers have acknowledged the benefits of joking.

"In the last century, an early play theory of humour was developed by Max Eastman (1936), who found parallels to humour in the play of animals, particularly in the laughter of chimps during tickling. He argues that "we come into the world endowed with an instinctive tendency to laugh and have this feeling in response to pains presented playfully" (45). In humour and play generally, according to Eastman, we take a disinterested attitude toward something that could instead be treated seriously."

Many psychologists confirm that humour can be virtuous. Engaging in humour can increase tolerance for diversity and ambiguity. It can also improve problem-solving skills. Humour can also be used as a tool to reduce stress and conflicts. It reduces the negative emotions attached to these types of information, and bad news can be less painful when accompanied by humour.

2.5 - Biological Explanations for the Grin Face

The main physical signs of humour are laughing and smiling. These signs developed from similar play signs from pre-human apes, such as gorillas and chimps, who usually play in the form of chasing, biting and tickling. These forms of play can be classified as mock aggression. In this type of playful fake aggression, all participants must be aware that it is not an actual aggressive activity. If there were no way an

animal could distinguish between mock aggression and a real one, the response would be with deadly force. Animals developed facial and gestural expressions such as smiles to distinguish these actions better. The “grin face” is a sign apes use to indicate playfulness. It consists of a facial expression in which the jaws are closed, but the corners of the mouth and lips are retracted to expose the gums and the teeth; there is no body movement, and the eyes look directly towards the interacting partner. The ape’s grin face is related to the human’s smile. Similar to the grinning face, apes also make sounds that help categorise their actions while playing with each other; these sounds are similar to human laughs. Our laughs would have evolved from these sounds and transformed according to our body anatomy and upright standing position. As described by Morreall (2020) in his article:

“The hypothesis that laughter evolved as a play signal is appealing in several ways. Unlike the Superiority and Incongruity Theories, it explains the link between humour and facial expression, body language, and the sound of laughter. It also explains why laughter is overwhelmingly a social experience, as those theories do not “ (Morreall, 2020).

2.6 - Visual Humour - A Short History

Humour can take many forms, especially visual formats. Exploring the history of visual comedy is essential to better understanding how visual humour works.

“Visual humour has historically used several media and formats for this communication to take place, engravings or prints on stones, ceramics, papyrus parchment or paper and in the present day the screen and digital transmission” (Junco, 2015, p. 11).

The first visual forms of comedy can be traced back to prehistorical times in the form of cave paintings. Some of these paintings depict humans with exaggerated features and animal heads or body parts, a common feature of visual humour. However, the

meaning of these visual codes will always remain speculation and cannot be assumed with certainty to be humorous.

Figure 1. Kylix with portrait of Aesop with a fox.
Source: Wikicommons/Vatican Museum Italy
(Unknown Greek artist, 470 BC).

There are indications that the grotesque and the extravagant have been used for thousands of years to produce amusement. Our ancestors from many cultures have used visual tricks, metaphors, parodic representation, and satires to convey many purposes, some of which convey social norms, while others amuse and produce laughter (Junco, 2015). These drawings usually depict distorted and exaggerated faces and bodies on purpose. These distortions aim to emphasise striking features generally for amusing or mocking purposes. Therefore, a vast variety of possibilities range from childish to intellectual and profound intentions, from private jokes to public mockery. Caricatures can utilise the following visual resources:

"Graphic winks such as thick lines, overly simple contours, artificial tone of voice, childish colour contrasts etc., but also comic forms (grotesque details, big nose,

absurd hair crossed eyes, exaggeratedly small bodies, weird clothing etc., in other words, everything that invites the complicity of the spectator which has a ridiculous, relaxing, friendly or ingenuous appearance (Junco, 2015, p. 6)"

One of the most common formats of caricature is the portrait, in which the image offers a radical synthesis of facial features and a hierarchical selection of the person's characteristics. It highlights what is most memorable about the person's body and face. The main objective of caricatures is to capture and depict the person using contradiction, reduction and exaggeration, excluding the dull characteristics and enhancing chosen features. To communicate successfully, caricatures must have the correct form according to the audience and express symbols and ideas that resonate with the audience. The visual codes of the image need to be shared by both the receiver and the message's producer, which ensures that the communication will be rightly interpreted. As mentioned by Junco (2015), "Any humorous work must therefore be carefully and precisely constructed, as it is not easy to devise the mechanism of comedy and make it work. Transgressing the normality of the habitual schemas used for daily life, whether through frivolity or formal imagination or through a new concept that distances and questions what came before, always implies a creative effort and an understanding of the temporal and cultural environment in which this action takes place" (Junco, 2015, pp. 8-9)

Humour has a social character and works as a collective action. Visual humour is a powerful tool for connecting groups of people. It questions established norms and patterns, suggesting something surprising, whether the novelty is silly and banal or intelligent and profound. It can have an aggressive or innocent tone.

2.7 - Humour as a Game

Humour can be understood as a game that involves established patterns, automatisms, and rules. These patterns can be readjusted and removed according to

the creator's preference. Consequently, these visual codes must connect the author and the audience since humour is a pleasurable communication with an undeniable social character (Junco, 2015).

Because of its collective nature, comedy is often linked to systems of signification. This connotation factor refers to the meaning something acquires through associating with another. These transfers of meaning are deeply attached to culture. The relativity of humour content depends on certain factors such as who created it, for whom, when and where.

“The concrete environment will determine the meaning of the message. Humour, by definition, questions collective patterns, eliminating established cultural clichés. Humour provides a way of escaping reality through surprise. The cultural paradigm within which it is realised is the place where concepts, symbols and forms are connected unusually. Thus, when it finds a complicit audience, the establishment and recognition of context will also lead to the possibility and recognition of its transgression” (Junco, 2015, p. 17)

Developing a formula for this reaction is more complex than it seems. Firstly, humour involves different ways of communication. It can englobe our visual and hearing senses and can be transmitted using only words, written or spoken.

2.8 - The Benign Violation Theory

One of the attempts to develop a theory of humour was made in the article called "*Theory of Humor*", written by Thomas Veatch (1998), who proposes a formula of N+V. The idea is that humour occurs when a (V) violation of a subjective moral principle occurs, and paradoxically, the situation is still labelled normal (N) (McGraw & Warner, 2015).

The violation idea is that there is some incongruity, an inconsistency between what people expect and what actually happens; this element of surprise then produces laughter, the disproportion between the expectation and what actually happened (McGraw & Warner, 2015).

According to the principle of normal violation, humour occurs when something seems unsettling, such as when a breach of an expected action happens. Still, the violation simultaneously seems okay and acceptable to the individual. The example of someone falling down the stairs could be used. The act of falling violates the expectation of taking the stairs step by step. The first element of humour in this story is the fall. Therefore, the violation of an expected action is not enough to produce humour; the detail of what type of fall is crucial for the humour to be created. What if the person got severely injured from the fall? If that is the case, there is no humour in the situation; however, if the person did not get hurt, the violation is considered benign; hence, the fall is humorous. Following this reasoning, humour is a violation that is judged to be acceptable and benign (McGraw & Warner, 2015).

Dirty jokes are an example of how tricky the formula of humour is. Typically, dirty jokes play with moral and social values; they violate social norms, but the joke will produce a laugh only if the listener is liberated enough to find the violation acceptable. " Sarcasm violates conversational rules by meaning the opposite of what's said " (McGraw & Warner, 2015, p. 28).

The same benign violation principles can be applied to visual humour. The humour created by memes depends on this subtle line between what is expected and acceptable for the viewer of the images. The large variety of digital subcultures enhances even more the complexity of what can be considered a benign violation. Certain violations that are acceptable for specific groups can be rejected by others.

2.9 - The Ambiguity of Humour

A good example of this ambiguity is the memes produced by rival political groups. For one of these groups, the same visual humour is judged to be an acceptable violation, but for others, the content can be classified as hateful. This ambiguity of visual humour has been used as a tool by far-right extremists to disguise their visual propaganda, as shown in a documentary called "*How the Far Right Use Memes to Recruit*" by VICE NEWS (2021).

The cultural elements of a joke can be adapted to different audiences while maintaining a similar structure. The internet joke "*Upgrading from Girlfriend 4.0 to Wife 1.0*" is a good example. The joke is in text form and has numerous variations in English and other languages. The story plays with the husband and wife dynamics, mixing with computer vocabulary. The joke was adapted according to the cultural context of different countries; nevertheless, "Out of the nine languages examined in the study, the only one in which translated texts for the joke was very rare was Arabic. A possible explanation may be culture-related norms encoded in the joke, such as premarital sexual relationships may be considered unsuitable in Arab-speaking societies" (Shifman, 2015, p. 122).

2.10 - Visual Incongruity

The incongruity between two elements is a common mechanism used in many genres of visual humour. A character frequently stands out as alien to the surrounding situation. This dissonance of elements can be caused by juxtaposing photos, connecting two opposing elements and generating a violation of expectation. The technical aspects of this juxtaposition also play a role in the humour effect. Memes usually look badly edited and clearly "Photoshopped". Cheap visual editing can violate expectations because the editing is supposed to be done correctly so that

the elements in the image match harmoniously, not causing dissonance (Shifman, 2015).

3.0 Memes and the Online World

"There is no such thing as Art. There are only artists."

(Gombrich, 1962, p.15)

There is no such thing as art, no specific definition that can englobe what art is; still, humans have created millions of objects that can be classified as art, from cave paintings to digital humorous images circulating on the internet.

We are aesthetic beings, and unlike most animals, we have the urge to communicate concepts visually. Thousands of years ago, our ancestors decided to mark lines on the walls of caves and on the surface of objects. These lines developed into more complex figurative drawings that allowed us to communicate our abstract and fictional concepts. From these mysterious markings inside caves to digital images, humans now have access to thousands of images inside our pockets. Images and videos are now our main form of communication.

Videos and pictures are now accessible almost everywhere. Our society is more visual than ever, and it is estimated that in the 21st century, humans have produced more information than in the entire history of humanity (Marr, 2018). The machine-made images allowed the visual culture to be mass-produced in enormous numbers; now, millions of people participate in the cycle of creation, reproduction and sharing of visual culture. The easy production of images gave rise to the possibility of multiple types of visual cultures used by different social groups as a way to identify themselves (Barnard, 1998).

With the development and popularisation of the internet, the creation and distribution of information became easily available to anyone with access to the internet and a computer-like device. The media generated by computers started to

become accessible to the public only by the late 1980s and allowed the production of images faster than any previous technologies. Consequently, computer-generated images became part of other media forms like magazines, films, and videos. The popularisation of the personal computer empowered individuals who could now run various image-creation tools on their private machines (Barnard, 1998).

According to the Website *Domo -Data Never Sleeps* (2024), in 2021, 65.000 images were shared every minute on the social media platform Instagram. That is a massive amount of visual content, especially compared to any other time in history. The amount of time people spend online has also increased in the past decade. The average person is estimated to spend 40% of their waking hours connected to the internet. The “screen time", which is the time people spend looking at screens either from a computer or mobile, is about 7 hours and 4 minutes per day in the United States. In Portugal, it is about 7 hours and 56 minutes, and in Brazil; 10 hours and 19 minutes (Howarth, 2024).

“As of April 2022, the internet reaches 63% of the world’s population, representing roughly 5 billion. Of this total, 4.65 billion — over 93% — were social media users. According to Statista, the total amount of data predicted to be created, captured, copied and consumed globally in 2022 is 97 zettabytes, a number projected to grow to 181 zettabytes by 2025” (Domo, 2023).

These images are created and shared by millions of internet users who live all around the globe and are part of different cultures and societies. The Internet allows different parts of the world to be connected and to influence each other, hence increasing the speed of information creation and dissemination.

The phenomenon of interconnections between societies, cultures and individuals worldwide can be described as globalisation. The speed and scale of these global flows have accelerated dramatically in recent decades, affecting almost every aspect of our societies. This worldwide network transcends national borders and sometimes unifies different geographical groups. However, this cyberspace is

shaped by the material world and the cultural and linguistic contexts of these groups (Shifman, 2015).

"The Web was invented so physicists could share research papers. Web 2.0 was invented so we could share cute pictures of our cats. The tools of Web 2.0, while designed for mundane uses, can be extremely powerful in the hands of digital activists. Especially those in environments where free speech is limited"

Ethan Zuckerman (2008), "*The Cute Cat Theory of Digital Activism*" (as cited in Shifman, 2015, p. 92)

These millions of internet users are consumers and producers of visual content. Social media platforms are the leading platforms for sharing and consuming content. Internet users can use these platforms to create their own content and share their images and ideas; they can also follow other social media users.

The content produced by the users can be validated through the "like button" and the comment section of each post. In the digital society, users with significant followers are considered influential and have the title of "influencers". The term influencer is more than a title; it is also a profession; influencers can be considered modern entrepreneurs; they have their brands based on their content and personality. Brands and companies pay influencers to access their audience and promote their products and services (Wilkinson, 2023).

Amid this flood of information, users share multiple sorts of content, from trivial information about their personal lives to more informative content, such as information about politics, sports, news, science and others. In the 1990s, during the early days of the Internet, it was possible to check all the existing websites in a directory. Now, there are more than a billion websites live on the internet stats.com (Sautoy, 2020).

It is estimated that this online world is inhabited by around 5 billion users, more than half of the world's population (Petrosyan, 2024). This massive number of users

can be divided into different groups, each with different tastes and interests. All the data produced and consumed on the Internet is stored and fed to algorithms.

The internet algorithms are a set of mathematical data organised into a series of steps. These instructions are the basis of the online world and also artificial intelligence. Moreover, Data has become the main ingredient and the main asset of many tech companies, 90 per cent of the world's data has been created in the past five years. It is estimated that 1 exabyte of data is made on the internet daily; this amount of information could be stored on about 250 million DVDs. We have reached a point where humankind is now producing in two days the same amount of data it took us from the dawn of civilisation until the early 2000s to generate. This massive flood of data is one of the most important catalysts for the age of machine learning (Sautoy, 2020, p. 66).

Computers started as huge machines developed to help process data; they were used by scientists to perform mathematical calculations; now, they are many times more powerful and are owned by millions of common individuals (Walker & Chaplin, 2015).

Most of the data produced and consumed by smartphones is visual content. The revolution caused by computer art was similar to photography; a computer could generate many copies of the same image and democratise visual culture. Nowadays, smartphones are equipped with high-quality cameras and are capable of executing advanced graphic tools so that ordinary people can become image creators (Walker & Chaplin, 2015).

Visual content production is part of our natural aesthetic behaviour, which is one of the ways we engage and stimulate ourselves. It is safe to say that vision is our dominant sense. Human beings are visual animals; their brain has a strong capacity to process visual information, and the area devoted to processing visual stimulation makes up about a third of the total size of the brain. (Høgh-Olesen, 2019). The new communication technologies such as the internet, social media and smartphones are

the perfect feast for our brains, they transform the world and how humans communicate, think, feel and behave (Walker & Chaplin, 2015).

What do internet users communicate?

Amidst the large types of content, humour occupies an important role. Humorous content is usually consumed in the form of memes.

The word meme was first coined by Richard Dawkins in 1976 to describe small pieces of culture spread by imitation (Shifman, 2015). Memes are like cultural genes that spread from person to person. The word "meme" derives from the Greek "*mimema*" which means something that is imitated. The term Internet meme is normally applied to describe the propagation of content such as videos, jokes and images from person to person via the Internet. Like genes, memes undergo variation replication, selection and retention (Shifman, 2015).

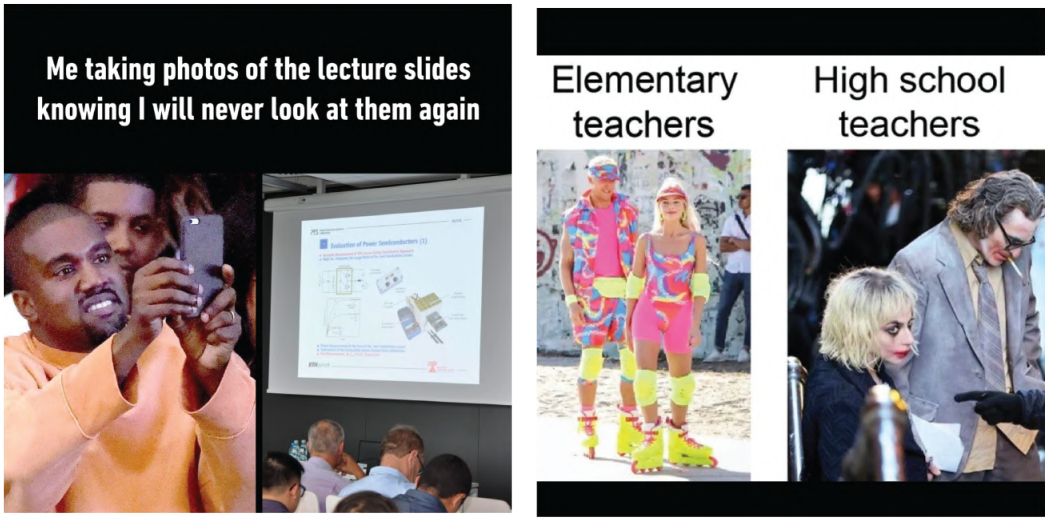


Figure 2. Examples of Memes (Source: Instagram @9gag, 2023)

The concept of meme englobes more than the propagation and imitation of images; however, for practical reasons, the term meme will be used in this investigation to refer to humorous images that are popularly shared and mimicked. The imitation can be articulated through parodies, remixes and mashups of the original content. Another essential attribute of memes is their intertextuality they can relate to each

other in complex and surprising ways, completely transforming the original meaning of images. This part of visual culture can mix different areas, such as pop culture, history, and politics, in unexpected ways (Shifman, 2015).

Over the past 25 years, memes have emerged as an essential form of online humour and communication. They have become universal and are used by many generations, from Gen Z to Baby Boomers. What started as a niche phenomenon on the internet is now a dominant part of the social media landscape. Nevertheless, meme culture did not begin with the Internet.

During the Second World War, the graffiti drawing "*Kilroy Was Here*" provoked a memetic effect similar to contemporary memes. The Kilroy drawing spread through a memetic evolution among American soldiers who actively contributed to its mimetic process by replicating it in various locations. The drawing became a well-recognised visual joke, inspiring future versions. Similarly to internet memes, *Kilroy Was Here* was adapted when spread to different social groups. The Kilroy drawing is just one example of a pre-internet meme among many. It is impossible to identify the exact first meme; the memetic attitude is part of human social behaviour (Hamilton, 2024).

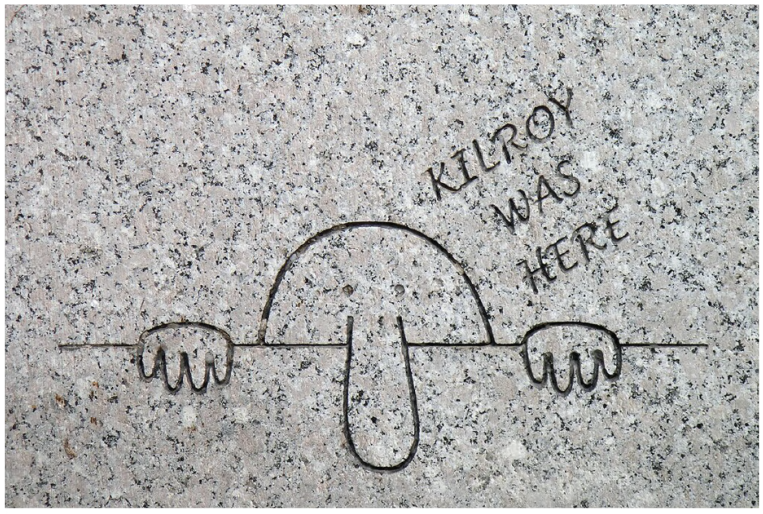


Figure 3. Engraving of Kilroy on the WWII Memorial. (Source: Wikimedia Commons, 2006).

Nowadays, the memetic culture has found fertile ground. In the Internet era, information is shared almost instantly and can reach large audiences in minutes, but what types of content do people like to share? Research shows people prefer spreading content that makes others feel good and, at the same time, reflects their personal self-representation or content that can be used for social purposes. The humour factor of memes usually increases their chances for "virality "due to their capacity to invoke emotional arousal and the surprise factor, two crucial features of viral content. The term viral is used to describe the rapid spread of content on the internet; when a piece of information, such as a text or an image, is rapidly shared and reaches a large audience, it can be named viral content. Nevertheless, not all viral content is a meme, but most memes are viral. Without the viral feature, these images would not be imitated and transformed enough to be considered memes.

The analogy with biology suggests that memes can undergo replication and adaptation and thrive in specific cultural environments. The characteristics that lead people to share specific content differ from the features that lead people to mimic or remix an image (Shifman, 2015).

Contemporary culture rewards participation with sparks of dopamine in the form of likes, which enhances the creation and sharing of memes. Our digital culture can be described as hyper-memetic, in which almost every major event sprouts a stream of commentary in the form of internet memes, sometimes even being the first form of information received. Memes also highlight the blurring line between interpersonal and mass communication (Shifman, 2015, p. 11).

Memes are now an essential part of the digital culture and play an important role in understanding a wide range of contemporary behaviours, ranging from worldwide political protests to funny dance movements (Shifman, 2015, p. 132). Even though memes are considered trivial mundane artefacts, they can be regarded as post-modern folklore for being constructed by cultural elements (Shifman, 2015, p. 17).

There are three main attributes ascribed to memes relevant to contemporary culture: the gradual propagation from individuals to society, the reproduction of the memes via copying or imitation, and, lastly, the distribution through selection and competition.

According to Hawkins's analysis, successfully spreading memes incorporates three properties: longevity, fecundity and copy fidelity. The internet facilitates fecundity, which is the number of copies made in a time unit. The web can also facilitate fidelity since it is easier to maintain the closeness of information in comparison to other types of media (Shifman, 2015, p. 20).

Sharing is a crucial activity on the Internet and often occurs through social media platforms. The act of sharing is central in contemporary culture; it converges two applications: when the users share their ideas and feelings through an image or text and when other users share the content of others as a way to show agreement. Hence, spreading memes is now an essential part of the digital culture. In the process of sharing, memes can be repacked and have their original form changed. The user then adapts the original meme into a new form.

The two main repacking mechanisms are mimicry and remix. The first mechanism, mimicry, consists of the recreating of an image. The second way is through remixing. The act of remixing involves the use of technology to manipulate the images. An example of manipulation is the use of the software Photoshop. Both forms of repacking are common in the digital environment (Shifman, 2015).

Memes vary in their degree of fitness, which is their ability to survive and thrive in a specific cultural environment. They can break their initial cultural boundaries and appear in completely different contexts compared to the original cultural context in which they were born. Memes have a multidimensional presence both in the digital spheres and non-digital environments. This hyper-memetic nature of our contemporary culture can be tested in the transformation and presence of old memes into new memes in totally different contexts (Shifman, 2015).

3.1 - Memes in Global Politics

To unveil the significance and impact of internet memes across various cultural domains, one could reference the documentary "*How the Far Right Use Memes to Recruit | Decade of Hate*" by VICE News (2021). The documentary reveals that silly visual jokes such as memes are powerful tools that strongly influence the political context of the United States. Different political spectrums use memes as an effective tool of persuasion. The humorous aspect of memes can be used as a disguise to spread hateful and discriminatory content, which is often protected by the ambiguity effect produced by jokes (VICE News, 2021).

Figure 4. Pepe the Frog original character
Source: Artnet News (Matt Furie, 2005).

The character "*Pepe the Frog*" exemplifies the evolution of meaning within the digital world. Initially created by cartoonist Matt Furie, the character had no political connotations until internet users decided to alter his expression by modifying his smile into a dismissive and self-satisfied smirk. This new facial expression became integral to the message intended by meme creators. Consequently, the character featured in countless meme images spanning various subjects and themes, eventually evolving into a symbol of right-wing internet propaganda. The addition

of the smirk gave the character an aura of pleasure derived from deviating from societal norms. Originally circulating among juvenile audiences on a handful of social media platforms, the meme gained widespread attention when former U.S. President Donald Trump shared a version of Pepe merged with himself on his social media account. These silly images of the internet moved from jokes between teenagers to global politics. The Pepe meme is now inhabiting more than just the digital world; it is affecting multiple aspects of society. Due to its spread among right-wing extremists, the meme was later classified as a hate symbol and added to the Anti-Defamation League's database (VICE News, 2021).

Figure 5. Donald Trump as Pepe the Frog (Source: Know Your Meme, 2015).

3.2 - Meme Culture

The meme culture is saturated with personal branding, and people use memes as a way to express their individuality and connectivity simultaneously. There is a common understanding that contemporary society is based on an "attention economy" in which the most important factor is not the information but the attention people pay to it. In this logic, users might mimic successful content using this structure to have more chances of views just as the original image. This mimetic activity is performed by a wide range of users and for multiple sorts of content (Shifman, 2015, p. 31).

In the digital world, all users are free spirits who can take their unique path to the hall of digital fame. These paths to fame are, in practice, similar to the tracks used in the creation of memes. These socially recognised types of communication share structures and stylistic forms that help the information be successful (Shifman, 2015, p. 77).

Memes are more than just viral images that are imitated; they may help us decipher cultural and social processes, but they do not arbitrarily flow on the internet. Behind each meme, there is an individual, a person who participated in the process of creating and diffusing the meme (Shifman, 2015, pp. 132-133). In this social logic of participation, individuals help shape social networks, simultaneously expressing their individuality and affiliation to larger communities. In this way, memes are places where the historical cultural production encounters the new affordances of the Internet (Shifman, 2015, p 32).

Memes are transported via various vehicles, such as images, gestures, and videos, passed from person to person. Using an analogy with biology, the visual characteristics of memes are like phenotypes, the tangible manifestation of cultural pieces of information.

Memes are not single entities that propagate well but individual units that are part of a bigger content group with common features. People tend to imitate three dimensions of cultural items: content, form, and stance. The first dimension is mostly related to the ideas and ideologies conveyed by it. The second dimension is the form, which regards the physical form of the message, what is perceived through our senses. The third feature is the stance, that is, the attitude of the item, the information memes convey about their communication; those express the linguistic codes and the relation to the text. According to this definition, it is important to discuss memes not as single formulas that are widely spread but as groups of content. Thus, Internet memes can be defined as a group of digital items sharing common characteristics related to form, content, and attitude that are created with the awareness of each other and shared and imitated through the Internet by different users. (Shifman, 2015, pp. 35- 37).

When Internet users imitate or produce a commentary on specific content, they create their own version, inevitably showing their interpretations and embracing or rejecting certain dimensions of the original meme. (Shifman, 2015, p. 38). The process of imitation includes copying and changing aspects of the original content. These changes may also switch the communication mechanisms. It may be that the most powerful memes use this process of recreation, mechanisms of ironic communication, which is, using visual codes and texts to imply the opposite of what the original item meant. (Shifman, 2015, p. 41).

3.3 - The Birth of a Meme

Since memes can be understood as a group of interconnected content units that share common features such as form, content and attitude, we may also include numerous variations and combinations of the same memes (Shifman, 2015). The same idea can be expressed in different forms, such as videos, texts, images and animations. As an

example, the meme "*Neil deGrasse Tyson Reaction*" demonstrates that there are multiple possibilities of formats to convey the same idea. The meme originated from the facial and gestural expression of the astrophysicist Neil deGrasse Tyson during an interview for the forum Big Think. In the original context, Neil was talking about the impressive achievements of Sir Isaac Newton. His reaction with funny facial expressions and his hands inspired internet users to utilise his expressions and appropriate them as their reactions.

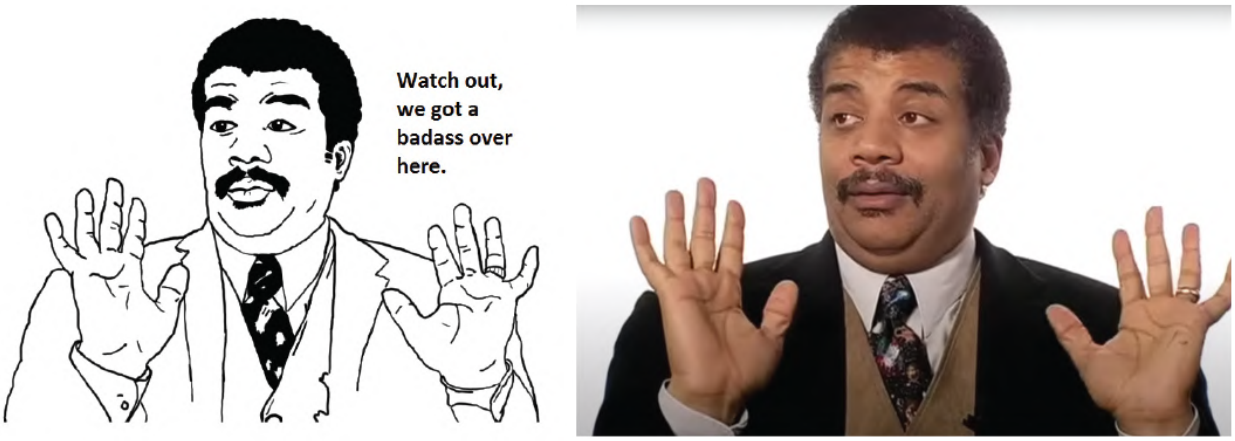


Figure 6. Neil de Grasse Tyson reaction meme (Source: Know Your Meme, 2012).

According to the website KnowYour Meme (2011), The creation of the meme can be traced back to the year 2011, when a user called "*Hippopoptimus_Prime*" replied to the thread with the comic version of Neil's face, as a response to a question in an Internet forum. The spread of the meme Neil deGrasse reaction mainly occurred in the context of poking fun at self-gratuitous comments, status updates or photos on social media websites. De Grasse Tyson's original reaction was later transformed and used to make fun of people who expressed their achievements on social media. In some variations of the meme, captions were added, such as "Watch out guys, we're dealing with a badass over here." and "Watch out, we got a badass over here"(Know your meme, 2018). The meme evolved into multiple variations, reaching audiences outside the English-speaking network. In January 2012, the meme started circulating

on the Spanish-speaking network, and a new caption, "Ay si ay si," was created, carrying a sarcastic tone. The Neil de Grasse reaction meme exemplifies the fluidity memes can take. Multiple variations of the meme and new memes reference the original one (Don & Brad, 2011).

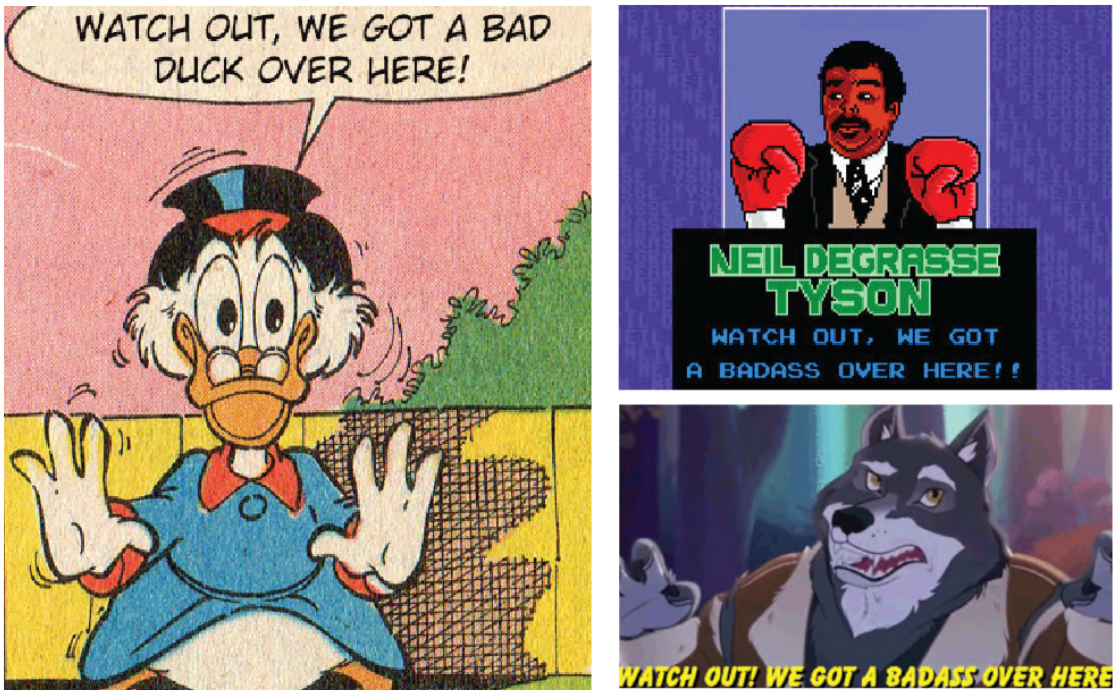


Figure 7. Neil de Grasse Tyson meme variations (Source: Know Your Meme, 2012).

3.4 - Genres of Memes

There are numerous genres of memes and different ways to classify visual humour. What constitutes a genre of meme is the similarity in form and attitude of the meme. Attitude can be defined as the similarity between concepts and socially recognised types of communicative action. Memes can share more than stylistic features; they can also share themes and specific audiences.

The genres of memes are vital in the construction of social group identities. Some genres of memes can be easily understood by a large audience, but others require additional knowledge of the communication codes of a specific subculture. In certain

genres, producers and consumers of memes must comprehensively understand the formats these groups use to communicate (Shifman, 2015, p. 77).

The Neil de Grasse reaction meme is part of a specific genre called "Rage Comics". This genre was created in 2008 and started as webcomics of characters made with basic drawing tools such as MS Paint. Those memes usually convey facial expressions exaggeratedly and simplistically (Don & Brad, 2011).

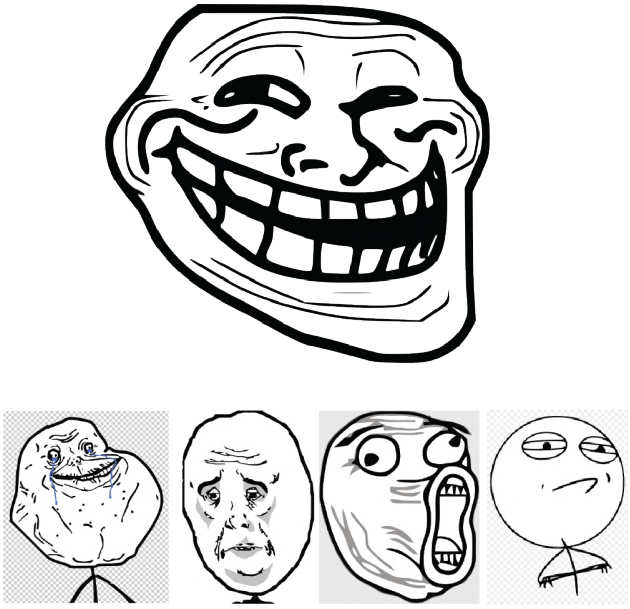


Figure 8. Examples of Rage Comics (Source: Developed by the author).

The success of Rage Comics can be attributed both to the simplicity and application of the content. Rage Comics depict expressive characters associated with typical behaviours and graphically translate emotions into simplistic amateur-looking comics. Another factor is the popularisation of "Rage Makers" websites, in which users could create rage comics easily by reappropriating ready-made characters, facilitating the creation and reproduction of memes. However, digital literacy is not enough for meme creation; as previously mentioned, the knowledge of the codes and norms of certain digital subcultures is an essential part of the process (Shifman, 2015, p. 88).

Despite the complexity of meme genres, it is possible to divide memes into three main groups: (1) Genre based on real-life moments (photo fads, flash mobs) and have less connection to the digital space and manipulation of images. (2) Genre based on the explicit manipulation of images, including manipulated images, videos and mass media content, regularly reappropriating popular culture and remixing into new forms. The transformation approach of mass media reveals criticism or enchantment towards contemporary cultural elements. (3) Genre that is purely digital and evolved directly from the digital universe, this group of memes usually emerged from the internet culture and requires a more profound knowledge of the specific elements used by subcultures to be understood. These memes often emerge in specific internet subcultures and might spread to broader audiences outside the initial group (Shifman, 2015, pp 90-91).

One of the main features of memes is the capacity to spread to large audiences. Larger audiences tend to increase the chance of mimicking and copying a meme. According to the dynamics of memes, communication is mainly a process of imparting information, expecting the effective transmission of the messages to the largest audience possible. In this procedure, communication is not an easily traceable unit but an ongoing process in which identities are continuously constructed (Shifman, 2015, p. 49).

3.5 - The Secrets of Viral Images

According to Shifman (2015), the researchers Jonah Berger and Katharine Milkman developed a search that revealed a few characteristics that increase the chance of content becoming viral, one of which is the preference for positive and humorous content. People prefer sharing content that is entertaining, motivating, and upbeat. This tendency reinforces the centrality of humorous content in viral processes. The category “jokes” is the largest forwarded type of content. In a similar study, Golan

and Zaidner discovered that more than 90% of advertisements produced by viral advertising companies include some sort of humorous element.

The element of surprise often present in humorous content is also a reason for the success. Surprises invoke emotional arousal, which is a key factor in making content go viral. (Shifman, 2015, p. 54) Another vital part of the viral diffusion of images is the packaging of the content. The researchers, Berger and Milkman, found that clear and simple stories tend to be more shared than complex ones. The simplicity of the content helps people understand the message quickly and then increases the chances of the content being shared. People assume that since they understood the message quickly, others will be able to do the same.

Another important factor related to the success of content is Prestige. The "prestige" element is related to people's knowledge about the creator of the content. The more famous the author is, the more likely people are to forward the content. People have more confidence in sharing stories of people with "prestige" than unknown authors. (Shifman, 2015, p. 55)

Another feature that might influence the "virality" of an image is its positioning. The place where the image will be displayed plays an essential role in the meme's potential. Positioning the image on the right social media, to the right audience, and on pages with more followers increases the chances of success (Shifman, 2015, p. 56).

Nowadays, content positioning is completely affected by the algorithms that rule the internet. Most social media and search platforms use algorithms to organise and distribute content.

Algorithms are a set of instructions to execute specific tasks. The internet algorithms are now vital for the whole functioning of the internet. With large quantities of data produced daily, it is humanly impossible to control the whole flux of information created and disseminated. Popular social media platforms like Instagram, TikTok, and YouTube rely heavily on algorithms to suggest and promote content to users

(Stockton, 2023). The algorithm can be recognised also as a form of artificial intelligence. The machine makes a decision based on the data it receives, and then it decides what to show to users, also based on the previous browsing history of each user. Every action taken in the internet world is now tracked and used to create a profile of tastes. Despite offering their services for free, social media platforms profit immensely from the information acquired from the users.

The internet is now a gigantic database divided into millions of segments, each with specific preferences and interests. The web is now the perfect tool for marketing and selling. Internet users are now simultaneously consumers and the products sold by the big Techs to other companies.

The exact mechanism also works for humorous content. The more an image is shared, the higher the chances the algorithm will promote it. The interaction (likes, comments and views) that an image gets is a sign of importance to the algorithm. The machine assumes the image is interesting and should be shown to even more people. This logic strongly interferes with the viral processes of memes and content consumption. The data collected about the user will directly influence what the social media platform displays in the user "feed". The user will be fed content based on previous interactions (Stockton, 2023).

3.6 - Anatomy of Memes

To better understand memes, it is necessary to analyse influential aspects of the anatomy of memes. The first feature of memetic videos is the expressive focus on ordinary people. There is a large quantity of successful memes that depict ordinary people. This could be explained by the fact that common people tend to relate more to ordinary people in everyday situations. The videos created by ordinary people usually receive more genuine interaction than videos of companies with celebrities. Another characteristic contributing to a higher number of memetic videos made by

ordinary people is the simple production, which is easier to imitate. Considering that social media platforms act as communities, people have more tendency to interact with the content of ordinary people. There is a higher connection to the ordinary person in comparison to the celebrities (Shifman, 2015, p. 62)

Another vital feature of viral memes is their simplicity. A vast majority of viral memes employ a simple construction of their topics; they convey simple ideas with simple visual production. Most videos feature one or two performers, with simple backgrounds and little or no editing, and many are filmed in a single shot. It is easier to mimic images and videos with little or no production. Reproducing content with expensive scenarios is out of reach for most internet users. Simple videos allow common users to emulate them in their own way, with low resources and limited digital literacy.

Repetitiveness is a complement to simplicity and is often used in meme content. The repetition of words and phrases helps the user absorb the content and encourages active user involvement. Repetition also enhances the memorisation of the content. (Shifman, 2015, pp. 64-66).

Simplicity and repetitiveness can also be applied to the written aspects of memes. Memes are made not only of visual codes but also of captions and words. It can be assumed that “bad” texts make “good” memes. Bad texts can create humour and facilitate the replication. The logic of the participatory contemporary culture contributes to the active involvement of users with the content. An incomplete text serves as bait for dialogue between users, who can complement the phrase. In this way, the unpolished amateur texts are invitations for users to fill the gaps and play with the puzzles generated by the memes (Shifman, 2015, p. 67).

Internet memes are a vital part of the Internet culture and are essential agents of the contemporary visual culture; they are products of the current technological and social context. Memes convey all sorts of information, from cute, funny images of animals to political symbols; they connect subcultures and promote globalisation,

creating a new form of communication which is fast, dynamic, visual and easily consumed. Can the new AI technologies interfere with this aspect of visual culture? To better explore this question it is important to understand how AI operates.

4.0 - AI and Computer Visuality

In only a century and a half, humans managed to move from big steam engines to magical pocket tools that can display, record and share images with anyone on the planet. These pocket devices, also known as smartphones, can perform various tasks. In a little more than 60 years, they evolved from gigantic machines used by the military and scientists to perform calculations to tiny devices present in the pockets of most contemporary human beings.

Humans are living in an exponential era of technological development. Many technological jumps that emerged in the last century would have been unpredictable to people living only a few decades before. In the early 90s, people would never have imagined that tools such as smartphones would be used to perform tasks such as finding a job, finding a sexual partner, booking flights, paying the bills, recording and broadcasting images to thousands of people in different parts of the planet among many others (Tyson, 2023).

In addition to the smartphone revolution, humans managed to develop artificial intelligence; a new technology that has the potential to change the history of our species profoundly.

Alan Turing was one of the first explorers of the idea of computer intelligence, he developed the concept that computers work similarly to human brains. Because of the similarity between computers and the human body, it was possible to develop some sort of machine intelligence emulating the human brain. However, the term artificial intelligence was coined later by John McCarthy (Copeland, 2016).

Computers were first used as efficient calculation tools, but with further development, they would soon be able to perform tasks that required other types of intelligence. This concept of intelligent machines with similar capacities to humans inspired people to create various movies and novels about this topic.

Many of these artistic representations depict futuristic machines with human intelligence, exploring our collective fears and aspirations regarding a shared future with intelligent machines. In some depictions, the machines are portrayed as benevolent, while in others, humans are depicted as enslaved beings. The famous movie "*The Matrix*" is a prime example of a dystopian future in which artificial intelligence reigns, relegating humans to mere sources of energy and subjecting them to living within a simulated reality (Grebey, 2021). Similarly, the movie "*Ex Machina*" offers a more subtle portrayal of artificial intelligence's potential dominance. In the film, an intelligent machine undergoes testing to ascertain whether it possesses full consciousness of its own existence (Universal Pictures Portugal, 2015).

4.1 - The New AI Tools

For a long time, the idea of computers being smarter than humans was seen as a distant future, but now, with the release of technologies such as ChatGPT, this distant reality seems closer than ever. Artificial Intelligence is now a cause of great alarm in many industries. Giant technological companies such as Microsoft, Meta, and Alphabet are spending billions of dollars to develop and implement AI tools (Leswing, 2023).

What has caused alarm about these new AI technologies? The impressive communication skills of ChatGPT and its capacity to simulate human-like answers. Only two months after its release, ChatGPT reached the mark of 100 million users, making it the most popular app in history (Hu, 2023). People were impressed by the fact they could have a text conversation with something that was not human. ChatGPT could create responses similar to humans and even provide creative outputs such as poems, song lyrics and jokes, according to the user’s request. ChatGPT was developed by OpenAI, a company financed by Microsoft (Hu, 2023).

The impressive texts produced by ChatGPT brought the question, “How close are we to developing a machine as intelligent as us?”.

The capacity of ChatGPT is undeniable. In many ways, it is more capable than humans, but in other tasks, it can fail miserably. It can produce a text or an explanation about almost any topic using sophisticated linguistic skills in only a few seconds. ChatGPT is fed with billions of written information from the internet and is also capable of writing in many human languages.

“Some have suggested that to program a machine to be artificially intelligent, you would first have to understand human intelligence”(Sautoy, 2020, p. 1).

The ability to surpass humans in certain activities is not new; computers were developed to assist humans in doing complex calculations and managing large amounts of data. For some time, humans believed that creativity was exclusive to humans and that machines would not be able to reach our capacity, but the recent achievements of AI put that into question (Sautoy, 2020, p. 2).

4.2 - The Creative Power

Humans have an extraordinary ability to imagine and innovate; they can create works of art capable of causing deep emotions and can solve problems by making machines and devices. Nowadays, humans cure diseases by combining chemical elements. “These are the outpourings of what I call the human code. This code is what people believe depends on being human because it is a reflection of what it means to be human. Mozart’s requiem allows us to contemplate our mortality. Witnessing a performance of *Othello* gives us the chance to navigate our emotional landscape of love and jealousy. A Rembrandt portrait captures much more than just the sitter's

appearance. How can a machine ever hope to replace or even compete with Mozart, Shakespeare or Rembrandt?" (Sautoy, 2020, p. 2).

Creativity can be defined as the ability to create something new or surprising that is valuable. However, creating something new is not a challenging task. The hardest part of creativity is proposing something new that surprises and generates value.

The element of surprise is what makes us feel a rush when we see an act of creativity that belongs either to us or to someone else. Creating something new can also be defined as originality, an original act that inspires others. For a long time, this creativity has been thought to be exclusively human, a product of millions of years of evolution that shaped our brains to acquire these skills (Sautoy, 2020, pp. 3-4). Can a machine learn how to be creative like us? Could it extend our ability and help us see new opportunities outside our scope? The impulse for creativity is part of what distinguishes human beings from other animals (Sautoy, 2020, pp. 5-6).

To be able to reach or surpass our capacity for creativity, artificial intelligence must create a work of art that is repeatable and is not the result of random error; also, human programmers should not be able to explain how the computer produced the output. The machine should come up with something new and surprising that has a contribution. Nevertheless, the value created should be more than the expression of the coder's creativity or the data which fed the AI, according to Sautoy (2020) is the Ada Lovelace test, a version of the famous Turing Test proposed by Alan Turing to measure the intelligence of a computer.

It is essential to distinguish the different types of creativity. Exploratory creativity is exploring the edges of something that already exists by extending the limits of what is possible while respecting the rules. This type of creativity accounts for 97% of human creativity and is where computers usually excel. The act of pushing a set of rules to the limit fits perfectly the mechanism of computers (Sautoy, 2020, p. 10).

The second type of creativity involves combination and is one of the most preferred by artists because it combines two completely different constructs that can generate something new. The third type of creativity is transformational creativity. The transformational type is the most mysterious and elusive, and it occurs in those rare moments when something completely unexpected happens, like the transformation of the state from liquid to gas. Art movements such as Cubism could be classified as a transformational creative type because of the complete change in the way of depicting figures. These transformational moments alter the rules of the game under which previous work had been working. Following Sautoy's (2020) reasoning, there is a way to instigate transformational creativity:

"You start dropping constraints and see what emerges. The art, the creative act, is to choose what to drop or what fresh constraint to introduce such that you end up with a new thing or value" (Sautoy, 2020, p. 11).

To perform a genuinely creative act, it is necessary to step outside the system and come up with a new way of seeing things. Can a computer do that? The concept of reality is attached to the current historical context of time. To define something as new, we should first consider the historical context in which the novelty emerged. Creativity is a relative activity attached to historical contexts. Sautoy (2020) says, "We are creative within our culture and frame of reference."

Can a computer initiate this kind of phase change?" Can a machine "think outside the box" and bring us to a new state in art, music, and science? The challenge for artificial intelligence is that the "how to act data" is based on the data the computer is given. Would this condemn the machine to produce more of the same? (Sautoy, 2020, pp. 12-13).

4.3 - Artificial Creativity

The creative skills of artificial intelligence have notoriously increased within the past few decades. Machines already managed to defeat humans in strategic games that require certain levels of creativity. AlphaZero, an artificial program developed by Google, managed to defeat Stockfish, another artificial intelligence program that had previously defeated the best human chess player of the time. Contrary to the expectation, AlphaZero had no preprogrammed moves; it developed its playing style based on the data it received. The program's creators instructed AlphaZero with the rules of chess and the command to maximise its proportions of wins to losses. The strategy used by AlphaZero was considered unorthodox, and they performed risky plays that humans normally would avoid; even so, AlphaZero became the current best chess player, and so far, no human has managed to defeat it (Kissinger et al., 2021).

The artificial intelligence advancements are not limited to games but also to science. In 2020, researchers from MIT announced the discovery of a new antibiotic used to kill strains of bacteria that so far had been resistant to human antibiotics. The novelty of this discovery was that the team of researchers had the help of artificial intelligence. The computer identified strains of bacteria that had eluded humans. To perform this discovery, the AI was first fed with data about thousands of molecules and was able to identify the bacteria's weaknesses. The new antibiotic was called "*Halicin*" (Kissinger et al., 2021).

The advancements and capacity of AI are undeniable. It can hugely impact all different aspects of human society. AlphaZero's victory, the "*Halicin*" discovery, ChatGPT, and the image-generating tools reveal AI's large potential to unveil previously imperceptible aspects of reality.

To assist AI machines in performing tasks that require human intelligence, they utilise methods such as "deep neural networks," which emulate the human brain

and allow the machine to produce images and texts that appear to be made by humans (Kissinger et al., 2021).

4.4 - Deep Neural Networks

Deep neural networks are machine learning programs that mimic how biological neurons work together. These neural networks consist of layers of nodes with an input layer, other hidden layers, and an output layer. According to the explanation provided by (IBM, n.d), "Each node connects to others and has its own associated weight and threshold. If the output of any individual node is above the specified threshold value, that node is activated, sending data to the next layer of the network. Otherwise, no data is passed along to the next network layer." This structure of layers allows the machines to perform very complex tasks. Despite the impressive outputs produced by AI, there are some aspects in which machines seem to perform poorly.

"In some tasks, AI achieves human or superhuman levels of performance; in others (or sometimes the same tasks), it makes errors even a child would avoid or produces nonsensical results. AI's mysteries may not yield a single answer or proceed straightforwardly in one direction, but they should prompt us to ask questions. What will AI's impact be on our culture, our concept of humanity and in the end our history? (Kissinger et al., 2021, p. 18)"

In the field of visual production, AI is now being used by many artists, designers, and illustrators to enhance their creative processes. Recent tools such as DALL·E 2, Midjourney and Stable Diffusion are capable of creating impressive images similar to humans but extremely fast, taking less than a minute; this would be inconceivable to humans.

Nevertheless, the images generated by these tools quite often require further editing and corrections. In 2022, Jason M. Allen won the Colorado State Fair annual art com-

petition. He created his artwork using Midjourney, an AI tool for images (Roose, 2022).

4.5 - AI in the Fine Arts

At a time of growing concerns about the impact of artificial intelligence on society, many artists are embracing the potential of AI and using it as creative tools. One of these artists is Refik Anadol, a Turkish-born and Los Angeles-based creator who works in new media. Over the past decade, Anadol has utilised data as his palette and algorithms as his artistic instruments. While his works may not alleviate our apprehensions about technology, they are often captivating with their mesmerising qualities and capacity to provoke thought. Securing a commission from a prominent modern art institution is a significant achievement for any artist, and achieving this milestone before the age of 40 solidified Anadol's reputation as a leading figure in generative AI art. "*Unsupervised*," a piece from his "*Machine Hallucinations*" series, takes all 138,000 artworks from MOMA's collection and subjects them to meticulously crafted algorithms, resulting in ever-evolving images projected onto a towering 24-foot screen in the museum's lobby (Whittington, 2023).

4.6 - Algorithms

To better comprehend how artificial intelligence works, it is vital to explore the concept of algorithms. According to the Britannica Encyclopaedia (2024):

"Algorithm, systematic procedure that produces—in a finite number of steps—the answer to a question or the solution of a problem. The name derives from the Latin translation, *Algoritmi de numero Indorum*, of the 9th-century Muslim mathematician al-Khwarizmi's arithmetic treatise "Al-Khwarizmi Concerning the Hindu Art of Reckoning."

The conquest of AI relies deeply on the way algorithms work. Algorithms can learn from their mistakes, and this is one of the secrets that led AlphaZero to win the chess tournaments. The self-learning ability of the algorithm played a vital role in developing computer vision. Algorithms are present in all aspects of contemporary life, as described by Sautoy (2020):

"Our lives are completely run by algorithms. Every time we search for something on the internet, plan a journey with our GPS, choose a movie recommended by Netflix or pick data online, we are being guided by an algorithm. Algorithms are steering us through the digital age, yet few people realise that they predate the computer by thousands of years and go to the heart of what mathematics is all about" (Sautoy, 2020, p. 44).

The integration of algorithms into our technologies has enabled efficient communication with computers. Through this new language, humans can convert numerous tasks into formulas. An algorithm aids in problem-solving by delving into underlying patterns and guiding us toward a solution. While performing specific tasks, the computer is not tasked with thinking; instead, it must follow the instructions encoded in the algorithm.(Sautoy, 2020, p. 47).

4.7 - Computer Visuality

The human brain has a very sophisticated visual sense. It is fine-tuned to find structures and patterns in images. Computer visuality has been one of the biggest challenges computer engineers face for a long time. The task of teaching computers "to see" was highly complex, and computers could hardly identify patterns easily recognisable by humans. On the other hand, humans have a strong capacity to read patterns and interpret them. These important human skills explain our admiration for music and art (Sautoy, 2020, p. 19).

Less than a decade ago, computers struggled to understand what they saw. Although they could analyse millions of pixels from images, programmers found it difficult to write an algorithm that could interpret such a massive amount of data and make sense of it.

Matching the human brain's capacity to interpret visual information is considerably challenging for any computer striving to rival human intelligence. While a digital camera can capture images with detail surpassing the storage capacity of the human brain, it cannot weave millions of pixels into a coherent narrative. The brain's ability to process and integrate data into a cohesive story remains mysterious to us. Simply examining individual pixels in isolation does not provide insights into the broader picture. (Sautoy, 2020, p. 76)

Unlike humans, who can effortlessly discern various pixel arrangements and recognise figures, computers perceive only pixels and the information they contain. Consequently, each computer image comprises a unique pixel arrangement. However, machine learning has significantly enhanced machine vision. For instance, to recognise an image of a cat, AI must undergo training with thousands of cat images. This training data aids the algorithm in learning to differentiate cats from non-cats and grasp the structure of a cat. With each mistake, the algorithm adjusts the questions it poses to itself to complete the task "Is this a cat?" to improve its accuracy in subsequent attempts. As a result, algorithms now possess the capability to adapt themselves to enhance their own performance and efficacy (Sautoy, 2020, pp. 69-70).

The algorithm builds itself and changes its structure by interacting with more data. The computer needs to analyse each pixel in an image and compare it to the neighbouring pixels, but what set of questions should the algorithm ask to identify the right objects? In the past, programmers had to think about the proper set of questions, but now, the algorithm learns the right questions by itself. The best way to improve the algorithm’s ability is to expose it to a large quantity of data; by

interacting with this data, the computer can learn the questions that work the best. All these questions help to build a decision tree. Each question adds another branch to the tree, Sautoy (2020) explains:

"The algorithm starts by choosing a series of arbitrary directions to head out from and some arbitrary depth threshold: for example, head north, if the difference in depth is less than y, go to the led branch of the decision tree, if it is greater, go right and so on. But over time, the algorithm starts refining the angles and the depth thresholds, and it will get better sorting of the pixels. By the time the algorithm has found the best set of questions, the programmers haven't really got a clue how it has come to this conclusion…There are over a million different questions being asked across the tree, each one slightly different. It is difficult to reverse-engineer why the algorithm ultimately settled on this question to ask at this point in the decision tree" (Sautoy, 2020, pp. 73-74).

This computer vision mechanism would be impossible to program manually, and it would require the programmers to come up with over a million different questions. As emphasised by Sautoy (2020), one of the challenges of machine learning is overfitting, when the program is too tailored to the data it has been trained on and is not able to learn something more widely from the data it was fed (Sautoy, 2020, pp. 74-75).

To acquire similar vision capabilities, the machine needs more than a digital camera with the capacity to capture images with detail that exceeds the human eye. The main point is to be able to turn these millions of pixels into a coherent story. The way our brains process visual data and interpret it is still beyond our full comprehension.

The computer method of reading one pixel at a time and trying to connect the dots sometimes is not effective in seeing the overall picture (Sautoy, 2020, p. 76). One of the main weaknesses of these AI image-generating tools is their incapacity to create human hands. Even though there are billions of images of hands, the complexity of human hands seems to be a big challenge for computer vision (Edwards, 2023).

4.7.1 - The Challenge of Depicting Human Hands

Figure 9. Hands holding an apple (Source: Developed by the author).

The challenge of accurately depicting human hands reveals more profound insights into the workings of computer vision. While humans learn to perceive the world by recognising patterns through a comprehensive 3D experience, AI relies on the information available on the internet. The paradox of AI lies in its inability to gain real-world interaction experiences despite having access to millions of hand images. It cannot comprehend how hands vary in appearance based on angles and interactions with objects. The difficulty of computer-generated hand depictions extends beyond its lack of a 3D experiential understanding. This challenge can be attributed to three main factors. Firstly, there is limited availability of relevant data; although countless images are online, not all are utilised for training, and hands exhibit diverse forms, unlike human faces, which follow a consistent structure. Secondly, the intricate nature of hand movements adds complexity. Hands can adopt various shapes based on interactions with objects, presenting numerous possibilities that require extensive data for machine learning. The disparity between the multitude of hand formats and the limited available images poses a challenge for

machines to master this task. Thirdly, the requirement for minimal error tolerance adds to the complexity. Unlike certain details like clothing patterns or human hair, which may not demand high precision, accurately rendering human hands necessitates precise depiction. To overcome these challenges, AI systems must be exposed to ample data to formulate appropriate questions and undergo human fine-tuning. This involves human training to provide feedback on the generated images, allowing the AI to refine its understanding and accuracy in hand depiction.

4.7.2 - The Translation Barrier

One of the primary challenges encountered with generative AI involves translating written words into visual concepts. This challenge stems from the complexity of understanding the nuances and context inherent in language, as words often carry multiple meanings and cultural references. Transforming abstract concepts into tangible visual representations necessitates a profound comprehension of language, context, and cultural subtleties. For example, words that denote abstract emotions may encounter difficulties when translated into visual concepts. Moreover, the inherent ambiguity of language can pose obstacles; a single phrase can evoke various images depending on the AI's interpretation.

Furthermore, ensuring that the generated visuals align with the intended message of the text requires a nuanced understanding of the AI system. While the technology is accessible to a broad audience, only a few experienced users may possess the expertise to generate satisfactory images effectively. Even minor variations in input can yield vastly different visual outputs, underscoring the significance of precise and detailed prompts. Users must strike a balance between specificity and generality in their instructions. The enhancement of image-generative AI tools in translating written words into visual concepts entails advancements in AI technology and user proficiency in utilising these tools.

4.7.3 - The Mona Lisa Experiment

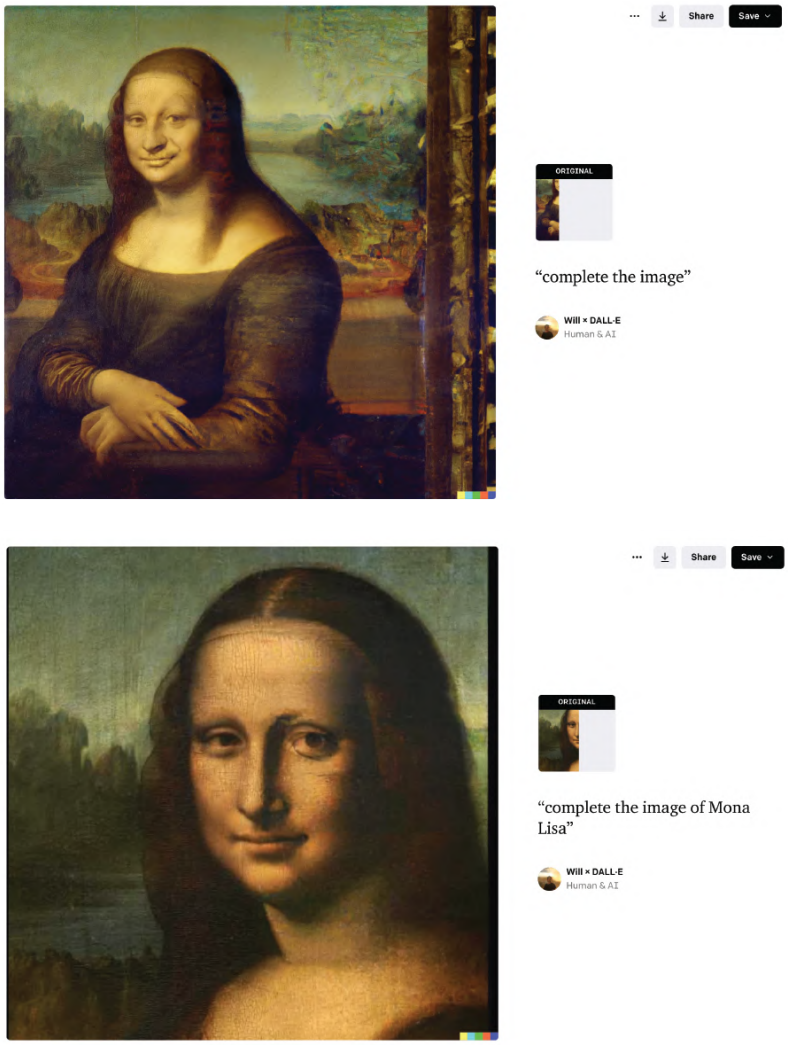


Figure 10. The Mona Lisa Experiment (Source: Developed by the author).

In a recent experiment involving generative AI, the model was tasked with "completing an image fragment" from the iconic painting of the Mona Lisa. Initially, the AI's attempt failed to reconstruct the missing portion accurately. It could not identify that the image provided was the famous painting of Mona Lisa by Leonardo Da Vinci, even though the Mona Lisa painting is extremely famous worldwide. The original Mona Lisa painting is one of the main attractions of the Louvre Museum in Paris, attracting millions of visitors yearly. How can an "artificially intelligent" tool have a poor knowledge of Visual Culture?

The first attempt to perform the task led the computer to create a distorted version of Mona Lisa. The face of Mona Lisa and the background look distorted, generating a different aesthetic experience in comparison to the original image. One of the possible reasons for this outcome was that the computer failed to properly “read” the image. When asked to “complete the image” the AI interpreted all the data present in the image in the form of pixels and tried to complete the pattern of the mathematical data it received. However, the computer’s data-reading skills were not enough to generate a pleasing image, on the contrary, the generated version could be described as a “creepy” version of Mona Lisa.

In the second attempt: the prompt became more specific, instructing the AI to "complete the image of the Mona Lisa." Remarkably, this simple change led to significantly improved results, with the AI almost successfully filling the image with a higher fidelity. Nevertheless, there are some subtle distortions in the image in terms of proportions and colour. The right eye of Mona Lisa looks bigger than the left and also seems to be extra red. Even with the subtle mistakes the results of the second prompt are significantly more satisfying than the first. This experiment underscores the crucial role of prompts in guiding generative AI to produce desired outcomes effectively. In addition, it sheds light on the capabilities of computer visuality. Without human intervention, the computer struggled to recognise one of the most famous paintings, even when provided with 50% of the image. Only after human intervention, the computer was able to recognise that the image was the famous painting of Mona Lisa.

4.7.4 - Google Deep Dream

Figure 11. Moonage Daydream: art created by Deep Dream (Source: The Guardian, 2016).

Even though computer visuality is significantly improving with time, and art created by computers is becoming more satisfying, it is still a product of the human mind. The machine is merely creating art through the lens of humans; even when it manages to be creative and effectively generate artwork, it is primarily done through human commands and prompts and, secondly, through the data it has acquired, which is created by humans. Machines are not inherently creative and expressive of their views. However, Google Dream, a tool developed at Google's Zurich labs in 2014, pushed the boundaries of machine self-expression. The tool utilised artificial neural networks, which emulate human nervous systems, to learn how to perceive images. The engineers aimed to see if the machine could create an image based on its interpretation of what was there rather than following a human prompt or detecting a clear image, such as a bird. The idea was to present an abstract image to the machine, such as a cloud, and prompt the computer to complete the image based on its interpretation. If the computer "perceived" the cloud as resembling a bird, it

would adjust the image to make it more bird-like, reinforcing its recognition of the bird. The results of this experiment yielded the most genuine type of machine art, created through the machine's perception. The images generated by the computer were filled with faces, swirling colours, and motifs (Rayner, 2016).

The current image-generating tools differ from those early experiments with Google Dream because they mainly rely on user guidance. To generate an image, the user must describe it to the computer using written words. These words are referred to as prompts; therefore, the more detailed the prompt, the more accurate the resulting image will be. Thus, to create an image, physical artistic skills are not necessary; the computer only requires the correct use of words. While there are already a few tools that operate on this mechanism, the most popular ones so far are DALL-E 2 from OpenAI, Midjourney, and Stable Diffusion. The primary difference between these tools lies in the data they have learned. This can lead to the creation of entirely different styles, even when the user provides the same prompts.

4.8 - AI in Visual Culture

Generative AI, particularly image-generating tools, have already caused a transformative shift in contemporary visual culture, offering new ways for artistic expression, storytelling, and creative exploration. They can generate a wide range of visual concepts, from photorealistic landscapes to abstract compositions. Generative AI fosters experimentation and innovation, challenging the conventional artistic order and processes. Moreover, these tools facilitate the entrance of people from outside the creative sphere, granting users with different technical skills access to image-generating tools. Beyond the art, generative AI also finds applications in various domains, including advertising, fashion, gaming, and film production. As generative AI continues to evolve, its potential to reshape visual culture remains a

matter of speculation; it might redefine how people perceive, create, and interact with imagery in the digital age.

5.0 - Visual Humour Analysis I: Quantitative

To evaluate the ability of AI to generate visual humour effectively, this study will undertake a data analysis approach. Firstly, it will conduct a quantitative data analysis focusing on visual humour in the form of internet memes. This analysis will be guided by predefined criteria, allowing for a systematic examination of humour-related variables in memes. By assessing these elements, this study aims to comprehensively understand the components and mechanisms involved in creating humorous content. The first methodology employed in this study is content analysis, a technical procedure known for its empirical and objective approach to analysing data. While content analysis provides a systematic method for quantifying representation through defined categories, it cannot support statements regarding the significance or interpreted meaning (Leeuwen & Jewitt, 2001). However, it serves as a valuable initial tool for testing a hypothesis by quantifying specific content categories, in this case, internet memes. The process begins with formulating a hypothesis and then defining the variables. This study aims to rigorously examine and interpret the data through content analysis, shedding light on patterns and trends regarding visual humour.

Unlike semiotic analysis, which examines signs and symbols, content analysis categorises based on specific dimensions to depict the entirety or scope of the chosen topic. To observe and quantify content categories, it is essential to define relevant representation variables first. These variables can include various aspects such as dimension, colour, position, and elements like the number of represented participants or settings. Understanding the concepts of variables and their respective values is crucial to comprehending the analytical process in content analysis. Each variable is logically or conceptually distinct from others identified within a specific research endeavour (Leeuwen & Jewitt, 2001).

5.1 - Defining the Variables

According to the methodology proposed by Leeuwen and Jewitt (2021), the first step to defining the variables is to formulate a hypothesis that will serve as a guide to the research:

1. Most memes do not use AI-generated images.

This hypothesis is one of the base questions of this research. Despite the popularisation and accessibility of image-generating AI tools, these technologies still seem to be absent in the production of visual humour. Somehow, most meme creators still prefer older methods of image creation, like editing software. To explore this topic, 120 memes will be analysed from 6 different content creators. For a better analysis, the memes were chosen in two different languages: Portuguese and English. All memes analysed were created and published between 2023 and 2024. The images will be divided into three categories: "*Yes*", "*No*" and "*Maybe*".

Due to the difficulty in analysing certain images, the category "*Maybe*" was created. The variable of the analysis is subjected to the eyes of the researcher, who will identify patterns, shapes, colours, proportions and perspectives that can be connected to AI tools.

Specific details such as the hands, face and style allow the observer to classify the memes correctly. However, due to the imprecision of the category "*Maybe*", the percentage of this category will be included in the margin of error.

Quantitative Analysis

@9gag

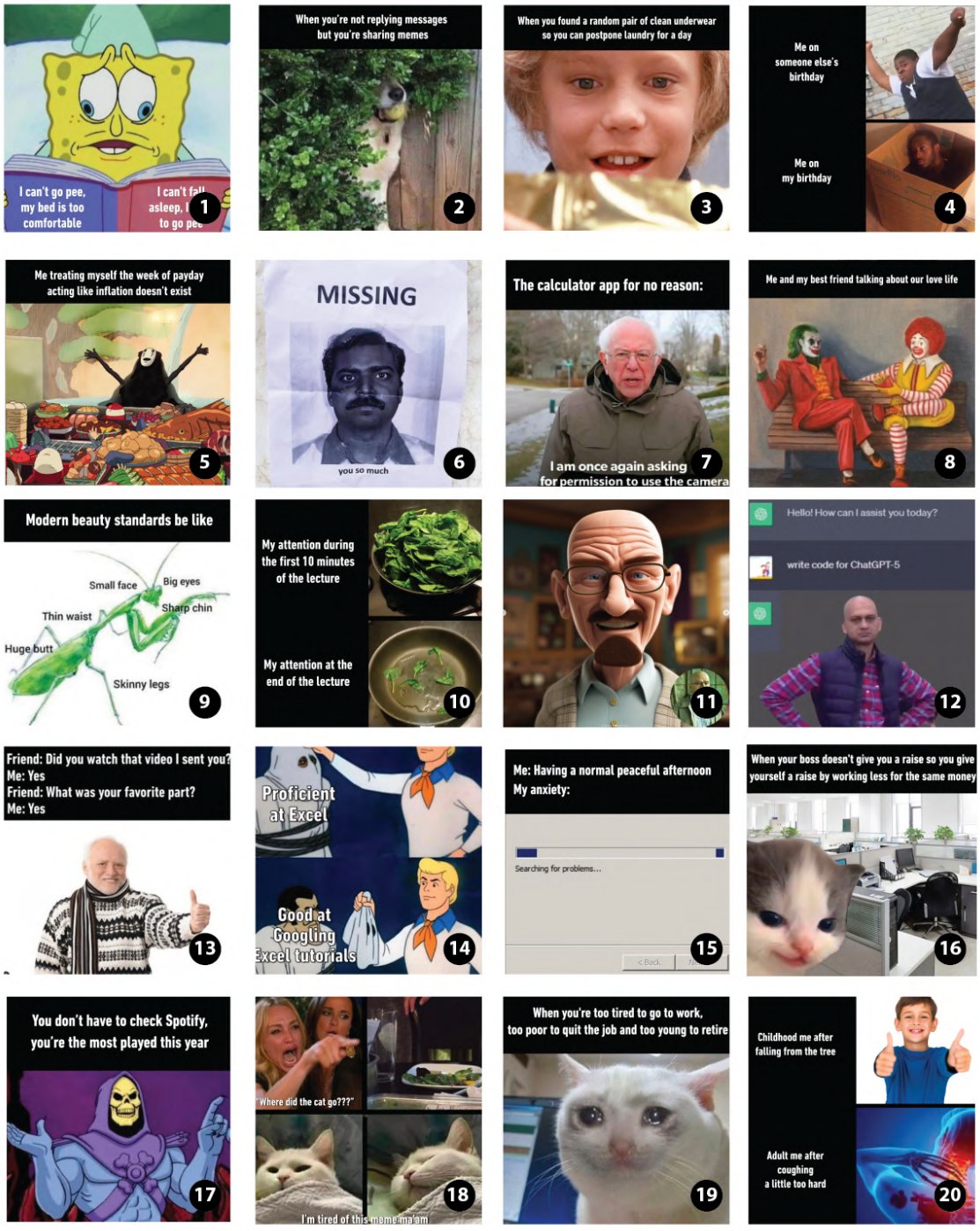


Figure 12. Memes from @9gag *(Source: Instagram @9gag, 2023-2024).*

Quantitative Analysis

@epic_funny_page

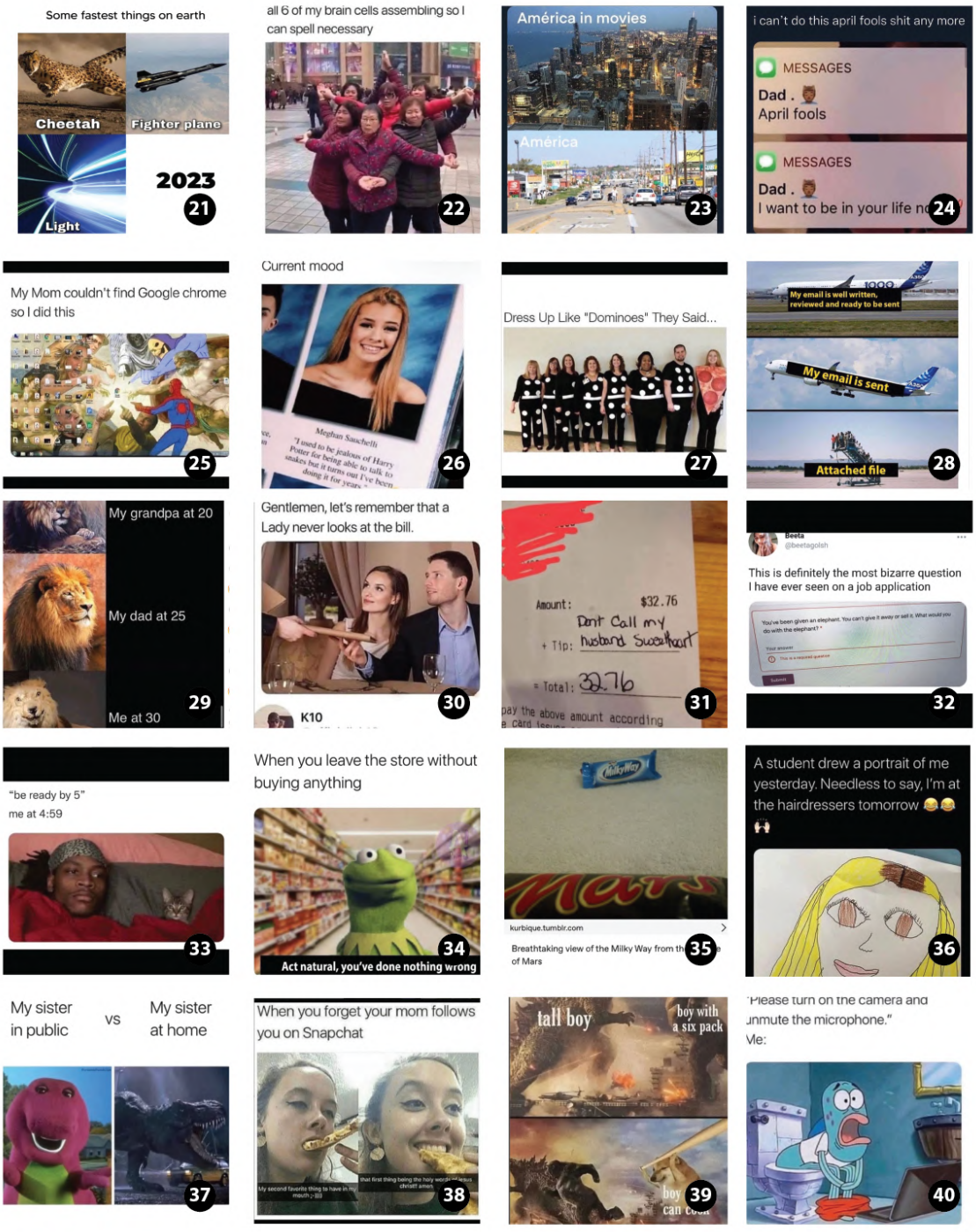


Figure 13. Memes Epic Funny Page
(Source: Instagram @epicfunnypage, 2023-2024).

Quantitative Analysis

@melted_videos

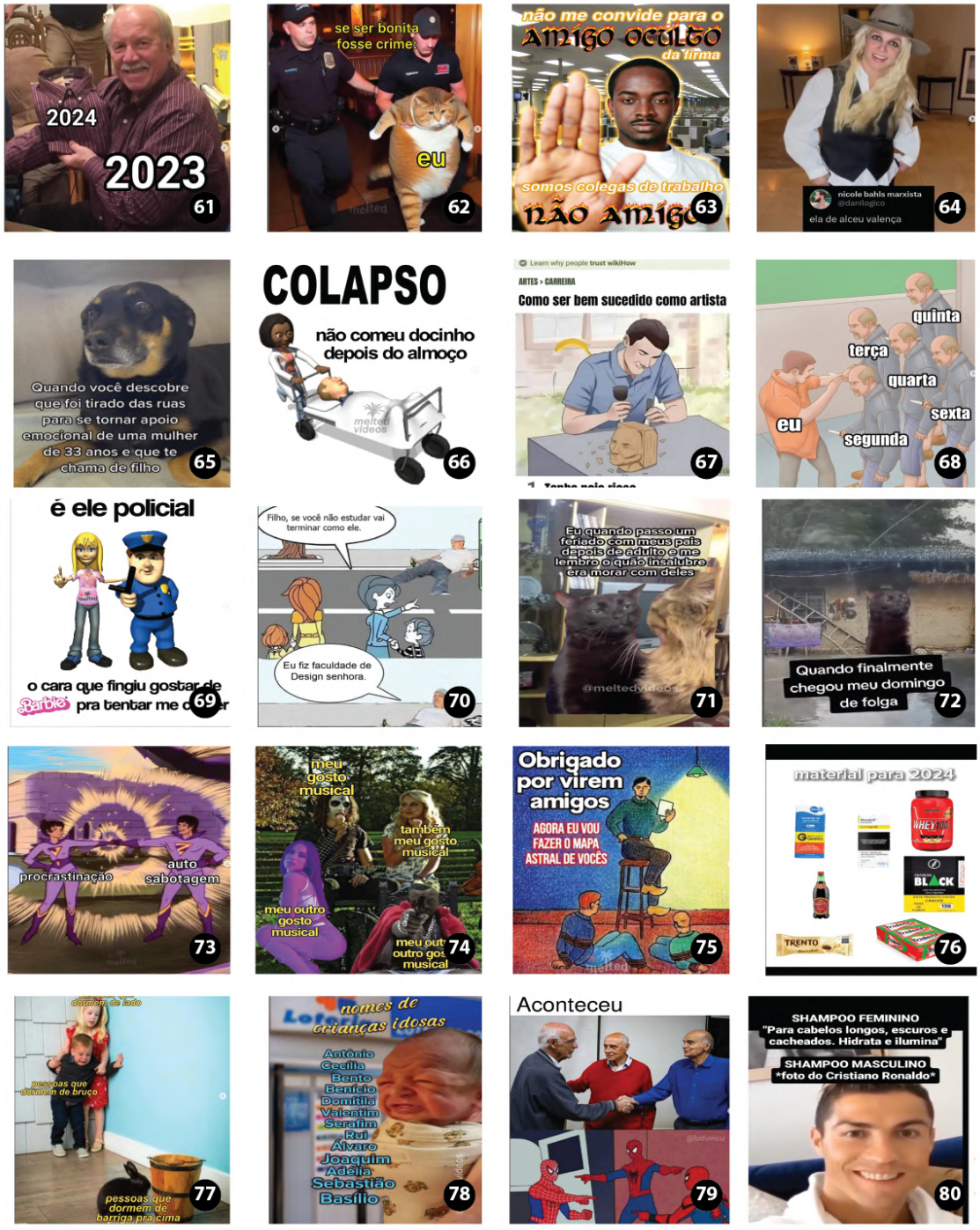


Figure 14. Memes from Melted Videos
(Source: Instagram @meltedvideos, 2023-2024).

Quantitative Analysis

@memes

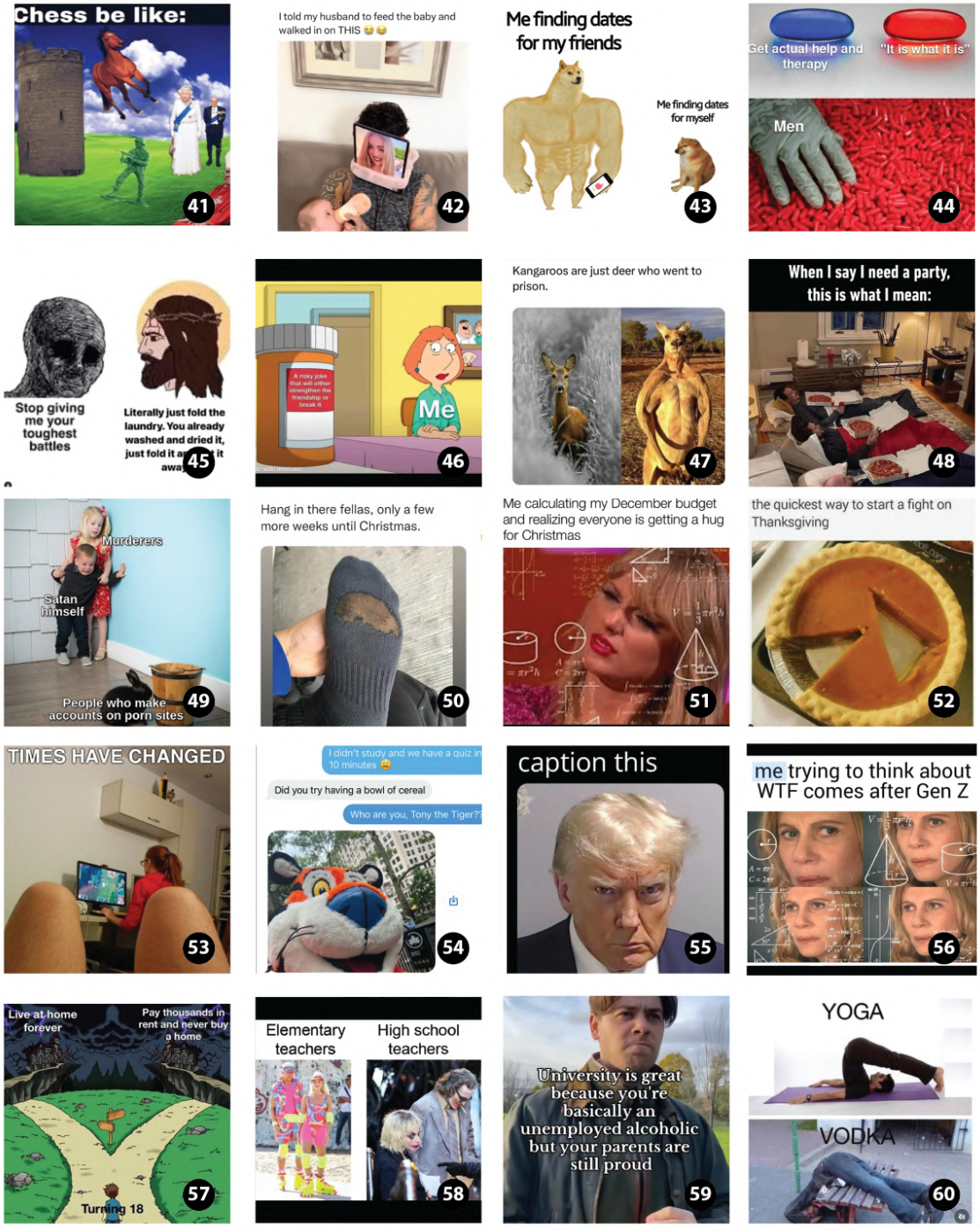


Figure 15. Memes from @memes (Source: Instagram @memes, 2023-2024).

Quantitative Analysis

@festa_da_firma

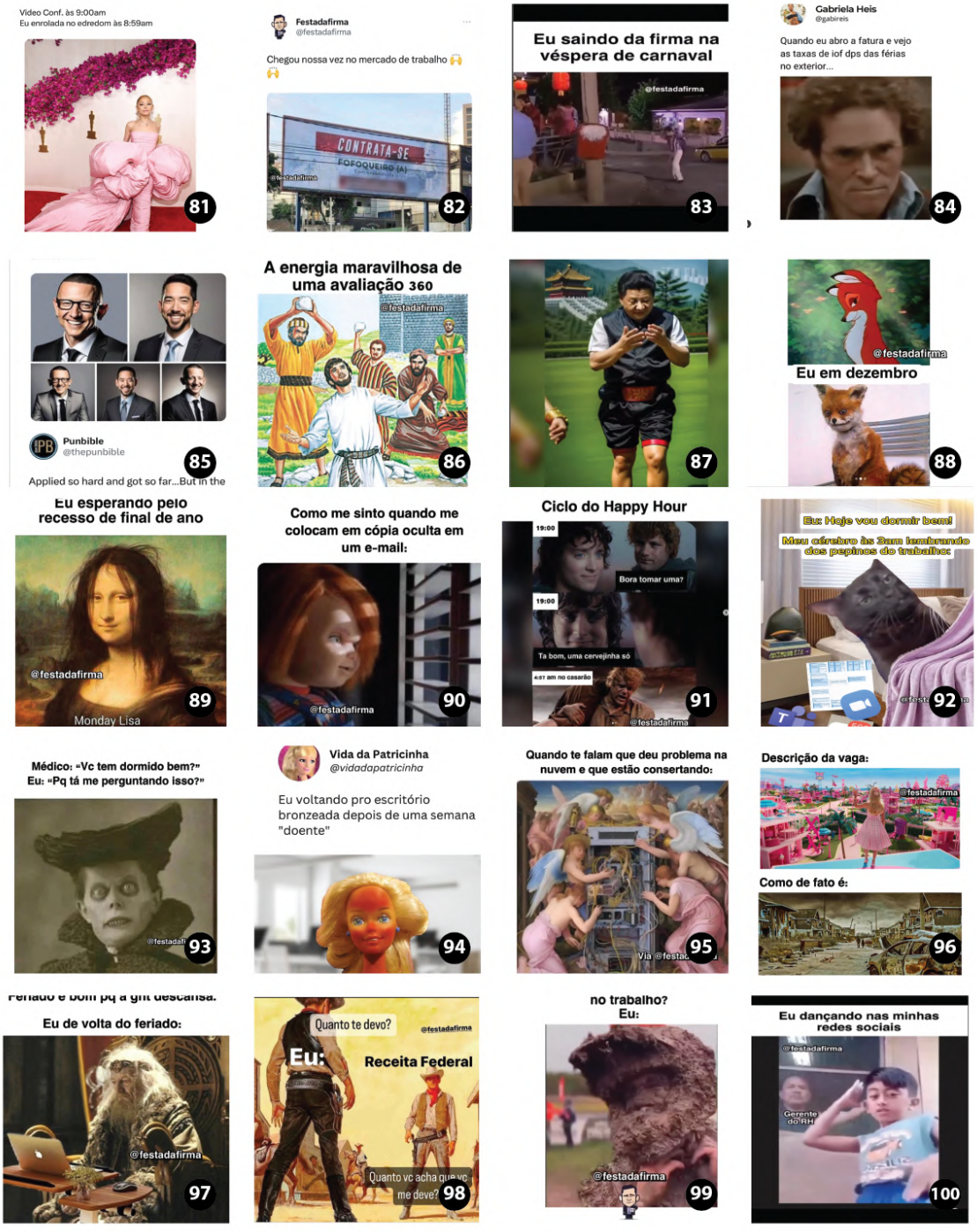


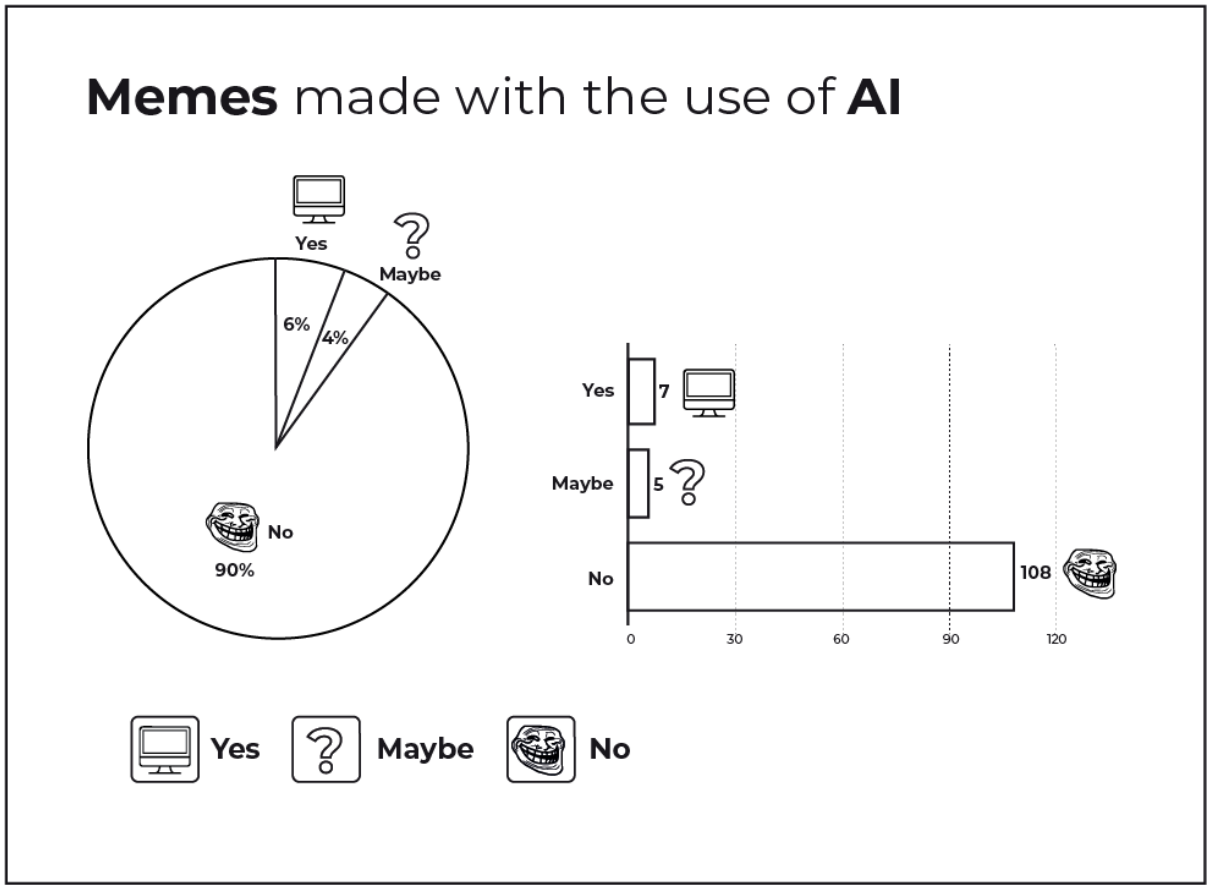
Figure 16. Memes from Festa da Firma (Source: Instagram @festadafirma, 2023-2024).

Quantitative Analysis

@south_america_memes

Figure 17. Memes South America Memes
(Source: Instagram @southamericamemes, 2023-2024).

5.2 - Data Results



Graph 1. Proportion of memes made with the use of AI tools.
(Source: Developed by the author).

According to the data analysis, only 10% of memes were created using image-generating AI tools. The number 10% includes the category "*Maybe*", which accounts for 4% of the images, and 6% of the category "*Yes*", which exhibited signs of AI use. If the category "*Maybe*" is excluded from the final results, the proportion of memes created through AI is significantly lower, only 6%. Even with the uncertainty of the category "*Maybe*", which could be classified as a margin of error, the disparity between humorous images created with AI tools and images created through other traditional methods is vast, with a difference between 90% - 94%. This data confirms the hypothesis that these tools are rarely used to create visual humour in the form of memes.

Despite the rise of AI tools, visual humour still seems less affected by AI. The majority of the analysed memes, 90%, showed no trace of these technologies, even with the benefits brought by these technologies, such as speed and simplicity. Curiously, these factors perfectly match desirable attributes of meme culture, which values speed because of its fast pace of creation and simplicity due to its democratic appeal, which enables any internet user to participate. However, according to the data, it is possible to affirm that most meme creators still prefer the old image creation methods, including image editing software.

What could be the disadvantages of these tools? It was thought that AI would replace humans in many different tasks. Why is a task that englobes poorly edited images and simple written captions an obstacle to AI? To answer these questions, firstly, it is necessary to develop a qualitative analysis to better comprehend the social contexts and history of how certain images became "viral".

6.0 - Visual Humour Analysis II: Qualitative

All visual representations are made in one way or another according to the circumstances of their production and time; this may contribute towards the effect they have (G. Rose, 2007). Unlike quantitative data, qualitative data provides insights into the depth and nuances of human experiences, perceptions and behaviours, exploring the circumstance in which images were made. Through a deep analysis of visual codes and themes, qualitative data analysis uncovers patterns and meanings embedded within the data. This deeper understanding enriches the interpretation of findings and helps generate proper hypotheses and theories. Moreover, qualitative data analysis facilitates the exploration of contexts and cultural codes which quantitative data alone may not capture. The social factor is one of the most important factors in understanding visual images (G. Rose, 2007).

Just as the grammar of a language describes how words combine into sentences and texts, visual grammar describes how visual elements combine into visual statements of greater or lesser complexity (Kress & Leeuwen, 2007).

Like linguistic structures, visual structures point to particular interpretations of experience and forms of social interaction. Certain ideas can, for instance, be said either visually or verbally, others only visually, and others only verbally (Kress & Leeuwen, 2007).

The term grammar refers to a set of rules one has to obey if one is to speak or write correctly, in socially acceptable ways; similarly, to communicate with specific groups, the visual creator must know the visual grammar of that specific community.

Contrary to common understanding, visual communication is not transparent and universally understood. It is culturally specific (Kress & Leeuwen, 2007). Mass media communication, such as memes, frequently combines elements from popular visual language systems, increasing the chances of being understood by mimicking specific patterns and structures of these communities. The quantitative analysis will explore these patterns and structures and the social context of 6 memes, 2 made with AI tools and 4 made with regular image editing techniques. The criteria for the selection were successful memes, images that provoked a large memetic behaviour using traditional methods of creation and memes that used AI tools.

6.1 - Pope in White Puffer Jacket

Figure 18. Pope in White Puffer Jacket
(Source: Know Your Meme, 2023).

The AI-generated image "*Pope in White Puffer Jacket*," also referred to as "*Pope Francis Drip*," portrays Pope Francis wearing a lengthy white puffer jacket. The image was created using image-generating AI tools. This image provoked great amusement, with numerous jokes and memes emerging about the Pope's new look. The realism of the image led many to mistake it for an authentic photograph, prompting discussions about the implications and capabilities of artificial intelligence.

The image depicts Pope Francis from the waist up, confidently walking and wearing a long white puffer jacket. The joke around the image was the confusion it generated around its veracity. Pope Francis is a well-known figure, famous worldwide and often recognised for being a progressive leader of the Catholic Church compared to previous popes. His popularity for being a more open-minded pope might

contribute to the first doubts about the image's veracity. The image of an old religious leader wearing clothing often associated with young men can be amusing, especially if the older man is the most crucial figure of a religion often associated with conservative ideas. In addition, the puffer jacket is not only a visual code related to youth but also a particular type of youth, the so-called “gangsters”, “thugs”, and “badasses”. The corporal posture of the pope also enhances the association with a badass attitude. In the image, Pope Francis walks confidently to the right with his arms and shoulder slightly open, projected forward, A posture often associated with strength. His face is depicted almost in a completely side view; his eyes are slightly closed, and his mouth is slightly open, giving an overall expression of confidence. The framing of the image also plays an important role in transmitting the idea of veracity. Many pictures taken by paparazzi similarly frame the subject as if the subject does not seem to notice that "the picture” is being taken.

In terms of colour and lighting, the image has a high contrast between the subject dressed in white and illuminated by the sun and the dark background. In the dark background, there is a subtle detail of a column, which is a prevalent building structure in the Vatican State; this also contributes to enhancing the “veracity effect” of the image.

The humour generated by this image can be associated with the benign violation theory. The image violates the viewer’s expectation, who usually associates Pope Francis with a “father caring figure”. Contrary to that, the Pope is depicted as a “badass character”, wearing a puffer white jacket. This “benign violation” of expectation generates the humour in the image. Since its first publication, many different memes have been created using this image, often making jokes about the association between a church leader and a gangster figure.

6.2 - Harry Potter AI Trend

Figure 19. Harry Potter AI Trend (Source: Know your meme, 2023).

The "*Harry Potter in AI Trend*" refers to a collection of AI-generated videos depicting Harry Potter characters in diverse scenarios and versions. The original Harry Potter story was situated in a fantasy version of the U.K. and tells the story of a magical world secretly functioning in British society. Harry Potter is one of the most famous fantasy characters and a recognised part of British literature. It is also well known worldwide, with the books selling around 500 million copies worldwide. The books were later adapted into a set of films that became one of the most popular fantasy movie series (Pottermore, 2019).

The trend started with a video portraying Harry Potter in the style of an "*80s Yakuza film"*. The initial video showcased a montage of AI-generated and animated clips featuring Harry Potter characters reimagined as figures from an "80s Yakuza film". After the first Japanese remake version that utilised AI, the video gained over 1.1 million views and generated numerous versions, including Harry Potter "Italian", "German" and "Brazilian" versions. This meme was shared mostly in the form of a video created using AI tools (Philipp, 2023).

Most of the scenes are the same in all videos; the sequence and structure of the video repeat itself in multiple versions. The video shows many characters from the Harry Potter story framed in a medium shot, looking at the camera. What differs one version from the other is how the characters are portrayed, which changes according to the stereotypes of the countries; the clothing, hairstyles and background change according to the version of the meme. In the Italian version, for instance, all the characters of Harry Potter are depicted wearing fancy suits or dresses with sunglasses and hats. In the background, the bright sunny skies of Italy are depicted. These elements refer to the stereotypes of Italian mafia movies like "*The Godfather*", where characters are dressed similarly. The characters' voices in the video trailer also change according to the language of the country represented.

Nowadays, with the help of AI, it is possible to recreate the voices of real actors in many different languages. One of the key elements in this meme is that the characters are already known worldwide. The second element is the humour produced by the violation of expectation. The viewers are used to seeing a British version of Harry Potter, and when exposed to a different version like Japanese or Italian, they might feel a violation of expectation, and in this case, a harmless and benign violation that produces humour.

This type of video is very engaging for portraying a pop culture figure and for generating curiosity and interaction if the country is represented in the interest of the viewer.

6.3 - Barbie Heimer

The meme "*Barbenheimer*", or "*Barbie Heimer*", originates from the merging of two famous movies, "*Barbie*" and "*Oppenheimer*," this is an allusion to the simultaneous release of the film Barbie and Oppenheimer on July 21st, 2023. This coincidence of synchronicity sparked a wave of memes that juxtaposed the divergent themes of the two films. The term "*Barbenheimer*" gained popularity across social media platforms, particularly in the original discussion about which movie people should watch. The memes often featured images incorporating elements of both films and their colour schemes: pink for Barbie and black for Oppenheimer. These two colours naturally produce a high contrast. In Western contemporary societies, pink is often associated with femininity and black to dark and gothic motif, which for some is naturally more "male". This created a visual contrast that blended Barbie's girlish aesthetics with Oppenheimer's explosive motifs. In April of 2022, Twitter accounts sparked the initial surge of "*Barbie vs. Oppenheimer*" memes with tweets using the term "*Barbenheimer*". Although the term was first used in 2022, its first-known viral post emerged in January 2023. As 2023 unfolded, the viral usage of "*Barbenheimer*" persisted on Twitter, particularly gaining traction closer to the premieres of the films in June 2023. In early July 2023, numerous Facebook and Instagram accounts began sharing *Barbenheimer* fan art, blending images of Margot Robbie, the leading actress from Barbie and Cillian Murphy, the main actor from Oppenheimer, along with other characters, reminiscent of the "*Barbie-Oppenheimer*". This meme plays with the duality of opposing things and can be interpreted as the rivalry between the feminine and masculine sides (Owen, 2023).

Barbie is a well-known pop figure. It is a doll created in the late 1950s and made especially for girls. It was present in the infancy of many women born in the last 50 years (Lord, 2024). The character gained popularity and expanded to many countries outside the Western cultural bubble.

Oppenheimer is the story of a real physicist, J. Robert Oppenheimer, who participated in the creation of the atomic bomb during the Second World War, the deadliest type of weapon that humankind has ever been able to develop (Rouzé, 2024).

Both movies also differ in the genre of the film. While Barbie can be classified as a comedy or musical, Oppenheimer is a drama mainly based on the real story of the physicist who worked on the project that developed the atomic bomb. The structure of the memes plays with this duality and contrast and often depicts "two sides of a person". Many captions often instigate the viewers to choose a side or play with the fact that many have these two opposing sides, the feminine and masculine. The meme created by the page "*Melted Videos*" clearly expresses this merging.

Figure 20. Barbenheimer (Source: Know your meme, 2023).

The meme makes an extra connection between another famous audiovisual piece often associated with masculinity: The "*Peaky Blinders*" series. This ability to create

intertextuality is one of the most exciting features of the memetic culture. The original idea of Barbie versus Oppenheimer was then slightly adapted in this meme. The image portrays three men dressed in clothes from the beginning of the twentieth century, driving one of the first models of cars, with their eyes looking forward and transmitting confidence. All men wear hats, and the main character smokes a cigarette, which is also a symbol of masculinity. In addition, the main character is Cillian Murphy, who also plays the main character in Oppenheimer. The confident "badass" attitude contrasts with the pink colours and sparks of hearts added in the editing, elements that often reference feminity.

In addition to the visual codes, the verbal code in the image in the form of a title plays directly with this odd combination. It says "Pink Blinders," a direct reference to the name of the series "*Peaky Blinders*". The verbal pun and the odd visual combination are what create the humour. In a more social interpretation, this meme plays with the two sides of each person: the Peaky Blinders characters symbolising roughness and the pink colour symbolising delicacy. This trend generated numerous variations that played with either the rivalry between opposing sides or the combination of masculine and feminine.

6.4 - Black Cat Zoning Out

Figure 21. Black Cat Zoning Out
(Source: Know your meme, 2023)

The meme "*Black Cat Zoning Out*," also referred to as "*Black Cat Zoned Ou*t" or "*Staring Black Cat*," originates from a video capturing a black cat gazing blankly at another cat. The video starts with a black cat looking at another cat and slowly closing his eyes while the other cat moves his mouth, simulating the act of “speaking”. The cat's facial expression is similar to a behaviour meme creators associated with the term "zoning out.” The verb zoning out is often described as the state where you become unaware of your surroundings, often due to relaxation or boredom. This led to pairing the video with captions related to losing focus or dissociating. The original video originated on March 2023, when the TikToker @vivizirumugireo shared a video featuring two cats, with the black one staring at the tan-coloured one before eventually “attacking” it.

Over the next four months, the video garnered approximately 242,800 views and 21,000 likes. Subsequently, on March 27th, 2023, another TikToker uploaded a

cropped version of the video, adding a remix of a Snoop Dogg song as the audio, which accumulated around 1.7 million views and 300,300 likes within four months (Owen, 2023). This modification in the video played a significant role in the meme's viral dissemination.

The meme template gained traction primarily on TikTok in 2023, where creators utilised a greenscreen template to overlay the cats onto various backgrounds. Additionally, TikTok users utilised the song "*Aquatic Ambience*", contributing to the video's humour. On April 15th, 2023, the video was shared, placing the cats in a different scenario, accompanied by a top caption reading, "When you randomly zoned out during a conversation and now have no idea what was said to you." The video reached approximately 4.6 million views and 611,500 likes within three months. After gaining popularity, numerous TikTokers created their own versions of the meme, resulting in over 11,500 videos being created within the same week. The meme format quickly spread to other platforms, including Twitter and Instagram.

Notably, the most viral use of the greenscreen was by TikToker @oscarchavezneri on June 30th, 2023, accumulating around 8.1 million views and 588,100 likes within five days (Owen, 2023).

The video's main humour element is the black cat's facial expression. The eyes slightly closing, gazing away, giving this zoned-out expression resonated with numerous social media users. "Zoning out" is a common practice among humans. It can occur in many scenarios explored by the meme creators, who would maintain the black cat but change the background and the captions, which is an excellent example of mimicking.

The theme of animals is usually very popular among internet users; the meme "*Black Cat Zoned Out*" is not the first cat to go viral; there were other animals with funny related facial expressions that gained popularity among the internet communities. The resonance with human behaviour is what made the meme so famous. Different variables produced the humour. First, the fact an animal is acting in a similar way

humans do. Domestic animals such as cats and dogs usually grab people's attention for their human-like behaviour, which is often seen as humorous because it violates the idea that animals should act like animals and not imitate humans. The second factor for the humour is the cat's facial expression. The specific way the cat closed its eyes, the subtle smile, and the angle at which the animal was recorded played a vital role in generating humour. The flexibility of the expression, which can be used in many different situations, is the perfect feature of a meme. The meme can be adapted to multiple different scenarios and contexts.

6.5 - Hide the Pain Harold

Figure 22. Hide the Pain Harold
(Source: Know your meme, 2015)

"*Hide The Pain Harold*", commonly known as "*Harold*" or "*Maurice,*" is the name of a Meme featuring a senior stock photography model whose facial expression suggests he is hiding feelings of pain or discomfort. This meme is over ten years old and is an example of a meme that has survived through time. Most memes are only visual commentary that lasts only a few months; however, "*Hide the Pain Harold*" can still be seen around the internet.

The earliest documented discussion about Harold can be traced back to a thread on the *Facepunch* forum dating back to 2011. In September of that year, a user shared stock photos featuring the older man, sourced originally from *Dreamstime.* After the release, users started playing with the images and developing humorous content. Harold's stock photos sparked a lengthy tribute thread on the *4chan* board on May 5th, 2014. This thread crafted a fictional narrative about an elderly man unhappy in his role as a stock photography model. Subsequently, on September 7th of the same year, an *Imgur* user compiled notable quotes from the *4chan* thread into a gallery post

titled "*Hide-the-pain-harold*," which garnered over 880,000 views in over three weeks. The meme evolved into thousands of different forms and expanded to multiple communities worldwide, all using the funny expression of the subject, the uncomfortable smile of an older man, which resonates with many people who feel the need to express a fake smile.

Most variations of the meme often depict the character Harold in a situation job related. The fact the real model also took several pictures for the stock photos site, increased the variations of the meme. *Hide The Pain Harold* is still part of the internet culture more than ten years after the first versions.

The real identity of the Harold character was later confirmed to be András Arató, a resident of Kőszeg, Hungary, who revealed himself as the man in the stock photo. In September 2018, Arató delivered a TEDx Talk in Kyiv, Ukraine, where he delved into his journey as a "meme-hero" alongside his career as an electrical engineer in Ukraine. The meme's origin unfolds with a vacation in Turkey, during which a photograph of him was captured. After sharing the photo on Facebook, a professional photographer contacted Arató regarding modelling opportunities, leading to a successful trial shoot. Pleased with the outcome, they continued collaborating, creating "a couple hundred" stock photographs.

In the subsequent months, Arató discovered his photos were being used in various contexts beyond the agreed-upon stock photography, and he ultimately concluded there was limited recourse available to deal with the situation. The remainder of his speech revolves around his gradual acceptance of his meme fame and his decision to embrace the notoriety, culminating in establishing a homepage for the photographs (Know your Meme, 2014).

6.6 - Philosoraptor

Figure 23. Philosoraptor
(Source: Know your meme, 2009)

"*Philosoraptor*" is a meme series characterised by an illustration of a velociraptor accompanied by captions that portray the dinosaur engaging in philosophical inquiries and quirky paradoxes. The image is an illustration that depicts a close shot of a green dinosaur, specifically a velociraptor, with the mouth slightly open, showing almost no teeth and its claws close to the jaw, imitating the body posture associated with a philosopher. This body posture resembles previous statues of human figures sitting with one of the hands supporting the jaw, such as "*The Thinker*" by Rodin and "*Lorenzo de' Medici, Duke of Urbino*" by Michelangelo.

The original illustration of the "*Philosoraptor*" was created and copyrighted by Sam Smith. It was initially released on October 8th, 2008, as a T-shirt design available on the online retailer Lonely Dinosaur. In correspondence to the website Know Your Meme, the author Sam Smith disclosed that the concept for the "*Philosoraptor*" image originated in the summer of 2008. To create the image, Smith combined various online images of velociraptors, compressing them into one-color representation. He altered the jaw of one raptor to appear open, resembling a thoughtful expression.

Additionally, he fashioned a claw using an eagle talon image, modifying it to resemble a raptor's claw. The finishing touch involved adjusting the eye slit, which, when shifted, gave the dinosaur a contemplative gaze (Kikinak, 2009).

The image was inspired by a friend of the creator named Devin, whose nickname was "*Philosoraptor*." At the time, they were unaware of similar jokes circulating online. Although the term "*Philosoraptor*" (a blend of philosopher and velociraptor) is featured on the original T-shirt design, it might have been created before the release of Lonely Dinosaur merchandise. Given the phonetic resemblance between the two words, it is plausible that someone else could also have independently developed the concept of a "*Philosoraptor*"(Kikinak, 2009).

The comedy generated by the meme is firstly associated with the humorous connection between the two words: velociraptor and philosopher. This strange combination, added to the illustration depicting a dinosaur, an extinct animal often associated with violent behaviour mimicking a human thinker, creates the humour. The subtleness of the facial expression of the dinosaur also increases the humour. The eyes of the dinosaur, the mouth open, and the claws strongly reinforce the philosopher's posture. The meme's verbal codes also play a vital role in generating humour. The different possibilities and interesting paradoxes proposed in the phrases helped the meme's popularity. The simple format of the me with the face of the dinosaur provides a structure that can be easily replicated. The meme appeared in many different versions and inspired new memes. Despite its popularity, the meme is currently decaying in popularity.

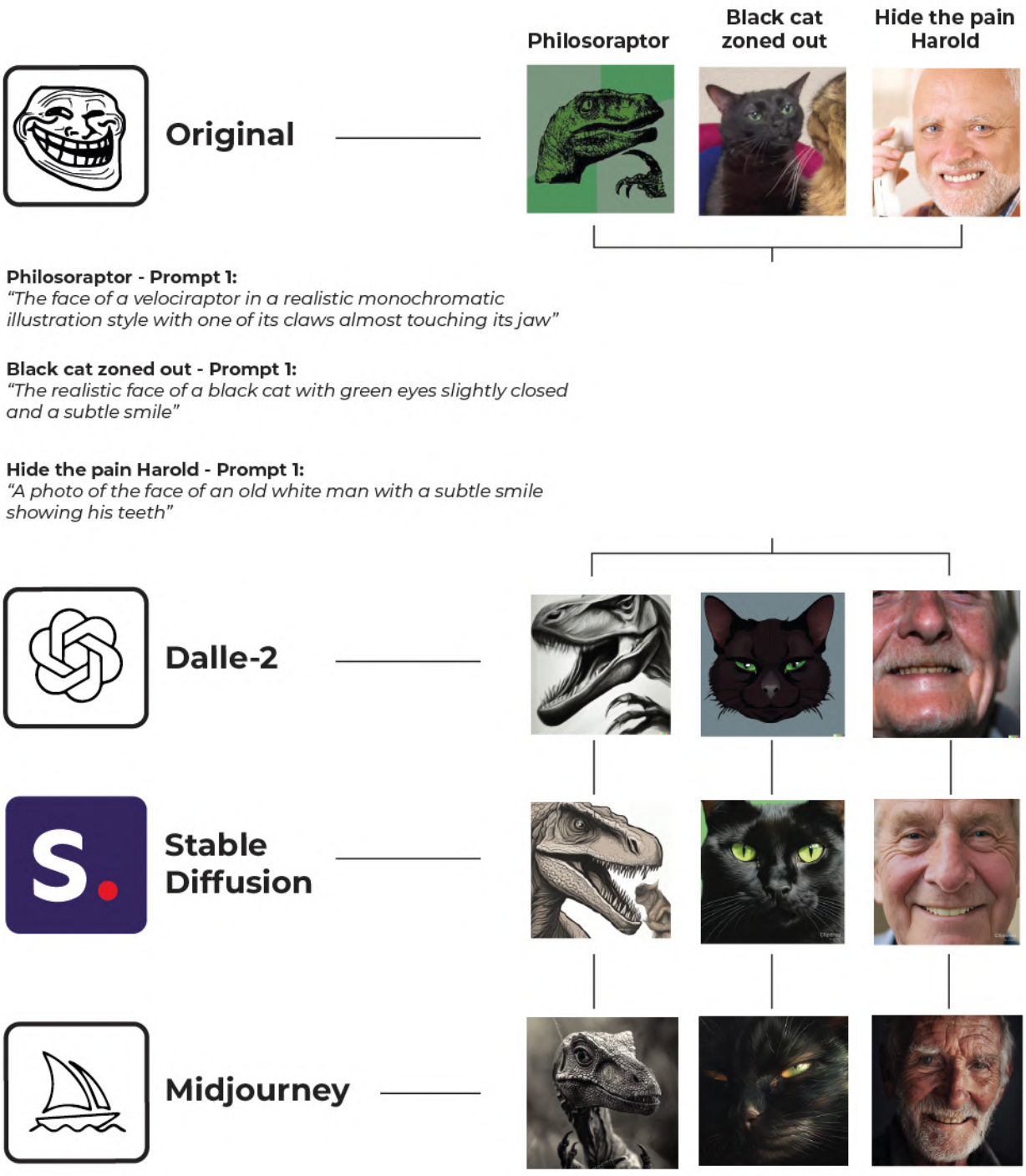
Knowing the social context and history of certain memes is essential to understanding how a "silly" image can become a mass instrument of communication, able to alter its original meaning and convey multiple ideas using almost the same visual codes.

7.0 - Recreating Memes with AI

The next step to assess the potential of AI is to test its capabilities to recreate memes. In this experiment, three memes will be reconstructed using generative AI tools. The selection criteria for these memes include popularity and longevity, intending to choose memes that have either endured over time or gained significant popularity. Additionally, the diversification part of the criteria included the depiction of a human and an animal and a category for the style of illustration. This experiment will measure the AI's proficiency in depicting different elements and styles by including a range of subjects, offering insights into how the machines see, interpret and create visual codes.

Recreating memes with **AI tools**

PART - 1



Philosoraptor - Prompt 1:
"The face of a velociraptor in a realistic monochromatic illustration style with one of its claws almost touching its jaw"

Black cat zoned out - Prompt 1:
"The realistic face of a black cat with green eyes slightly closed and a subtle smile"

Hide the pain Harold - Prompt 1:
"A photo of the face of an old white man with a subtle smile showing his teeth"

Diagram 1. Recreating memes part 1.
(Source: Developed by the author, 2024)

7.1 - Attempt 1 - Black Cat Zoned Out

Figure 24. Black Cat Zoned Out made with AI (Source: Developed by the author).

The attempt to recreate the meme "*Black Cat Zoned Out*" revealed that different platforms generate different images even with the same prompt; this could be explained by the fact that each AI tool is fed and trained with different datasets. In the attempt to recreate the "*Black zoned-out cat*", DALL·E 2 managed to produce the best result, taking into consideration the humour criteria. In this attempt, it is possible to recognise a similarity between the original meme and the recreation. DALL·E 2 depicted the cat in the illustration style even without a prompt; this differs from the original meme, which includes an image of a real cat. The DALL·E 2 image is framed in a close-up of the face of the cat and portrays the cat with a cunning look, in a certain way, similar to the original meme; however, the eyes are too closed, and the facial expression gives the cat a malicious look rather than a "zoned out". One of the flaws in the image is the direction of the look; the cat is looking directly at the spectator; also, the original black cat had a more silly expression in the eyes, and it is looking away from the scene, giving the impression that the cat was not paying attention to the present moment, the famous zoned out expression.

Stable Diffusion generated a more realistic representation of the cat, contrasting with DALL·E 2, which represented the cat in a more cartoonish way. In this specific point, Stable Diffusion represented the cat closer to the original form. However, the depiction of Stable diffusion fails considerably in delivering a funny expression,

which is one of the most important features of the meme. The funniness was ignored entirely by the AI, which only depicted a close-up of the face of a black cat gazing directly at the viewer, expressing no comedy. Again, the zoned-out look is absent even though the prompt described a smile and the eyes “slightly closed”. The output generated by Stable Diffusion does not depict the eyes closing, at least not enough; it might have misinterpreted the word slightly. In addition, the cat's smile differs strongly from the original, and it is so subtle that it cannot be perceived.

Midjourney first generated a close-up image of the face of the black cat with dramatic lighting illuminating only part of the cat’s face. The image is rendered realistically but resembles more a digital painting than a photograph. The eyes are slightly closed but do not convey humour; the cat's look is again more malicious than funny. There is no connection to the idea of zoning out; on the contrary, the cat seems observant. This attempt was probably the most distant depiction from the original meme because it does not convey humour in any way despite the impressive skills of the AI tool, which rendered a highly technical, dramatic digital painting of a black cat in less than a minute.

7.2 - Attempt 2 - Hide the Pain Harold

Figure 25. Hide the Pain Harold made with AI (Source: Developed by the author).

In the attempt to recreate the meme "*Hide The Pain, Harold*", DALL·E 2 rendered a realistic photographic close-up image of an old white man smiling. The first issue can be pointed out to be the framing of the image; somehow, the angle at which the image was rendered seems strange, almost from the point of view of someone looking at the man from below. DALL·E 2 also excluded most of the subject's face. The eyes are almost out of frame, making the image unpleasant. The lighting resembles a flash effect with white points of light reflected in the nose, jaw and checks of the subject. The smile is subtle but does not transmit humour; neither can it be associated with the idea of a "*hiding the pain smile*."

The attempt of Stable Diffusion generated a centralised image of an old white man smiling in a photorealistic style. The lighting is uniform, and the shadows create little contrast on the subject's face. The subject's eyes are directed at the viewer, and the image is also a close-up but not as claustrophobic as the depiction from Dalle 2. The subject's expression is of joy, and the image passes the idea of tender happiness but does not convey the idea of humour.

Midjourney created a more dramatic depiction of the subject with a dark background and contrasting shadows. The man's face is centralised but slightly tilted. There are many details and textures on the man's skin. The facial hair,

wrinkles, and skin pores can be seen. The man has a beard similar to the original meme, even though the prompt did not suggest it. The man's look is of joy despite the contrasting light and overall dark tones of the image.

7.3 - Attempt 3 - Philosoraptor

Figure 26. Philosoraptor made with AI (Source: Developed by the author).

The image created by DALL·E 2 as an attempt to recreate the meme "*Philosoraptor*" delivered a pencil illustration style of a velociraptor with a mouth wide open, showing its teeth and the impression of a roar. The image was rendered in black and white, and it also depicted the dinosaur's claws close to the mouth according to the prompt, but the eyes are completely black, and there is a sense of dynamism. The overall impression of the image is of a dinosaur in movement, angry, and ready to attack, which is completely different from the meme, in which a dinosaur is thinking and with a "philosophical" posture.

The "*Philosoraptor*" of stable diffusion is an illustration style of a dinosaur with a wide open mouth and many sharp teeth. The image frames the face of the dinosaur in a close-up, and the overall colour is beige with low saturation. The main flaw of this attempt is probably the dinosaur's claw, which was supposed to be close to the mouth, but instead, the AI rendered an abstract, distorted shape. The image also

rendered a lot of texture on the face of the dinosaur. Similarly to DALL·E 2, the overall impression is of anger instead of a contemplating dinosaur.

The attempt from Midjourney is probably the closest to the original meme. The image generated is a photo-realistic depiction of a velociraptor in almost black and white, with a very subtle inclusion of the beige colour. Differently from the other AI, the dinosaur does not have its mouth wide open and does not have the facial expression of anger and attack. The mouth of the dinosaur is slightly open, showing only a few front teeth, giving the impression of a smile. The framing of the subject is more centred and does not cut parts of the dinosaur face like DALL·E 2 and Stable Diffusion. The claw of the dinosaur is well positioned close to the jaw. The skin of the dinosaur was rendered with a lot of texture. The overall impression of the image is more comic and friendly in comparison to others, but it does not pass the idea of a "philosopher".

7.4 - Further Attempts

Due to access limitations, the additional attempts were made using only the paid version of Midjourney; other AI tools like DALL·E 2 and Stable Diffusion were excluded from this experiment.

To obtain different results, different prompts were used to explore the AI's strengths and weaknesses and to try to reach the closest resemblance to the original memes. Another objective of this series of attempts was to reach a resemblance and a similar effect of humour, even if the image is not perfectly similar to the original memes. The main objective is to transmit a humorous idea even if slightly different visual codes are used.

Recreating memes with **AI tools**

PART - 2

Original

Philosoraptor

Philosoraptor - Prompt 1:
"The face of a velociraptor in a realistic monochromatic illustration style with one of its claws almost touching its jaw"

Black cat zoned out

Black cat zoned out - Prompt 1:
"The realistic face of a black cat with green eyes slightly closed and a subtle smile"

Hide the pain Harold

Hide the pain Harold - Prompt 1:
"A photo of the face of an old white man with a subtle smile showing his teeth"

Extra attempts & new prompts

Midjourney

Hide the pain Harold - Prompt 2:
"A 60 year old white man with a smile that looks like he is hiding pain"

Philosoraptor - Prompt 2:
"A philosopher in a form of a velociraptor"

Black cat zoned out - Prompt 2:
"The face of a black cat zoning out"

Philosoraptor - Prompt 3:
"a velociraptor posing like a philosopher in a one colour style illustration

Black cat zoned out - Prompt 3:
"A black cat talking to another cat but with a zoned out expression"

Diagram 2. Recreating memes part 2. (Source: Developed by the author, 2024)

7.4.1 - PROMPT 1: “The face of a black cat zoning out.”

Figure 27. Black Cat Zoning Out extra attempt 1 (Source: Developed by the author).

The idea for this prompt was to test the ability of the AI to identify what “zoning out” means. It was thought that if the term indicated, the AI would render an image closer to the meme. Differently from the expectation, this attempt generates a close-up, realistic image of a black cat looking at the viewer, but with no “zoning out” expression. Contrary to the first prompt, adding the words “zoning out” did not help the computer express this concept visually. The images generated were even more tedious than the first attempt. One of the images created with this prompt frames the cat in a side view and gives the impression that the cat is looking away, which can be interpreted in a certain way as a “zoning-out” expression.

7.4.2 - PROMPT 2: “A black cat talking to another cat but with a zoned-out expression.”

Figure 28. Black Cat Zoning Out extra attempt 2 (Source: Developed by the author).

This prompt was an attempt to enhance the zoning-out idea by adding another cat to the image. The original meme contains a second cat that plays an important role in setting up the context even though the second cat is only represented partially and from a back point of view. This modification in the prompt made the computer generate a second cat but with no zoned-out expression. Even with the ideas of “zoning out” and “talking to another cat” added to the prompt, the AI generated a dull image of two carts looking at each other.

7.4.3 - PROMPT 3: " A 60-year-old white man with a smile that looks like he is hiding pain."

Figure 29. Hide the pain Harold extra attempts (Source: Developed by the author).

The objective of this prompt was to provide a hint to the AI about what type of smile the subject should have. For this, some words like "hiding the pain" and "smile" were added. This attempt increased the difficulty of the AI to generate satisfying images, especially because of the conflicting ideas such as "pain and smile". The first significant change the AI performed was that in all depictions, the man has his eyes closed. This might have occurred because of the word "pain". According to this output, it is possible to say that for the AI human beings always close their eyes while in pain; hence, a "hiding the pain" smile should depict a man with his eyes closed. The second feature was the closed mouth; for the computer, hiding the pain type of smile should depict the lips closed and no teeth exposed. This feature strongly contrasts with the original meme in which the front teeth of the subject are exposed. The lighting of the images varied, but in most examples, there is a significant contrast between dark and light areas; this increases the dramatic effect of the image. The man's facial expression transmits more the idea of pain than humour. Moreover, in all three images, the AI failed to depict the age factor mentioned in the prompt. The depictions do not seem to match the idea of a "60-year-old man", but instead the men portrayed appear to be above 75 years old. With these results, it is possible to say that the AI failed to recreate a "Hide the Pain Herald" type of meme.

7.4.4 - PROMPT 4: “A philosopher in the form of a velociraptor.”

Figure 30. Philosoraptor extra attempt 1 (Source: Developed by the author).

This prompt proposed a different way to achieve the idea of a “*Philosoraptor*”. Differently from the first attempt, this second prompt provided the concept of “philosopher” and “form” to the machine. However, even with those new words, the final results did not reach a similar effect compared to the original image. Instead, The computer created a more literal combination of both concepts and generated an image of a “reptile figure” in an 18th-century clothing style, holding a book. The “reptile figure” is more similar to a lizard than to a velociraptor. The character is portrayed from the waist up with its face in a side view. The framing of the image slightly reminds a painting from the 18th century.

The same prompt also generated a similar image of a dinosaur reading a book. Still, the velociraptor is dressed in a toga, even though there was no direct reference to clothing in the prompt. It is possible to find a certain degree of humour in these images. The combination of concepts is indeed interesting, though the lack of

expressive facial and body posture diminishes the humorous effect of these attempts. The facial expression and the hand gesture were essential parts of the original meme.

7.4.5 - PROMPT 5: “A velociraptor posing like a philosopher in a one-colour style illustration.”

Figure 31. Philosoraptor extra attempt 2 (Source: Developed by the author).

In this prompt, the idea was to add the words “one colour illustration style”, with the expectation that the AI could generate something more similar to the original meme. The visual impact of the term “illustration style” contributed to the humour in the image. In this attempt, the rendered image illustrated a velociraptor with its mouth closed and its face depicted in a side view. The eyes stare away, and the image is painted in greenish tones. The dinosaur has its arms crossed and is wearing clothes that appear to be from the early 20th century. Despite the interesting and high-quality results, the image does not resemble the original meme; however, in two examples, it is possible to consider a certain degree of humour. In one of the versions, the dinosaur is sitting with its legs crossed, similar to a meditation posture; the dinosaur is dressed in a robe, which could reference Asian cultures. The facial

expression of the dinosaur is calm; the mouth is shut, and the eyes are slightly closed. The other successful attempt also depicted the dinosaur sitting with its legs open in a relaxed posture. The dinosaur's claws are depicted similarly to human hands; one of its hands is holding the dinosaur’s jaw while the subject looks away, giving the impression that the velociraptor is thinking.

These two attempts can be considered successful in presenting images with a potential for humour. The facial expressions of these two images convey the idea of a calm, thinking dinosaur contrasting with the animal itself, which is often portrayed as a ferocious animal. This contrast of ideas violates the expectation of seeing a violent velociraptor; thus, it could generate humour. Even though these images do not accurately resemble the original meme, they transmit a similar idea, the idea of a dangerous dinosaur having a philosophical moment; hence, these images could work as memes, especially if the right text was added and shared in the right context.

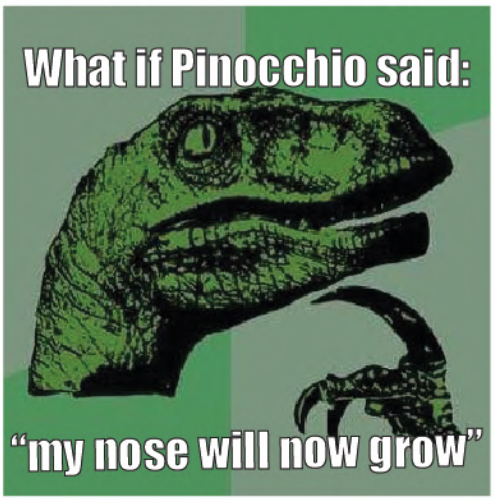
Recreating memes with **AI tools**

PART - 3

Original

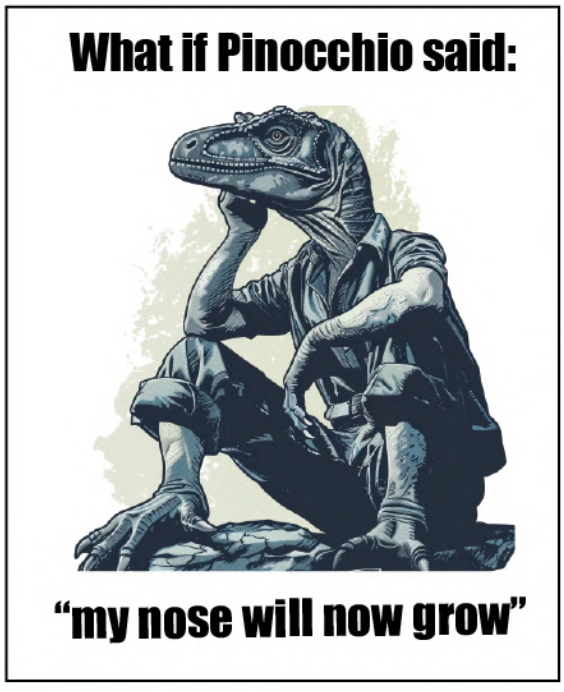
Philosoraptor

Example Philosoraptor with captions



Midjourney

Philosoraptor - Prompt 3:
"a velociraptor posing like a philosopher in a one colour style illustration



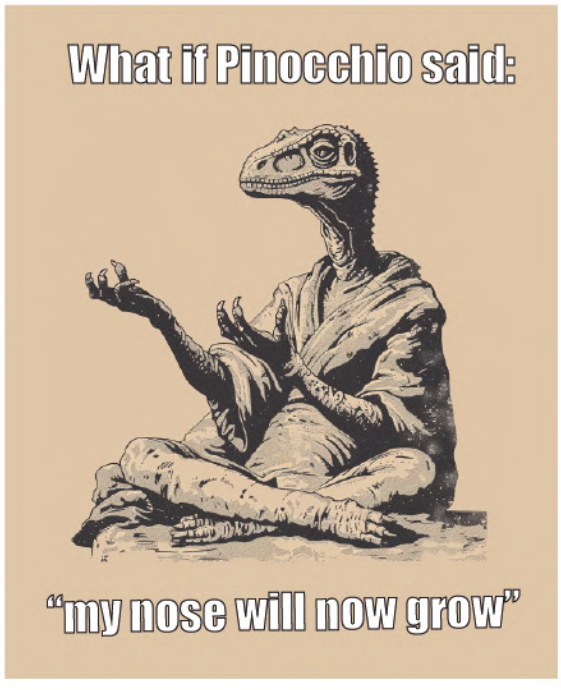


Diagram 3. Recreating memes part 3 (Source: Developed by the author).

8.0 - Discussion

Contrary to popular belief, visual communication is not universal; it is inherently encoded and appears transparent only because people are already familiar with the underlying codes presented in certain images. Can a computer fed with billions of images from multiple cultural sources, translate verbal prompts into effective visual codes? Can AI tools be an important part of the creation of visual humour? The data shown in this research expressed that it might not entirely happen.

Without explicit knowledge of these codes, people may struggle to comprehend certain visual cues unless they are integrated into the cultural contexts from which these images originate. How can a computer be fully immersed in a culture to master its visual codes? To create and properly see visual content, the machine should consider the context and history of these images in relation to the various other images that coexist in the online environment. The changing dynamics of the visual landscape are influenced by social, cultural, and economic factors, including the increased linguistic and cultural diversity within nation-states and the diminishing significance of national boundaries due to multiculturalism and electronic media (Kress & Leeuwen, 2007). Even in a more multicultural world, the plurality of human groups still exists. Is artificial intelligence able to differentiate the types of visual codes used by these groups? Can it understand how they think and communicate using the right visual grammar? The first flaw of these technologies is the lack of connection to human feelings.

8.1 - No Body No Connection

Any semiotic mode has to be able to represent aspects of the world as it is experienced by humans (Kress & Leeuwen, 2007, p. 42). Many people see art as a way to connect with the artist; they are attracted by how another human sees the

world. This lack of perspective from AI art can cause a lack of interest among the spectators of AI images. The images created by AI may give the impression that the author has no point of view or perspective. This lack of personal vision can distance the AI and the viewers. Since the machines do not possess a physical body or sensors, they only experience the world through data. This lack of real-world participation can also interfere with the cultural knowledge of the machines. To master the cultural visual codes, one must be fully immersed in a culture. Even with all the verbal and visual data collected online, it is not enough to generate immersive participation.

Visual communication also possesses tools for establishing a different form of interaction, one that occurs between the creator and the observer of the image. Visual communication engages two types of participants: the represented participants (including individuals, locations, and objects portrayed in images) and interactive participants, those who engage in communication with one another, namely the creators and viewers of the images (Kress & Leeuwen, 2007, p. 114). In the world of memes, this interaction is vital for the propagation and development of memes. The whole mimetic culture runs around images that are imitated and transformed, provoking the involvement of the observer who feels connected to the idea and engages as a participant, sharing or recreating the memes. This engagement is core to the circuit of culture, in which images are socially produced, distributed and consumed. Within this cycle, there are processes of transformation taking place, as well as struggle and contest of what they mean and how they are used.

This immersion in the culture also plays a role in visual humour's intertextuality, especially in the meme culture, in which a meme can generate another meme. These discourses are articulated through different forms, which can be visual or verbal. The diversity in the forms of discourses is an example of how dependent meanings are from other images and texts (Rose, 2007, p. 142).

In humour, the audience plays an important role in decoding the significance or creating new memes for the same visual codes. The previous example of a zoned-out black cat reinforces this idea. The memetic process only occurred after an interpretation of the information that resonated with the public, after one of the viewers associated the cat's face with the act of zoning out. This interpretation, added to other resources, such as the music and the flexibility of the meme theme, were important reasons for its success. The creation of this meme would not be possible through the use of AI alone. As shown in the attempts to recreate memes, the subtleness of the facial expression is extremely hard to reproduce. The machine does not possess its own knowledge and understanding of the experience of seeing.

It could be argued that the machine does not need a physical body to generate art since it is only a tool used by humans. The human artist should guide the machine to achieve the best visual results. However, this mechanism might raise the problem of translating verbal concepts into visuals.

8.2 - Lost in Translation

One of the reasons for the low presence of these tools might be the difficulty of translating with accuracy verbal to visual code. Image-generating tools operate using verbal code to generate visual codes. The user of these tools must first write a prompt to guide the machine in the process of image generation. Depending on the chosen prompt, the AI might generate different results for different prompts. To generate properly, these technologies require specific verbal details. The possibility of ambiguity in verbal codes can be a barrier to visual creation. Different from images, words convey broad meanings and instigate the imagination. A word can be visualised in multiple ways according to the reader's imagination. This might create a problem with the translation from verbal to visual. The user of these tools must first visualise mentally the idea of the meme and then translate it into verbal code.

This method can be a problem since many artists are not used to visualising ideas in their minds and translating them into words.

Describing images is part of the concept of "*ekphrasis*", which is the verbal representation of visual representation. Essentially, *ekphrasis* serves as a literary depiction of art, using verbal language to paint a picture. What sets it apart from other forms of imagery, like pictorialism, is that the image created is a visual representation made through verbal communication. *Ekphrasis* orchestrates a meeting between representations in two different mediums: one visual and the other verbal. One of the earliest examples of *ekphrasis* can be found in Homer's "*The Iliad*," where the poet extensively describes Achilles' shield. Contemporary writers frequently draw inspiration for their verbal visualisation from digital or physical artworks (Olson, 2020).

The necessity of describing the image verbally can create an obstacle for AI users. Not all artists are familiarised with the concept of *ekphrasis*, and not many have the skills to paint an image using only verbal communication. As mentioned by Ben Davis (2023) in his article about the future of AI and art:

"Although the very accessibility of these generators is what has made the subject of "AI Art" such a white-hot topic. With minimal effort, anyone with an internet connection can pump out pictures that look like artworks that would have required some sophisticated illustration skills just a little while ago. This fact has injected a strange, sometimes delusional, probably temporary euphoria into the recent conversation. The internet is being flooded with low-effort A.I. junk, trying to capture your attention with the least possible effort. Since the impression of general artistic coherence can be generated easily, but it still takes effort to nail down all the details."

Mastering these tools requires a knowledge of verbal description of visual concepts, a skill that not many artists have mastered. Even subtle changes in the verbal description can generate completely different outcomes. "The Mona Lisa

Experiment" shown previously is a good example of how a simple word can alter how the computer sees. The illusion that these AI tools required no technical skills seems to be only euphoria. Even though memes do not require advanced artistic skills, creating humorous visual concepts through translating verbal concepts is not simple. Both visual and verbal structures can be used to express meanings drawn from a common cultural source; the two modes are not simply alternative means of representing the same thing. Only a detailed comparison can bring out how, in some respects, they represent similar types of meaning and in others, they represent the world quite differently (Kress & Leeuwen, 2007).

8.3 - Bad Vision and Humour by Chance

Computers still have inferior "eyes" compared to humans; even though AI has shown an amazing capacity to create images, its visuality still has some evident flaws. The limitations of computer vision technology obstruct its ability to accurately perceive and render subtle details, such as specific facial expressions required to capture the essence of a meme. Maybe due to the lack of training or the absence of a physical presence, the AI eyes lack a certain sensitivity to details. During the meme experiments, the AI failed to create a smile that could also be associated with pain, two ambiguous and contrasting ideas. However, even if this task was given to a group of human artists, they might also struggle to depict these concepts. The main difference in these contexts is that the human eyes can recognise a "hiding the pain" type of smile, while a computer is still unable to perceive such subtleness. The birth of these memes happened by chance, through the eyes of humans who identified certain characteristics of these images and associated them with humorous ideas. Both the "*Black Cat Zoned Out*" meme and "*Hide The Pain Harold*" were not initially created to become memes. Their initial meaning was reinterpreted and transformed into something humorous.

How could the AI differentiate a smile of joy from a smile of pain? Technically, they look almost the same. Maybe the key element is how wide the smile is or the expression of the eyes. Similarly, in the case of the meme *"Black Cat Zoned Out"*, the subtleness of the eyes closing and the smile are imperceptible to the machine. The results produced by the AI were realistic and technically advanced depictions of a black cat. However, they still lacked the invisible ingredient that makes the expression of the original cat funny. On the other hand, AI tools are very effective in creating alternative versions of already-known pop culture figures. In a specific trend of memes, the AI tools effectively generated the "Pixar/Disney style" of other pop culture characters. This trend proved the powerful skills of AI in rendering images and its ability to recognise certain visual features of the given characters. This ambiguity in potential is part of the AI technologies; on the one hand, they can perceive details and render high-quality images extremely fast, easily outperforming humans; on the other, they lack a precise vision to communicate and depict subtleness essential for expressing feelings and emotions visually.

8.4 - AI Limited Aesthetics

As is the case with any visual representation, AI-generated images also end up developing a specific style. The images created by AI tools tend to output similar patterns and structures, depending on which AI tool is used. The aesthetics of AI often reflect the programming and data input received, resulting in distinct styles for each tool. The reliance on learned patterns and replication could potentially restrict the creative output of AI-generated art. These systems can become conditioned to produce images within predefined parameters, limiting their ability to explore more unconventional or imaginative depictions. As a result, AI may struggle to break free from established rules and develop innovative artistic expressions. Many memes do not follow a precise technical approach to image creation. In some cases, the fact that the representation is badly depicted is one of the reasons that the image gains

funniness and goes viral. The limitation of experimenting with new ways of depiction can be seen as one of the major flaws of AI art. The transformational type of creativity, which is the most rare, the rule-breaking type, is quite limited by the functioning mechanism of AI. On the other hand, the explorational type is perfect for AI tools, as it allows the user to explore all the possibilities of already defined visual boundaries.

8.5 - Legal Issues

In the world of visual humour, ownership of content is almost nonexistent. Memes are created and distributed with the original author rarely being mentioned. Many meme creators appropriate images from the internet, modifying them without the original author's permission. The meme "*Pepe the Frog*" is a good example of how an illustration was transformed without the author's consent. Despite this cultural practice of the Internet world, the AI industry has been facing criticism from multiple artists. Laws are being discussed to regulate this market since, with these tools, it is possible to steal not only an image but the whole style of an artist. The computer can learn how to replicate an artist's style, which has caused immense turmoil in the creative field. Many artists feel like their years of study and practice have been stolen by these tools. Given that these tools create images based on existing data and patterns, issues regarding ownership and originality arise. Governments and societies must discuss how to deal with AI developers and artists. Determining the creator of an AI-generated image is a difficult task. Who is the author? The machine, the human who wrote the prompt or the artist who inspired the visual style? Currently, there are no specific parameters to judge authorship when AI systems are involved. As AI technology evolves, legal frameworks must adapt to address these copyright concerns and establish guidelines for ownership, attribution, and fair use of AI-generated content.

8.6 - AI Successful Attempts

During the attempts to recreate the "*Philosoraptor*" meme using generative AI, two images were identified as having "meme potential". Even though they lacked accuracy in terms of reproducing exactly the original meme, they had elements that could generate humour and produce a similar effect to the original meme. In these attempts, both images were created in the illustration style and reproduced the dinosaur in a humanised manner. The facial expressions, body posture, and clothing increased the comic effect. In one of these images, the dinosaur is depicted as an Asian monk sitting with legs crossed and the palms of the hands up, similar to a meditation posture. The humour in this image is generated by the contrast between two opposing ideas, the dinosaur as a dangerous animal and the Asian monk, a figure usually associated with calmness and kindness. The combination of these two ideas, together with the details of the eye, which is slightly closed, enhancing the impression of calmness, and the mouth, which has a subtle grin, generate a possible humorous image. The second attempt, which depicted the dinosaur in a classical philosopher pose with the hands supporting the jaw and the subject looking away, can also be recognised as an image with "meme potential". In this second image, the facial details enhance the touches of humour, such as the eyes looking away and the mouth of the dinosaur, which is closed, showing no teeth and revealing a subtle smile.

However, it is important to notice that the judgement of "meme potential" is subjected to the personal interpretation of the researcher. In the world of visual humour, specifically memes, there are numerous possibilities of making a joke, even the flaws or lack of humour of these AI images could be used as a way to generate humour. In addition, other social factors can influence the meme's potential, such as the context in which the meme is created, the timing of the message, and the content, for instance, if the meme is about a trending topic or who is sharing the meme. A meme shared by a creator with a reduced audience may not be successful.

Nevertheless, even with successful attempts, the use of AI to create visual humour still seems to be limited. The quantitative data also supports this claim, as well as the other failed attempts to recreate memes.

8.7 - Conclusion

After the development of this study, it was possible to identify a gap in the potential of AI in the field of visual humour. There is no doubt that the achievements of AI are remarkable and that, in the near future, these technologies will become more present in the lives of contemporary human beings.

However, according to the findings of this study, the field of visual humour seems to be partially excluded from the growing presence of these tools. Artificial intelligence is an inherently unpredictable and speculative field, with numerous variables playing a role in the future trajectory of these technologies. While many current issues exist, future advancements in the field could drastically challenge the current status quo. Ten years ago, the development of these tools would be unimaginable. Nowadays, these tools are easily available to almost anyone connected to the internet.

Establishing a true potential for these tools is susceptible to a high degree of speculation. The discourse surrounding AI will continually evolve, reflecting the technological innovation and its profound impact on society. In the creative field, these technologies have proven to have impressive capacities to render images closely to the skills of a mid-level designer/artist. Still, if taking into account the factor of time, the machines easily surpass human capabilities.

Hopefully, art is not about technical skills in the shortest time possible; there are subtle details regarding expression, point of view and resonance that can only be created through a physical embodied experience of the world. Art and humour cannot be translated into mathematical terms and algorithms, which are the basis of

the computer "brain". Moreover, these technologies have great potential to assist humans in numerous tasks, including the creative field. Nevertheless, this study has revealed that the image-generating AI technologies that operate in the text-to-image dynamics still have a timid participation in the process of visual humour creation. The reasons for that may be explained by the fact that humour and visual communication operate as a language that requires previous knowledge of the acceptable and popular codes, both visual and verbal. The lack of knowledge of these grammar rules might obstruct effective communication. The creators of popular memes have a strong capacity to formulate content that resonates with people. The resonance is one of the most important parts of communication; the viewer must somehow feel connected to the idea and the visual codes presented in the image. In the internet culture, the viewers are not passive receivers of messages but participants who actively engage in the sharing and transformation processes of memes.

The act of creating humour is highly subjective and deeply rooted in cultural context, social norms, and linguistic nuances. Creating memes often requires an understanding of these elements to convey humour effectively. While AI algorithms can learn patterns and generate content based on datasets, they may struggle to comprehend the subtleties of visual codes that create humour.

Secondly, memes often rely on juxtaposition, irony, and unexpected associations. These elements of humour can be challenging for AI algorithms to emulate accurately, as they require a deep understanding of context, semantic meaning, and sometimes paradoxical ideas. The theories of humour, such as the "*Theory of Benign Violation*", sustain this idea. According to this theory, to create humour, the machine needs a high level of cultural knowledge to perform a violation that is considered benign. Additionally, memes often incorporate elements of spontaneity. Many memes are comments about ongoing situations or events that either just happened or will happen in the future. The meme "*Barbie Heimer*" is an example of this

attachment to time. The meme started before the event and lasted a few months after the movies were released. The problem of time is a great challenge for generative AI tools. AI can generate content based on predefined parameters and datasets; they are still not able to generate images based on events that just happened or will happen. According to how these technologies operate, the computer must first be trained with visual data to create; this would require massive processing power and a high cost of energy, in addition to changes in the functioning systems of these tools that would need to be fed with images in real time.

In summary, while AI tools have made significant advancements in various creative tasks and have the potential to assist humans in creating visual humour, its presence in this field remains timid; visual humour is complex due to its subjective, contextual, and spontaneous nature that is deeply ingrained in human culture and cognition.

The main objective of this dissertation was to propose a discussion about the capabilities of artificial intelligence in the field of visual humour. These new technologies have sparked numerous debates about their influence in society, especially among artists and designers who, for a long time, believed to be exempt from the influence of automation.

One of the main challenges of this project was the complex nature of the topics approached. The difficulty of this research was increased by the lack of consent regarding the theory of humour and the fact that humour crosses multiple fields, including psychology, biology and sociology. In addition, the field of artificial intelligence is deeply complex as it englobes the realms of computer science and mathematics. Furthermore, these technologies still lack a complete understanding, even among specialists in this field; and they are constantly evolving, making it extremely difficult to develop an accurate analysis.

It is important to notice that the analysis proposed by this research is only temporary and might require adjustments in the future. The current flaws and obstacles for

these technologies presented in this paper might no longer be present in future versions of these tools.

Numerous factors might influence the potential of AI since these tools are still a novelty and require time to be fully implemented in society. Moreover, humans are still figuring out ways to use and benefit from these tools. There are also multiple legal and environmental issues regarding copyright regulations and the enormous amount of energy required by these technologies to operate.

Undoubtedly, AI's current achievements are impressive. Its presence in society is certain, but its true potential remains a matter of speculation.

***Figure 32**. Ogres are like cotton candy; both of them have sweetness* (Source: Instagram @wearearada, 2024)

9.0 - References

Arada [@wearearada]. (2024, January 25) "Ogres are like cotton candy; both of them have sweetness, Shrek" [Online Image] Instagram. https://www.instagram.-com/p/C2im17toAW3/?img_index=1

Artnet, A. (2022, March 3). The true story behind the origins of Pepe the Frog. Artnet News. https://news.artnet.com/buyers-guide/pepe-art-angle-transcript-2077584

Barbenheimer [Online Image]. (2023) Know Your Meme. https://knowyourmeme.-com/photos/2620827-barbenheimer

Barnard, M. (1998). Art, design and visual culture: An introduction. Macmillian Press.

Black Cat Zoning Out [Online Image]. (2023) Know Your Meme. https://knowyourmeme.com/memes/black-cat-zoning-out

Britannica, T. Editors of Encyclopaedia (2024, March 8). algorithm. Encyclopedia Britannica. https://www.britannica.com/science/algorithm

Copeland, B. (2016, November 20). artificial intelligence, situated approach. Encyclopedia Britannica. https://www.britannica.com/technology/artificial-intelligence-situated-approach

Davis, B. (2024, January 27). *10 predictions about the unexpected impact of A.I. on art (part 1 of 2)*. Artnet News. https://news.artnet.com/art-world/10-predictions-a-i-art-2411013

Domo. (2023). Data never sleeps 10.0. https://www.domo.com/data-never-sleeps

Donald Trump as Pepe the Frog [Online Image]. (2015) Know Your Meme. https://knowyourmeme.com/memes/pepe-the-frog.

Don, & Brad. (2011, November 8). *Neil deGrasse Tyson reaction*. Know Your Meme. https://knowyourmeme.com/memes/neil-degrasse-tyson-reaction

Edwards, P. (2023, April 4). Does AI still struggle with drawing hands? Vox. https://www.vox.com/videos/2023/4/4/23669625/ai-data-learning-hands-difficulty-patterns

Engraving of Kilroy on the WWII Memorial. [Online Image]. (2006) Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Kilroy_Was_Here_-_Washington_DC_WWII_Memorial.jpg

Epic Funny Page [@epicfunnypage]. (2024) Funny Memes & Videos. [internet memes] Instagram. https://www.instagram.com/epicfunnypage/

Festa da Firma [@festadafirma]. (2024) Corporate Memes. A Firma como ela é! [internet memes] Instagram. https://www.instagram.com/festadafirma/

Furie, M. (2015). *Pepe the Frog original character* [Online Image] Artnet News. https://news.artnet.com/buyers-guide/pepe-art-angle-transcript-2077584

9GAG [@9gag]. (2024) 9GAG: Go Fun The World [internet memes] Instagram. https://www.instagram.com/9gag/

Google Deep Dream. (2016). *Moonage Daydream* [Online Image] The Guardian. https://www.theguardian.com/artanddesign/2016/mar/28/google-deep-dream-art

Gombrich, E. H. (1977). Art and Illusion: A Study in the Psychology of Pictorial Representation. Phaidon.

Gombrich, E. H. (1962). *The story of art*. London: Phaidon.

Grebey, J. (2021, December 17). Everything To Remember About the Matrix Trilogy Before You See The Matrix Resurrections. TIME. https://time.com/6129306/matrix-trilogy-recap/

Hamilton, P. (2024, March 20). *What is a meme? the history and evolution of Memes explained*. Know Your Meme. https://trending.knowyourmeme.com/editorials/insights/what-is-a-meme-the-history-and-evolution-of-memes-explained

Harari, Y. (2023, November 9). *Ai could cause a "catastrophic" financial crisis, says Yuval Noah Harari*. The Guardian. https://www.theguardian.com/technology/2023/nov/09/yuval-noah-harari-artificial-intelligence-ai-cause-financial-crisis

Harry Potter AI Trend [Online Image]. (2023) Know Your Meme. https://knowyourmeme.com/memes/harry-potter-but-in-x-ai-trend

Hide the Pain Harold [Online Image]. (2015) Know Your Meme. https://knowyourmeme.com/memes/hide-the-pain-harold

Hong, J., & Curran, N. M. (2019). Artificial Intelligence, artists, and art. *ACM Transactions on Multimedia Computing, Communications, and Applications, 15*(2s), 1-16. doi:10.1145/3326337

Howarth, J. (2023, December 4). *Alarming average screen time statistics (2024)*. Exploding Topics. https://explodingtopics.com/blog/screen-time-stats

Hu, K. (2023, February 2). ChatGPT sets record for fastest-growing user base - Reuters.https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

Høgh-Olesen, H. (2019). *The Aesthetic Animal*. New York, NY: Oxford University Press.

IBM, (n.d.). *What are neural networks?* IBM - United States. https://www.ibm.com/topics/neural-networks

Junco, M. Á. (2015, April). Caricature Before Caricature An archaeology of visual humour since Prehistory. https://www.academia.edu/36382337/Caricature_Before_Caricature_An_archaeology_of_visual_humour_since_Prehistory

Kaput, M. (2022, July 27). Dall-e 2 and the future of Design. Retrieved November 2022, from https://www.marketingaiinstitute.com/blog/dall-e-2

Kikinak. (2009, March 21). *Philosoraptor*. Know Your Meme. Retrieved February 7, 2024, from https://knowyourmeme.com/memes/philosoraptor

Kissinger, H., Schmidt, E., & Huttenlocher, D. P. (2021). The Age of AI: And our human future. London: John Murray.

Kress, G. R., & Leeuwen, V. T. (2007). *Reading images: The grammar of visual design* (Second Edition). Routledge.

Know your meme. (2014, September 27). *Hide the pain Harold*. Know Your Meme. https://knowyourmeme.com/memes/hide-the-pain-harold

Know your meme. Know Your Meme. (2023). https://knowyourmeme.com/

Leeuwen, T. V., & Jewitt, C. (2001). The Handbook of Visual Analysis (First). Sage.

Leswing, K. (2023, April 28). Tech giants are pouring money into A.I. as they cut costs elsewhere. CNBC. https://www.cnbc.com/2023/04/28/tech-earnings-calls-show-mega-cap-companies-going-big-on-ai-.html

Lord, M. (2024, April 9). Barbie. Encyclopedia Britannica. https://www.britannica.-com/topic/Barbie

Marr, B. (2018, October 12). How much data do we create every day? the mind-blowing stats everyone should read. Retrieved November 09, 2022, from https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/?sh=77c45f0b60ba

Mazzone, M., & Elgammal, A. (2019). Art, Creativity, and the Potential of Artificial Intelligence. Art, Creativity, and the Potential of Artificial Intelligence.

Memes [@memes]. (2024) Memes: the universal language of laughter [internet memes] Instagram. https://www.instagram.com/memes/

Melted Videos [@meltedvideos]. (2024) Melted Videos [internet memes] Instagram. https://www.instagram.com/meltedvideos/

McGraw, P., & Warner, J. (2015). The humor code: A global search for what makes things funny. Simon & Schuster Paperbacks, an imprint of Simon & Schuster.

Morreall, J. (2020, August 20). Philosophy of humor. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/humor/

Neil deGrasse Tyson Reaction [Online Image]. (2012) Know Your Meme. https://knowyourmeme.com/memes/neil-degrasse-tyson-reaction

Neil deGrasse Tyson Reaction variations [Online Image]. (2012) Know Your Meme. https://knowyourmeme.com/photos/1206356-neil-degrasse-tyson-reaction

Olson, Rebecca. "What is Ekphrasis?" *Oregon State Guide to English Literary Terms,* 17 Sept. 2020, Oregon State University, https://liberalarts.oregonstate.edu/wlf/what-ekphrasis.

Owen. (2023, July 12). *Barbenheimer*. Know Your Meme. Retrieved April 19, 2024, from https://knowyourmeme.com/memes/barbenheimer

Owen. (2023, July 5). *Black cat zoning out*. Know Your Meme. Retrieved April 3, 2024, from https://knowyourmeme.com/memes/black-cat-zoning-out

Oxford, D (n.d.). *Meme noun - definition, pictures, pronunciation and usage notes | Oxford Advanced Learner's dictionary at Oxfordlearnersdictionaries.com*. Oxford learner Dictionaries. *https://www.oxfordlearnersdictionaries.com/definition/english/meme*

Petrosyan, A. (2024, May 7). *Internet and social media users in the world 2024*. Statista. https://www.statista.com/statistics/617136/digital-population-worldwide/

Pope in White Puffer Jacket [Online Image]. (2023) Know Your Meme. https://knowyourmeme.com/memes/pope-in-white-puffer-jacket-pope-francis-drip

Pottermore. (2019, October 9). *500 million Harry Potter books have now been sold worldwide*. Wizarding World. https://www.wizardingworld.com/news/500-million-harry-potter-books-have-now-been-sold-worldwide

Philipp. (2023, December 1). *Harry Potter but in X AI trend*. Know Your Meme. Retrieved April 10, 2024, from https://knowyourmeme.com/memes/harry-potter-but-in-x-ai-trend

Philosoraptor [Online Image]. (2009) Know Your Meme. https://knowyourmeme.com/memes/philosoraptor

Rayner, A. (2016, March 28). Can google's deep dream become an art machine? The Guardian. https://www.theguardian.com/artanddesign/2016/mar/28/google-deep-dream-art

Roose, K. (2022). An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy. The New York Times. https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html

Rose, G. J. (1992). The power of form: A psychoanalytic approach to aesthetic form. Madison: International Universities Press.

Rose, G. (2007) Visual methodologies: An introduction to researching with visual materials. Second. London: SAGE.

Rouzé, M. (2024, April 24). *J. Robert Oppenheimer*. *Encyclopedia Britannica*. https://www.britannica.com/biography/J-Robert-Oppenheimer

Sautoy, D.M. (2020) The creativity code: How ai is learning to write, paint and think. London: 4th Estate.

South America Memes [@southamericamemes]. (2024) South America Memes. [internet memes] Instagram. https://www.instagram.com/southamericamemes/

Shifman, L. (2015). *Memes in digital culture*. Massachusetts: MIT Press.

Stockton, M. D. (2023) The Power of Algorithms: How They Drive the Internet Experience. Medium. https://medium.com/@stocktonmattdavies/the-power-of-algorithms-how-they-drive-the-internet-experience-fb6753d697c2

Tyson, N. deGrasse. (2023). *Starry Messenger - Cosmic Perspectives on Civilisation*. Harper Collins.

Walker, J. A., & Chaplin, S. (2015). *Visual culture: An introduction*. Manchester: Manchester University Press.

Wilkinson, F (2023) Influencers: The modern entrepreneur. Education National Geographic. https://education.nationalgeographic.org/resource/influencers-modern-entrepreneur/

Whiddington, R. (2023, March 1). Digital Artist Refik Anadol Has Hypnotized the World With His Swirling A.I. Art Landscapes—Here Are 5 Essential Works to Know. Artnet News. https://news.artnet.com/art-world/5-refik-anadol-works-to-know-2259071

Universal Pictures Portugal. (2015, February 27). *"Ex_Machina" - trailer oficial Legendado (Portugal) HD* [Video]. YouTube. https://www.youtube.com/watch?v=54nzsdLPs9I

VICE News. (2021). *How the Far Right Use Memes to Recruit | Decade of Hate*. Retrieved from https://www.youtube.com/watch?v=gc2fIeVsucM&ab_channel=VICE.