CODEX OF THE FUTURE SERIES

THE FUTURE OF INNOVATION AND CREATIVITY

UNLIMITED THINKING . EXPONENTIAL POTENTIAL

BY MATTHEW GRIFFIN

311 Institute Global Advisory : Forecasting : Innovation : Strategy

Version 3.0 JAN 2022

ABOUT THE AUTHOR

Matthew Griffin, an award winning futurist and author of the Codex of the Future series, is described as "The Adviser behind the Advisers" and a "Young Kurzweil." Matthew is the Founder of the 311 Institute, a global Futures and Deep Futures advisory, as well as the World Futures Forum and XPotential University, two philanthropic organisations whose mission it is to solve global inequality and the world's greatest challenges.

Regularly featured in the global media, including AP, BBC, CNBC, Discovery, Forbes, Netflix, RT, ViacomCBS, and WIRED, Matthew's ability to identify, track, and explain the impacts of hundreds of exponential emerging technologies and trends on global business, culture, and society, is unparalleled.

Recognised for the past six years as one of the world's foremost futurists, innovation, and strategy experts Matthew is an international advisor and speaker who helps many of the world's most respected brands, governments, investors, and institutions, explore, envision, build, and shape the future of global business, culture, and society.

BE BOLD. MAKE FIRST CONTACT.

mgriffin@311institute.com
www.311Institute.com

PROUDLY BUILDING THE FUTURE WITH ... & MANY OTHERS

accenture
AIRBUS
Bloomberg
citi
Disney
HM Government
LEGO
Legal & General
RSA
SAMSUNG


AON


ARM


BCG


BENTLEY


centrica


COLAS


DECATHLON


Deloitte.


hp


HUAWEI


interdigital


KPMG


LinkedIn


McKinsey
& Company


Qualcomm


T-Mobile


Willis
Towers
Watson

A LETTER FROM OUR FOUNDER

MATTHEW GRIFFIN

WE LIVE in extraordinary times, in a world where individuals, organisations, and technology can impact the lives of billions of people and change the world at a speed and scale that would have been unimaginable just twenty years ago.

We also live in a world full of challenges, and a world where all too often negative news gets amplified at the expense of good news, and where tales of hope, inspiration, and positivity get drowned out and lost in the noise. It's no wonder therefore that today more people are more anxious about the future than ever before. And, arguably, a society which believes it's marching towards the darkness, rather than the light, has a poorer future than one that doesn't. Hope, however, is all around us and it's our purpose to light the way so all of us, people and planet, can prosper.

TESTIMONIALS

THANK YOU EVERYONE!

#FUTURENOW

EXTRAORDINARY!
Peter K., EMEA Managing Director
Accenture

ASTOUNDING!
Peter B., COO
Aon

AMAZING!
Abdessamad K., Head of Derivatives
Bloomberg

INSPIRING!
Jay C., CHRO
Dentons

SIMPLY GREAT!
Isaac H., Country Manager
Google

ESSENTIAL!
Christina L., Director
Her Majesty's Government, UK

BLOWN AWAY!
Nicola P., Global Procurement Director
Lego

WORLD CLASS!
Ana C., CMO
LinkedIn

EXCEPTIONAL!
Robert D., Global Strategy Director
Qualcomm

VISIONARY!
Catherine C., Managing Director
WIRED

BLOWN AWAY!
Mark R., Director
Willis Towers Watson

OUR MISSION.

FOR PEOPLE & PLANET: BUILDING A BETTER FUTURE

OUR MISSION is to be a driving force to help solve the world's greatest challenges, help organisations build sustainable and lasting legacies, and democratise access to the future so everyone everywhere, irrespective of their ability or background, can benefit from it.

We do this by surfacing essential future-focused insights and open sourcing our content, by amplifying inspiring stories and voices, and by bringing people together.

OUR BRANDS.

UNLIMITED THINKING . EXPONENTIAL POTENTIAL

OUR BRANDS compliment one another and align with our core mission, they include:

311 INSTITUTE

Our **globally renown** Futures and Deep Futures advisory working with the world's most respected brands, governments, and investors to explore, co-create, and shape the future of global business, culture, and society.

WORLD FUTURES FORUM

Our **philanthropic organisation** working with the United Nations and other world leading institutions to find solutions to the world's greatest challenges including all 17 UN SDG.

XPOTENTIAL UNIVERSITY

Our **philanthropic university** working with academia, governments, and regulators to create and deliver accessible future focused curricula and educational content for business executives and students from around the world.

“TECHNOLOGY USED TO BE THE TOOL AND HUMANS WERE THE CREATORS, IN THE FUTURE TECHNOLOGY WILL BE BOTH TOOL AND CREATOR.”

- Matthew Griffin, Founder

311 Institute
World Futures Forum
XPotential University

EXPLORE MORE

WANT EVEN MORE INSIGHTS INTO THE FUTURE AND DEEP FUTURE?

Explore More

BLOGS • BOOKS • EDUCATION • EVENTS • KEYNOTES
NEWS • PODCASTS • SHOP • VIDEOS & More ...

www.311Institute.com
www.WorldFuturesForum.com

CODEX OF THE FUTURE SERIES _FUTURE PROOF YOUR BUSINESS CODEXES:

FUTURE PROOFING your business has never been harder. So we've made it easy for you.

HOW TO BUILD EXPONENTIAL ENTERPRISES.

Exponential technologies have accelerated the global rate of disruption, now it's your turn to disrupt the quo.

CODEX OF THE FUTURE SERIES

EXPLORE MORE

LEADERSHIP LESSONS FROM ORGANISED CRIME.

Explore the techniques Organised Crime groups use to grow despite being subjected to huge "competitive" pressures from governments.

THE FUTURE OF EDUCATION AND TRAINING.

Explore how we prepare our children and society for a future where Science Fiction becomes Science Fact.

THE FUTURE OF EXPONENTIAL DISRUPTION.

Explore the forces accelerating global disruption and how close we are to disrupting industries in just a day.

CODEX OF THE FUTURE SERIES

EXPLORE MORE

THE FUTURE OF INNOVATION AND CREATIVITY.

Explore what happens when creative machines and not humans are the dominant creative force in the world.

XPOTENTIAL UNIVERSITY

UNLEASH YOUR EXPONENTIAL POTENTIAL

Explore More

FUTURES EDUCATION FOR STUDENTS AND LEADERS

www.311Institute.com

CODEX OF THE FUTURE SERIES _FUTURE TECH & TRENDS CODEXES:

WITH HUNDREDS, thousands, of emerging technologies and trends it can be hard to identify them all and understand their implications. So we put them all right at your fingertips.

311 TRENDS CODEX.

Explore the hundreds of megatrends and microtrends impacting and shaping your world and everything you care about.

CODEX OF THE FUTURE SERIES

EXPLORE MORE

EXPONENTIAL TECHNOLOGY CODEX.

Explore the hundreds of exponential technologies that are emerging in detail and learn about their implications for global culture, industry, and society.

THE EMERGING TECHNOLOGY STARBURST COLLECTION.

Use our Griffin Emerging Technology Starbursts to explore the future and find new ways to disrupt the status quo.

CODEX OF THE FUTURE SERIES

_FUTURE DEEP DIVE CODEXES:

OUR DEEP Dive Codexes explore the future of individual topics in depth. So now you're the expert.

THE FUTURE OF THE CONNECTED WORLD.

Explore the future of the technologies and trends that will connect everything and everyone and shape our connected future.

CODEX OF THE FUTURE SERIES

EXPLORE MORE

THE FUTURE OF GAMING.

Explore the future of gaming and what happens when the simulations become people's new reality.

THE FUTURE OF INSURANCE.

Explore the future of insurance, and the dangers of a future where global risk becomes systemic.

THE FUTURE OF SMARTPHONES.

Explore the future of smartphones and smartphone formats, and discover what's around the corner.

CODEX OF THE FUTURE SERIES

EXPLORE MORE

THE FUTURE OF SPORT.

Explore the technologies and trends shaping the future of sport and sports performance.

THE FUTURE OF SYNTHETIC CONTENT.

Explore the technologies and trends revolutionisning how content is made and consumed.

CONTENTS

Third Edition published Jan 2022. To request this Codex in an alternative language please contact the author.

INTRODUCTION

TODAY, AS I lay out in more detail in my Future of Exponential Disruption codex, the companion guide to this codex, we are getting closer to the day when we'll be able to realise global culture, industry, and society being disrupted in just a single day, and in part that vision is thanks to the emergence of a new breed of Creative Machines and exponential on demand manufacturing technologies that are already helping us tear up the increasingly antiquated innovation and business rule books.

While their mere existence might be a surprise to many Creative Machines that can design, innovate and then produce new digital content, as well as increasingly sophisticated physical and digital products, from cars and clothing to computer chips, drugs, and software, in real time are now on the ascendancy, and there's no going back to the old status quo.

In an increasingly connected society, where the global rate of innovation is continually accelerating the fact shouldn't be lost on you that we're nearing a tipping point - a time when companies, individuals, and machines, either in isolation to one another or in harmony, are able to design, innovate, produce, market, and then sell new products and product concepts in an internet heartbeat, and all in an increasingly automated and autonomous way.

In this first of a kind Codex we'll take a look at the exponential technologies that are already revolutionising how we create content and innovate products, and I'll walk you through examples of how companies have already leveraged these amazing new tools to accelerate their own rates of innovation by millions fold - all of which will help bring about a new creative revolution unlike anything the world has ever seen before.

Explore More

MATTHEW GRIFFIN
Founder

THE ERA OF CREATIVE MACHINES

WE CAN now see a day when Creative Machines, either autonomously or in unison with humans, will be able to design, innovate, and mass produce new digital content, and new digital and physical products, in real time and on demand at global scale. And while this reality is already here and being leveraged by companies, to some extent, given the inter-connectedness of our society it's no small leap to say that these machines will, sooner than later, revolutionise every corner of global culture, industry, and society, and forever change how we create and innovate products from A to Z - from AI software and Augmented Reality gaming arenas to Zebra burgers made from lab produced meats, and way beyond.

A NEW NARRATIVE

Machines can never be creative. And machines can never innovate or be innovators. That's been the false narrative spun by analysts and experts alike over the past few decades. It's also a narrative that, at best, has been counter balanced by analysts and experts elsewhere who have spun the narrative that posits when, not if, machines are capable or creativity and innovation then that time will be a long time coming with the average estimate being that these so called Creative Machines arrive in the year 2035.

However, as any of you who have ever watched my keynotes or read my blog know, firstly never say never, and secondly, remember that everything is accelerating exponentially. Needless to say, not only is it possible for machines to be creative and innovative, but they're already here, and ironically they're here because of our own insanely brilliant human creativity and inventiveness.

There is an important point to highlight here though. While today's Creative Machines can all innovate and produce a wide variety of different types of content and products how they innovate them isn't too dissimilar from how we humans do it, except for the fact that at this point in time their process is much more logical and data driven rather than emotional or empathetic.

As a species humans have long been of the opinion that what sets us asides from all other living things on this blue planet of ours is our seemingly unique ability to create and innovate new products and tools - including technologies. But what all the aforementioned analysts and experts seem to have forgotten is that creativity, without entering into a theological debate about the act of

Notes:

creation or the soul, is subjective, and innovation is merely a process - a process that can be broken down and replicated in algorithmic form and embedded into a machine.

INNOVATION AS A PROCESS

Today Creative Machines are only capable of iterative innovation and not Primary or Disruptive innovation, yet, so just to drive home the point that most innovation is a process, asides from all of the innovation methodologies that lay it out as a process, if I asked you to reduce the weight of a cup by half, which is a simple example of iterative innovation, in your head you're already going through a series of, albeit complex, mental steps and weighing them all against the desired objective. And here's the magic, here's how we now create our own Creative Machines.

If I can understand what those individual process steps are and their consequences and outcomes then, with some data science magic, I can convert them into an algorithm - a Creative Machine. Furthermore, the more complex the algorithms and models are the more capable these Creative machines become and the better able they are to design and innovate increasingly sophisticated products, whether it be new movies or interplanetary rovers, or new drugs or even new AI's and robots.

A CROWD OF MACHINES

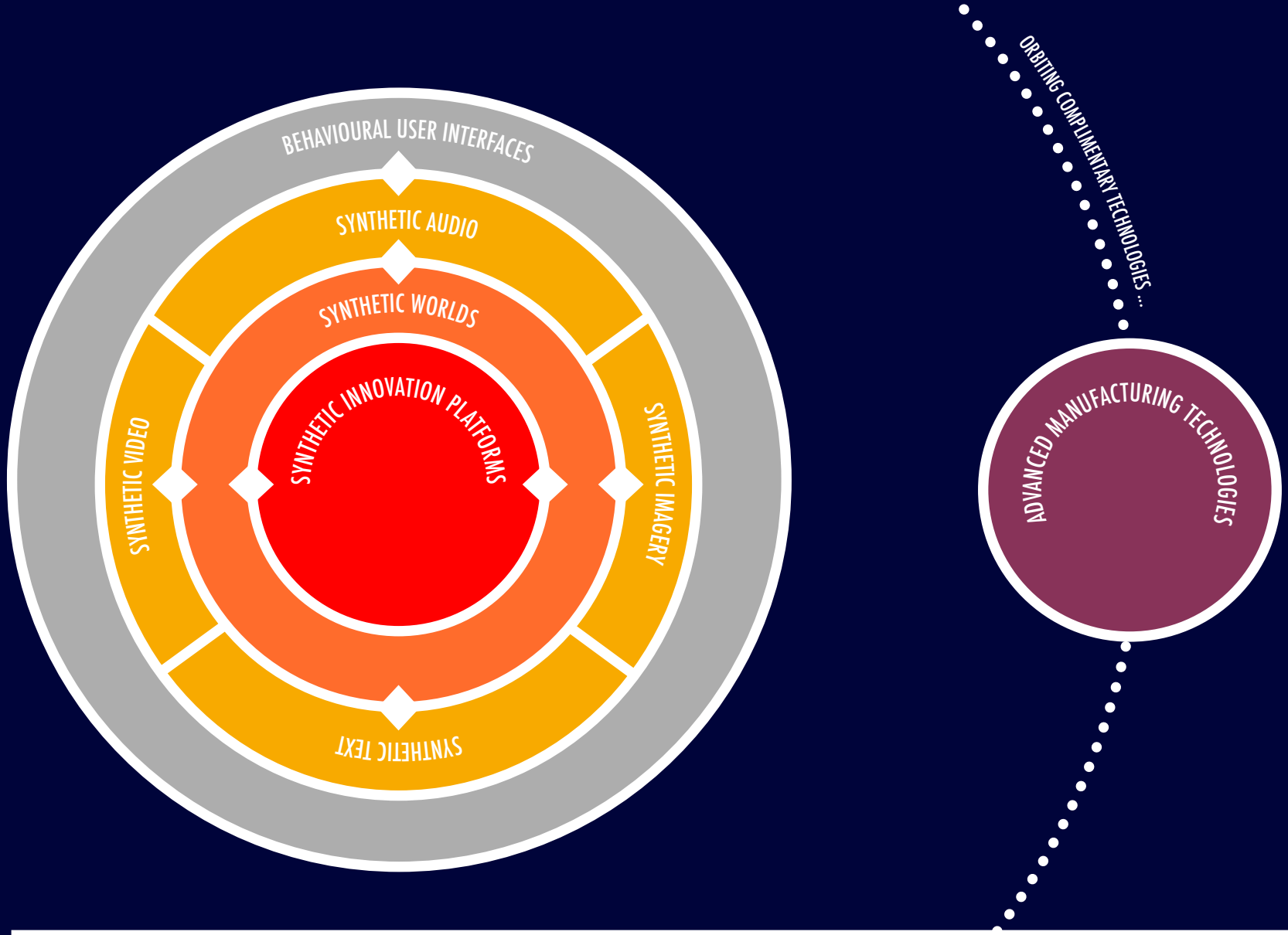
There's more to developing Creative Machines than you might imagine though because it's not simply enough for these machines to be good at or to be competent in one discipline or one skill. They have to be competent in many different disciplines and skills, and the best way to explain this is by using an example.

A Creative Machine that's been tasked with generating a piece of new Synthetic Video content will need to have a grasp of atmosphere, context, flow, human emotions, natural language, perspective, scripts, tempo, tone, and much more, as I'll discuss in later chapters.

Meanwhile a Creative Machine that's been tasked with creating a new physical product, for example, will perhaps need to have a grasp of bio-mechanics, context, design thinking, durability, function, materials science, manufacturing processes, physics, reliability, supportability, usability, and, again, much more.

Notes:

In short, just like we humans, in order to master their respective “trades” these creative machines have to be well rounded, be well versed in multiple disciplines, and then be able to intelligently combine all of these disciplines together within the context of the task in order to achieve their desired objective. And that’s what researchers around the world have been developing, and that’s what we’re going to explore in more detail in this Codex.



EVERY TECHNOLOGY ACCELERATES THE OTHER

These technologies, whether they are used to generate synthetic audio, imagery, text, video, or Synthetic Worlds, will disrupt the entire global content industry - from the production of art, books, games, and movies, to Virtual Influencers and VR. They also compliment one another and accelerate each others development, including the development of Synthetic Worlds which Synthetic Innovation platforms will use to virtually innovate, model, and test new product concepts at speed.

MULTIPLE FORCE MULTIPLIERS.

WHEN THINKING about how to harness the potential of these Creative Machines, that can arguably create and innovate anything, for maximum advantage and impact, whether you're interested in machines that can innovate new content in all it's forms, from audio, imagery, text, and video, to Synthetic Worlds, or machines that can innovate new digital and physical products you should also think about what happens when you combine those powerful exponential tools with other powerful exponential technologies because this is where you really unlock their true potential.

For example, combine these Creative Machines with 3D and 4D printers and you now have a way to innovate and produce products in real time and on demand at a rate and speed that was previously unimaginable.

Furthermore, not only does this mean you can dramatically reduce your time to market by multiples, but it could also have a dramatic impact on your business model, go to market, and open up a wealth of new opportunities you'd never dreamt of before - all this is only possible though because there have been multiple developments in multiple technology disciplines in the field that are now becoming mature enough for companies and researchers to integrate together to create increasingly capable and usable creative machines.

TECHNOLOGY COMBINATIONS

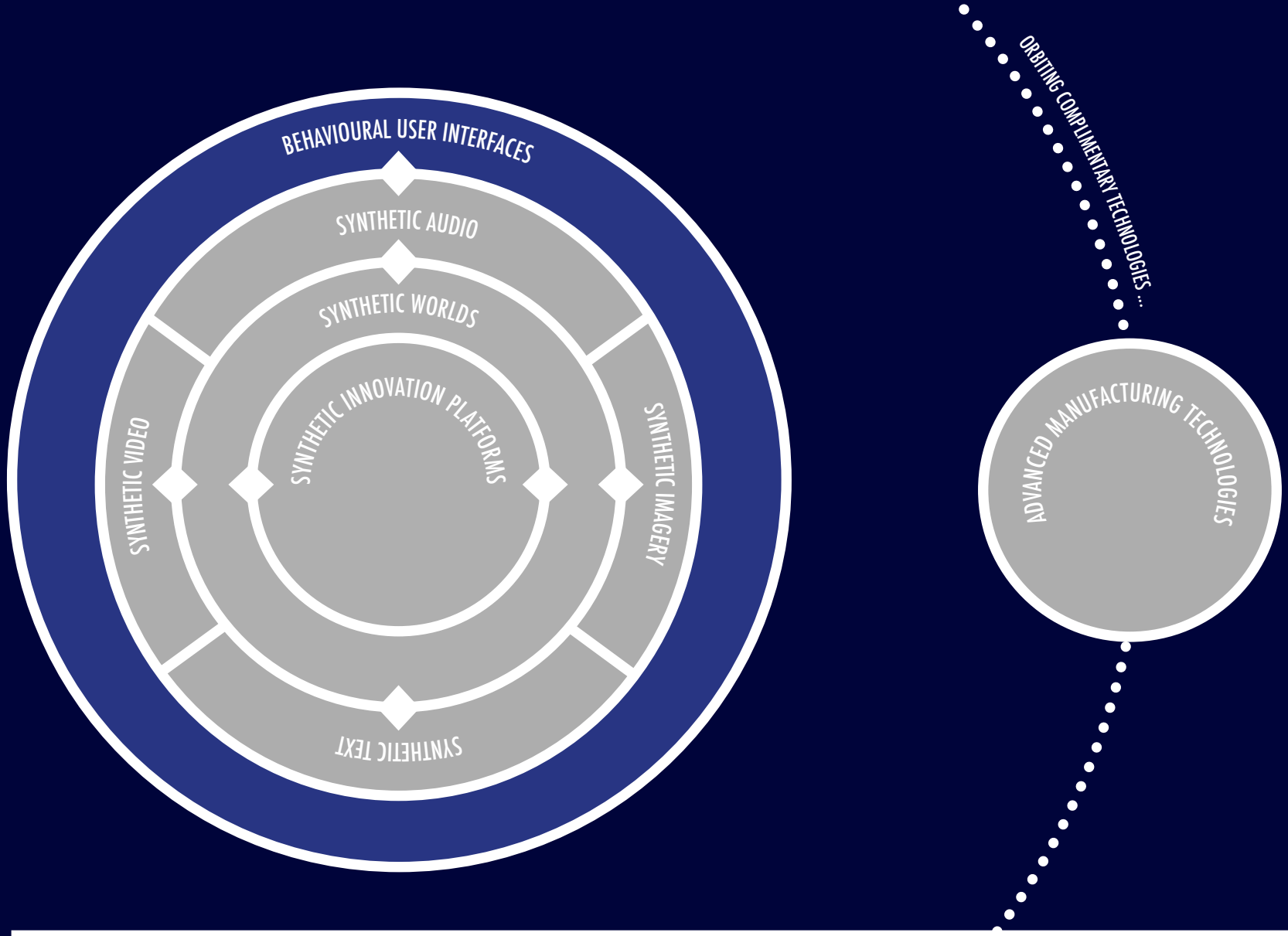
When you look at the diagram to the left you will see six distinct technology categories that each, within their own right, will transform the world and revolutionise how we create content and products. All that said though, and as impactful as they all are, when they're combined together that impact is multiplied thousands fold, if not more.

For example, Synthetic Text generators can write authentic news articles, novels, research papers, and scripts, for starters, with the obvious implications on industry and society.

When integrated into Synthetic Video generators though all of a sudden these generators, these machines, can write the script for a movie and then the Synthetic Video generator can use that input to auto-generate the corresponding movie.

Combine the same Synthetic Text generators with Synthetic Image generators and you now have a way to convert text into imagery. Combine them with Synthetic Audio generators and now you have a way to create music with lyrics and human vocals. And so we can keep going through the wheel, so try thinking about the potential of these combinations for yourself.

Furthermore, and diving further into the wormhole, by combining all these generators together we then use them to create increasingly sophisticated and realistic synthetic virtual worlds that on the one hand could become new Virtual Reality game environments, as well as even movie sets, as we recently saw with the production of Disney's Lion King, but that also become the ideal digital environments that let Synthetic Innovation platforms design, model, test, and then train and evolve their new product concepts at extreme, or "digital" speed.



UNLOCKING CREATIVITY AND INNOVATION FOR ALL

As Creative Machines become increasingly advanced behavioural user interfaces will let everyone with a smartphone and an internet connection tap into their power using just their voice, as well as a variety of other means. In short, these behavioural user interfaces will be the secret ingredient that democratise access to creativity and innovation for the billions of people on the planet who today have great ideas but have little or no way of making them a reality.

DEMOCRATISING CREATIVITY AND INNOVATION.

AS THE development of these Creative Machines, in all their variants, continues to accelerate the next step will be to simplify their user interfaces and make them increasingly accessible and easy for people to use, at which point the technology will be democratised.

EVERYONE IS A CREATOR

The impact of the democratisation of these technologies on human society, as well as industry, will be nothing short of revolutionary. The result of which will be that we will be able to create and innovate content and products not at traditional human speed, but at exponential digital speed.

In just a three year period, for example, we've already seen convincing DeepFake content, which is just one form of synthetic content, move from being a high tech lab experiment that requires a team's worth of PhD's to create to being integrated into smartphone apps that allow anyone with a camera and a face replace Hollywood A-Listers likenesses for their own.

Now take this development a step further and all of a sudden the billions of people with nothing more than a smartphone and a lousy internet connection will be able to just ask an app to design a new product for them and make modifications to it using nothing more than their voice.

Furthermore, when this technology is combined with 3D printing, for example, those products can then be manufactured anywhere on demand - time and time and time again.

Ultimately, we are now starting to bear witness to the exponential acceleration of global innovation, but it's not just constrained to content or physical products, in time it will apply to everything - from AI software and computer chips all the way through to new pharmaceutical drugs and other increasingly sophisticated products.

We are nearing a tipping point and make no mistake that the rate of global change, that we already think is fast today, is about to be kicked into overdrive.

AUTONOMOUS CREATORS AND COMPANIES

However, as fast as human creativity and innovation is at digital speed it will be nothing when compared to the rate of creativity and innovation when autonomous Creative Machines and autonomous companies, driven by machine entrepreneurs not human entrepreneurs, emerge and combine - something that is already starting.

At this point not only will the machines be able to identify problems to solve, or opportunities, and create the content or products to match, but they will then, thanks to the nature of our increasingly digital and interconnected society, be able to build and scale companies, and market and sell these goods to a global audience and at global scale at exponential speed.

MACHINES THAT INNOVATE PRODUCTS

INNOVATION IS one of humanity's most exalted skills and crowning achievements - one that, as we are almost continuously told, separates us from many of the other animals and arguably makes us the "most unique species in the known universe."

Naturally then you can imagine the fallout that the arrival of so called Creative Machines, machines that share this trait and that can autonomously design any kind of product, from a building or a computer chip, to a piece of AI software or an item of clothing, will have on human culture and society. And that's before we discuss their impact on the global business environment, the impact they'll have when they eventually exceed our own amazing talents, and then are able to innovate and produce products at digital speed, millions of times faster than humans could.

Furthermore, and as if this breakthrough alone wasn't enough, as I'll discuss in more detail later in this chapter, when these machines are combined with new advanced manufacturing technologies like those laid out in the Griffin Exponential Technology Starburst, not only will they be able to scan social media and other big data sources for problems to solve and new market opportunities and then innovate new product concepts but they'll be able manufacture and evolve them too - irrespective of whether those products are hardware or software based.

The result of all this will be that eventually they'll be able to identify opportunities, innovate and manufacture new products at such speed that they'll be able to take the Concept to Market times down to just days, and potentially, as I alluded to earlier in this Codex, disrupt global industries within just hours and days not decades or years as it takes today.

WE CREATED THE IMPOSSIBLE. AGAIN.

Ironically the arrival of these types of creative machines has only been made possible because of our own human ingenuity and inventiveness, and given the scale of the task of creating them it's little wonder that so many people never thought these types of machines would ever exist. But, here we are, and here they are.

Humans are amazing at many things - especially our ability to break large tasks down into simpler, replicable steps that can then be scaled up and developed further, and that's precisely what's happened in this instance.

Notes:

After all, as I explained in the previous chapter innovation is a process, and processes, no matter how complex they are, once understood can be replicated, and if you skipped that chapter then let's run through it again with a simple exercise because it's a very important point. Let's begin: Take a cup and make it half the weight. This is an example of what's known as Iterative Innovation where we take a product and improve it, and it's just one of several types of innovation.

In your mind you're already going through a process and weighing the results against the desired outcome. So, for example, no doubt you're wondering if you can cut the glass in half, change it's size, or the materials that it's made out of, and so on, until eventually you settle on a solution.

By understanding this process and by mapping out all of the individual steps involved we can now digitise them using machine learning to create an AI capable of mimicking it and this AI becomes our first basic Creative Machine. Then, in order to improve it, we iterate it and continue improving it until we have a machine that can solve the problem and innovate a product.

And as for creating machines that can create well designed products that meet the briefs, rather than creating products that noone wants to use or buy, well, that's an innovation problem, and we're solving that too so let's move onto the next section.

BUT IT'S NOT AS SIMPLE AS THAT.

Creating a good creative machine that's capable of innovating well designed products that meet the brief, whatever that brief is, isn't just as simple as digitising a few steps though - just think about all the different permutations and thoughts your mind had to run through just to create a lighter glass, let alone a more complex product.

Not only did you have to understand the context of the brief and the outcome we were looking for but you also had to take into account the purpose of the product and the final products usability. However, while you were running all the permutations in your head you also had to consider all of the different properties of the materials you could have used - concrete, for example, wouldn't have met the brief from a weight perspective, meanwhile paper would have met the brief, but the product would have been all but unusable. And so on.

Notes:

Innovation is anything but simple, even for a human, and it's even more difficult for a machine whose knowledge and experience is starting from scratch.

So, while developing a creative machine might sound straight forward in order to mimic even the simplest skills of the human mind researchers in the field have had to develop complex multi-disciplined AI's capable of understanding everything from context and natural language, through to materials and the laws of nature, and much more besides. Then, naturally, as they have to code in more information and "intelligence" into these machines the complexity of the models increases exponentially and that's a real challenge.

That said though we are now at the point in time where AI's capable of Iterative Innovation have arrived, and they are only going to improve and get more capable from here.

FROM ITERATIVE TO DISRUPTIVE

Being able to develop machines capable of basic, iterative innovation is obviously just the first stage of what will be a very long and protracted journey, but as the platforms develop they will in time be capable of much more including, eventually, Disruptive and Primary Innovation.

Disruptive innovation is where they are skilled enough to design new product concepts that are significantly better than anything we've ever seen before that can disrupt existing products and markets, for example, a disruptive new game, healthcare treatment, material, or vehicle, and many millions of other things besides.

Next up comes Primary Innovation - the point in time at which they become capable of ingesting data about problems to solve and designing new product concepts we've never seen before - a skill that could also help usher in new industries. An example of primary innovation could be anything from the development of interstellar space travel through to cracking the secrets of fusion and human immortality, creating a new healthcare treatment, and again many million other things besides.

As researchers, and then the creative machines themselves, as we're already seeing today, work together in tandem to develop more capable machines in time we will see them become so good at what they do and so plugged into the digital fabric of our society that they'll be able to spot problems and opportunities

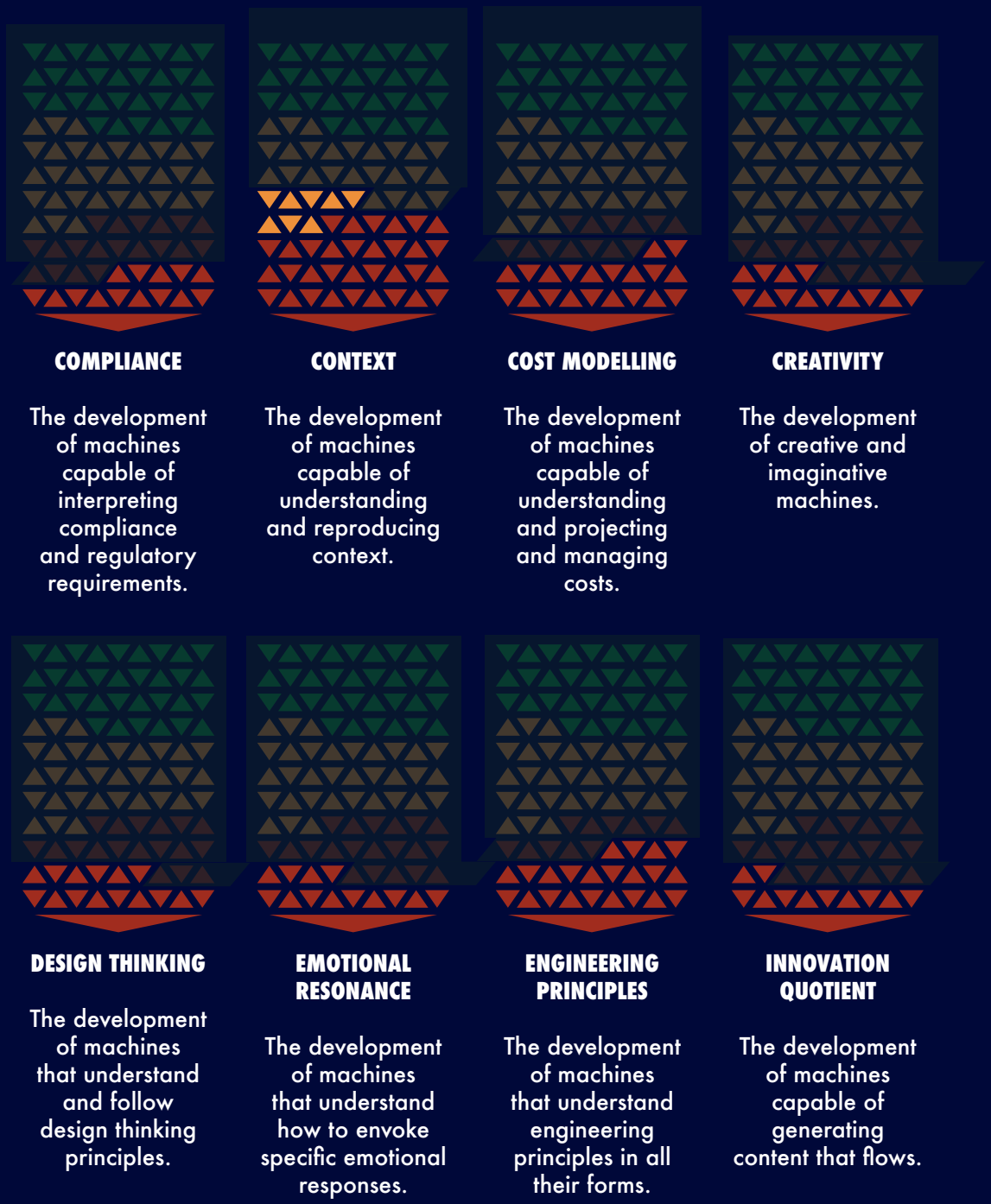


Figure 3. Current state of Creative Machine development by individual AI discipline.

Source: Multiple Sources

as they appear and design new products that solve them in real time.

For example, imagine everyone complaining about a product on a social media platform, whether it's a hardware based product or a digital one, and a Creative Machine being able to understand the context of their complaints and design a better product - something that leveraging the massive computing power and resources of AI and the cloud it could do in real time.

STILL WORK TO BE DONE

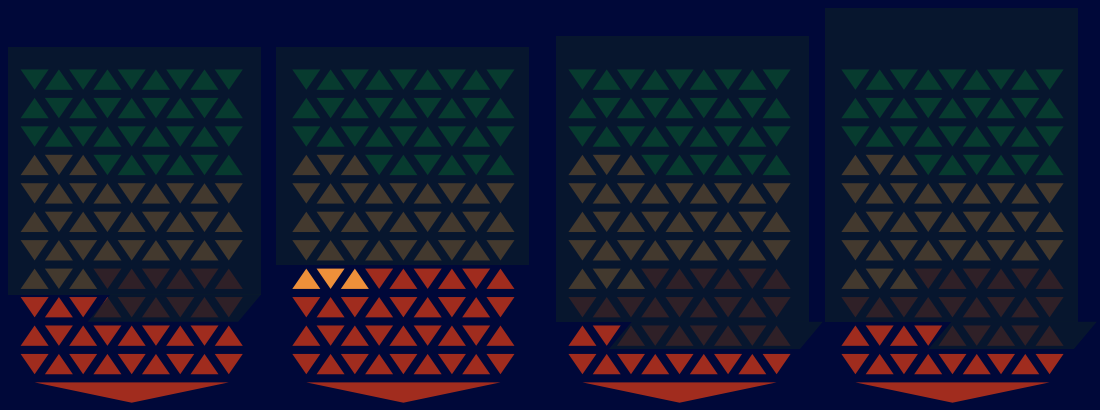
While many of us might think that building effective Creative Machines might be a herculean task in itself the truth is that it's actually harder than that, and it goes without saying that the more complex the products that these machines are tasked with designing the more complex the algorithms underpinning these Creative Machines need to be.

To the right and left of this column you'll see just a small number of AI disciplines, out of many, that all need to mature and then be combined and integrated with one another before we get close to creating what will be known as a General Creative Machine, in other words, a Creative Machine that can best

any human in almost any innovation task.

Take my glass example in the above chapters, in order to design an effective glass the machine first has to understand the context of our request, it also has to have an understanding of the purpose of the glass - which is no small feat for a machine that doesn't need to drink and doesn't have any way to physically hold or interact with a glass with in the first place. Ironically, this latter problem is where virtual modelling, in the form of Simulation Engines, which are being generated by another form of Creative Machines that I'll discuss in a later chapter in this Codex, become especially helpful during the innovation process.

Then, add into this that while it's always possible for the machine to meet the black and white brief, namely to reduce the weight of the glass, the final product it designs has to be not just functional but also has to meet other criteria as well, such as affordability, desirability, usability, and more. And, as mentioned, as the products become more complex it's therefore easy to see how the level of difficulty in creating creative machines that can meet the briefs demanded of them become exponentially more difficult.



LAWS OF NATURE

The development of machines that understand the laws of nature, including materials and physics.

MATERIALS SCIENCE

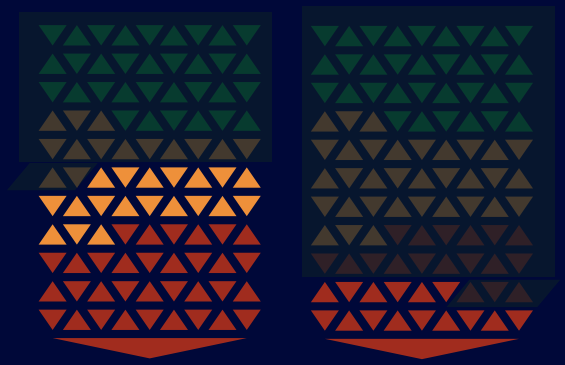
The development of machines capable of understanding material properties.

SUSTAINABILITY

The development of machines capable of understanding and meeting sustainability requirements.

SYSTEMS THINKING

The development of machines capable of systems thinking and the impact it has on the design process.



SYNTHETIC SIMULATIONS

The development of machines capable of creating realistic virtual environments to aid modelling.

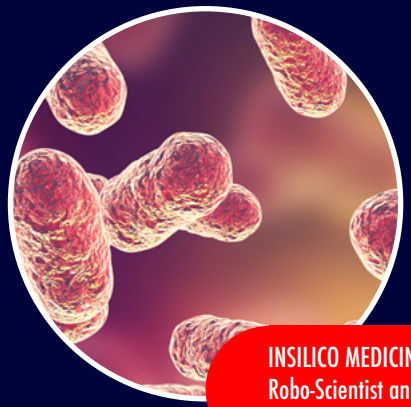
USABIILITY

The development of machines that understand the principles of usability and that can create usable products.

KEY:

▲ Each triangle represents 1%. The more triangles the more mature the discipline, until it reaches 100%.

AIRBUS are using creative machines to design and innovate new ultra-lightweight components for their A330NEO and A380 aircraft.



INSILICO MEDICINE are used an early example a Robo-Scientist and it designed 30,000 new drugs in 21 days, some of which were winners.

NASA are using creative machines to help them design new ultra-lightweight interplanetary landers and martian habitats.

UNDER ARMOUR are using creative machines to design new trainers and sports apparel which are then 3D printed in shops.

PRODUCTS DESIGNED BY CREATIVE MACHINES

As the capabilities of Creative Machines improve more and more companies are starting to experiment with them to create new product concepts, so here are some examples. And I could have also included Amazon and General Motors who are using creative machines to help them design fashion lines and cars - the floodgates are opening.

Notes:

BROADER AND BETTER

As our ability to create machines that are good at innovating specific products improves naturally over time the breadth and the complexity of the products they'll be able to innovate will also naturally improve.

Today, for example, they're already innovating everything from aircraft parts and apparel as well as cars and furniture. But, even though it's still early days we're already seeing the emergence of so called Robo-Scientists, creative machines that are designing new healthcare treatments and drugs, as well as new materials, so if you think that creative machine innovation will simply be confined to specific product categories then you'll be very wrong and at some point they'll likely cross every field.

SUMMARY

We are starting to see the emergence of a new world order and economy, one that is built and led by machines, and as humans we need to gather a point of view and understand the strategies our enterprises need to design and embrace in order to reap the significant benefits of this new reality while minimising the threats and the down sides.

MACHINES THAT INNOVATE CONTENT

IT GOES without saying that not all content is created equal. That said though every piece of content, whether it is an individual piece of content or an entire virtual arena, environment or world, has its place and its own special intrinsic value, from a pamphlet whose commercial value is relatively low and fleeting, or a blockbuster movie or AAA game whose value is substantially greater and much longer lasting.

AS RESEARCHERS get better at creating Creative Machines that generate Synthetic Content the next battleground will be creating the techniques and tools to direct, edit, manipulate, and control the outputs.

Over the past few decades there has been a paradigm shift in how the world creates and consumes content as people embrace new content formats, techniques, technologies, and tools, and move from creating and consuming mostly physical content to creating and consuming digital content instead.

It's this shift from physical to digital, combined with the emergence of increasingly democratised and powerful exponential technologies that are helping developers build the first generations of Creative Machines that are capable of generating a wide variety of so called Synthetic Content - either automatically or on command - which will change the way content is created forever.

Furthermore, asides from being able to automatically generate synthetic content on their own in time these creative machines will themselves become democratised which in turn will then help democratise content creation for everyone on the planet who has a smartphone and an internet connection. And one of the major implications of this, of course, will be that today's seeming torrent of content will end up looking like a mere trickle - especially when you consider that these creative machines will be able to generate and pump out content non-stop 24-7.

Today, we stand on the brink of a new era where these machines, that will be both the tools and the creators, will help transform every sector and corner of society in new and previously unimaginable ways - from the generation of DeepFakes that threaten to undermine democracy to putting the power to create blockbuster movies in everyone's hands - and where the rate of content creation only continues to accelerate exponentially.

Notes:

Furthermore, it should not be lost on you that as the growth in digital content continues its meteoric rise many analysts estimate that, on the one hand it will help drive over $2.7 Trillion in GDP, and that the emergence of creative machines, like the ones I'll discuss in the following chapters, will impact the careers of over 375 million people.

DEFINING SYNTHETIC CONTENT

As the way we create content continues to evolve, moving from human generated content to machine generated content, and moving from the generation of physical to digital content, one of the pedantic challenges that we face is accurately defining what we mean by the term synthetic content, so let me clear that up now - bearing in mind of course that today there still isn't a formal definition of the term.

In some people's minds digital content is already synthetic content, irrespective of who or what creates it, but for the purposes of this codex when I refer to synthetic content what I am specifically referring to is the digital content that is synthesised by machines.

THE NEW ERA OF SYNTHETIC CONTENT

As creative machines get more adept at generating synthetic content from scratch, whether it's in the form of synthetic adverts, articles, books, documentaries, imagery, movies, soundtracks, or video, they will bring about a new content revolution unlike anything we've ever witnessed before where the variety, velocity, and volume of the content being produced will be unfathomable by modern standards.

That said though, in order for future creative machines to conquer the proverbial content mountain and become the primary way that all digital content is generated - whether it's generated automatically or with some level of human involvement - not only will all of the individual content types that these machines need to draw on need to be mature, for example, synthetic audio, imagery, text, and video, but furthermore they will need to be able to combine them together accurately, and then also - as if the above isn't enough - draw on a whole variety of skills that are completely alien to them, including contextual, emotional, and semantic understanding, flow and rhythm, storytelling, and much more. And it's the combining together of all of these individual content elements, within the bounds of a storyline, that will be researchers greatest challenge.

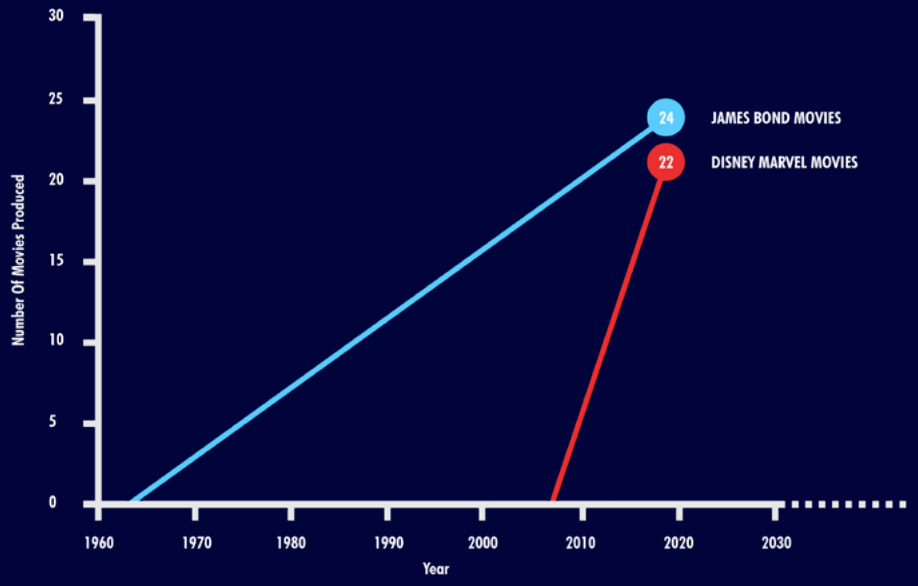


Figure 1. The rate of movie production.

Source: Visual Capitalist

While it can be said that there are creative and technological differences between the two studios when it comes to the production of movies the impact that new technologies, techniques, and tools, have on the speed of movie production cannot be denied.

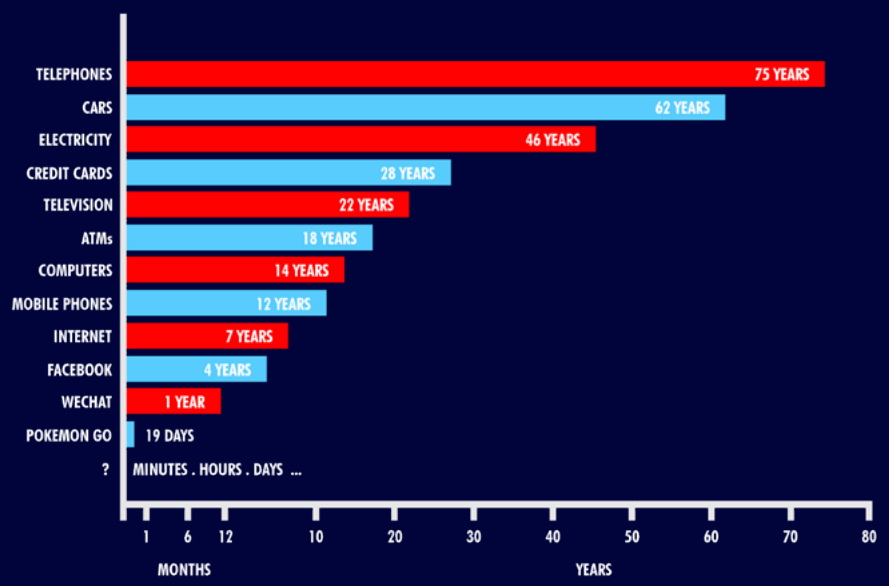


Figure 2. The rate of content and technology adoption.

Source: WSJ

However the data is presented it is undeniable that, thanks to the collision of different trends, including a more connected and digital society, that the pace of product and service adoption, in all its forms, has accelerated dramatically.

DIGITAL TECHNOLOGIES AND AN INCREASINGLY CONNECTED SOCIETY ACCELERATES DISRUPTION.

GOOGLE used synthetic audio generators to create Duplex, an AI assistant whose voice blew past uncanny valley.

NVIDIA used synthetic simulation generators to create the world's first computer game intro made by an AI.

OPENAI used synthetic text generators to create the world's first procedural text based games, which one day will include graphics.

SPRINGER used a synthetic text generator to create the world's first research book written completely by an AI.

SYNTHETIC CONTENT GENERATED BY CREATIVE MACHINES

As the capabilities of Creative Machines improve more and more companies and individuals will use them to push the boundaries of synthetic content creation and use them to generate everything from simple adverts and art, all the way through to using them to generate books, games, movies, and even digital humans.

Notes:

However, once researchers manage to create the first creative machines adept enough at combining all of these individual elements together to create realistic, high quality content - whether that content is 2D or 3D, static or dynamic - then, and only then, will the balance of creativity and creation finally and fully tip into the machines favour.

AND THEN COMES PROCEDURAL CONTENT

When researchers finally manage to crack the code of developing creative machines that can produce content that rivals the content put together by today's human professionals, especially video content, then fortunately they'll still have work to keep them occupied. And this is where now they start ratcheting up their pursuit of creating what's known as Procedural Content, and it's arguably one of the most exciting content categories.

For those of you who haven't heard of procedural content before it's the ability of machines to take in a variety of different data sources, which I'll come to in a moment, and then automatically and in real time adapt the content they're generating accordingly.

So let's run through an example, and, just in case you are wondering everything I'm going to run through in this example is based in real technologies that we're seeing emerge today.

It's Friday night and you're sitting by yourself on your favourite sofa - the one that you've sat in so many times before that it's now permanently moulded to the shape of your body. You finally flick on a movie from your favourite provider, and while you're sitting on your sofa the Creative Machine platform that's running in that providers cloud environment is ingesting data about all of your behaviours, emotions, mood, and preferences, from multiple sources - everything from your Facebook profile and so called "Big Data" footprint, all the way through to the UHD video feed from the camera on your Smart TV that's capturing all of your biometric information and micro-facial movements.

All the time the movie's playing the platform's analysing huge volumes of information in real time and after a while it notices you're looking bored - your attention's waning, you're facial expression's changed, and your heart rate's slowed. As a result of this analysis it makes the decision to dynamically re-write the movie and generate new more exciting content in order to get

Notes:

your attention back, and now rather than showing you the stock version of the movie that everyone else is watching elsewhere it's creating a perfectly tailored movie, just for you.

That is what we mean when we discuss procedural content - the ability of a creative machine to dynamically generate new content and storylines on the fly with no human involvement.

THE ENDGAME

The ultimate endgame of all this work and research, of course, will be to develop creative machines capable of generating human grade, and better than human grade synthetic content in real time without the need for any human involvement or intervention - a Master Creator, the impact of which will be revolutionary, in every sense of the word.

SUMMARY

Today, we are creating a whole new range of Creative Machines that are capable of generating increasingly high quality and sophisticated synthetic content in a wide variety of forms and formats. This is the first battle ground for developers to conquer. And when they do they will then swiftly move on to the next task of creating the technologies, techniques and tools, needed to control, edit, and modify all of the various outputs.

However, like any good story the story doesn't end there, because once we have mastered creating these sophisticated machines it is inevitable that many of those same researchers will then turn their attention to integrating the technologies and creating the machines that can edit and modify content in real time, on the fly, in response to consumers behaviours - at which point the era of Procedural Content will have arrived.

Notes:

UNCANNY VALLEY

Notes:

ONE OF, if not the greatest challenge, that researchers developing synthetic content generators face is that of being able to get their creations to create content that passes the famous Uncanny Valley test - the hypothesised relationship between the degree of an object's resemblance to a human being and the emotional response it generates. In short, the machine's ability to generate content that is good enough into fooling people that it's the real deal.

ACHIEVING IMPERFECTION

Ironically, however, in the past most researchers in the space did their best to develop creative machines that generated flawless output, but as they quickly found out, even though the outputs were world class the critics were still able to quickly identify the fakes because they were, frankly, too perfect. And life, especially humans, aren't perfect.

As a result one of the greatest turning points in generating passable content, asides from the amazing advances in the technologies, tools and techniques, was a shift in researchers attitude and their sudden determination to introduce subtle imperfections into their work.

One of the best examples of imperfection at work is human speech. For example, while early synthetic speech models were designed to be perfect in both their flow, pronunciation, tone, rhythm, and vocabulary, the fact of the matter is that no human alive today speaks perfectly. Our speech is imperfect and our conversations are filled with stutters, slightly random tone and rhythm changes, pauses, and Umm's and Ahh's.

As a case in point when Google first demonstrated Duplex, now part of Google Assistant, on stage at their I/O conference it was the introduction of these imperfections that amazed and fooled the people listening in the audience into thinking they were listening to a real person booking appointments. And synthetic voice isn't the only type of content that is now benefiting from this lack of perfection, we're seeing it emerge in imagery and other content.

All that said though there are naturally times when you want your content to lack imperfections, such as music performances, and other types of content, and now that researchers have figured out how to include imperfections in their work their next challenge is to be able to turn it on and off, like a switch, at will, and embed enough intelligence into their creative machines so that hose

Notes:

machines understand when, and when not to go off the grid.

Notes:

SYNTHETIC AUDIO

EXPLORE MORE. Click or scan me to learn more about this emerging tech.

Notes:

OVER THE years there have been many attempts to accurately synthesise the nuances of different sounds as well as the human voice, with plenty of enterprises finding it difficult enough to create high fidelity creative machines capable of realistically synthesising just a single voice - let alone systems that are accurate and responsive enough to fool people into thinking they are in a genuine moment listening to a wide range of real audio streams.

However,as this particular field becomes table stakes rather than just a nice to have, and as we continue to see the rise and adoption of Behavioural Computing platforms, including Digital Personal Assistants, enterprises around the world have significantly stepped up their efforts to conquer it. And, supported by increasingly capable and powerful exponential technologies they're now making significant headway on a wide range of fronts.

One of the most significant breakthroughs in this space in recent years has been enterprises experimentation with Deep Learning and Wavenets which, after a significant amount of testing and development, have now largely helped researchers in the field conquer one of the fields greatest challenges - dynamically synthesising raw audio's extremely fast tick rate which, for human speech, for example, is over 16,000 samples per second and where each sample is influenced by all of the previous ones. And now that some of these foundational technologies and tools continue are maturing it becomes easier to adapt them to synthesise an increasingly wide range of sounds, from speech to music and beyond.

However, as we look further out, it could be argued that this challenge will pale in comparison to creating the world's first truly intelligent so called Conversational AI's that not only synthesise realistic human-like audio but will also be able to have expansive and wide ranging conversations with us.

USE CASES AND TRENDS ...

KEY USE CASES AND TRENDS

There's no question that as soon as the various synthetic audio generators in development reach good enough quality, let alone mature, that they will instantly have access to a global market worth hundreds of billions of dollars. And that's just the exclusive audio market.

When we factor in other content markets where audio is a component of the finished product, whether it's computer games and virtual gigs, movies or vlogs, that figure will swell by multiples.

TECHNOLOGY USE CASES

To be blunt, and from a futures perspective, if you can hear or listen to something then it's a use case, whether it's using a synthetic audio generator to create the perfect car sound, or using it to generate the sounds of animals in a field.

From a developer and market perspective though, let's face it, we're more interested in the use cases with the biggest market opportunities and returns.

That said though we can split them into two main categories.

Firstly we have the markets where audio is the exclusive content type. And secondly we have the markets where audio is a complimentary component of the overall finished product:

Exclusive Use Cases

- Accessibility
- Audiobooks and narration
- Computer-Human interfaces
- Conversational interfaces
- Decision and debating platforms
- Digital Assistants
- Educational content
- Foley and SFX
- Music
- Noise cancellation
- Podcasts
- Post editing
- Robo advisory
- Social audio
- Sound and voice cloning
- Stock audio
- Universal translators
- Voice coaching
- Voice overs

Complimentary Use Cases

- Gaming
- Movies
- Resurrection technology
- Video content
- Virtual beings
- Virtual simulations
- Vlogging

TRENDS BEHIND THE TREND

The impact of synthetic audio on global culture, industry, and society, will be nothing short of revolutionary and will include benign and nefarious use cases - both of which will have wide ranging implications that will need to be discussed and mitigated accordingly when the time is right.

It will also bring about everything from the development of new interesting use cases such as being able to advise and converse with your co-workers and family from beyond the grave, while at the same

Notes:

time leading to mass job losses across a wide range of career categories that will include everything from call center operators, sales people, and translators, through to influencers, musicians, and voice over artists.

From a trajectory perspective, as with all of the synthetic content generators we discuss in this codex, we will naturally continue to see the capabilities and scope of the technology mature, and then in time see it integrated with other complimentary synthetic generators at which point its commercial appeal and mass market adoption will accelerate again until we reach the end game of finally realising Creative Machines that are capable of generating tailored procedural content on the fly in real time.

FUTURE HISTORY ...

THE FUTURE HISTORY OF SYNTHETIC AUDIO

2018: A revolutionary new WaveNet AI emerges capable of synthesising human voices with uncanny accuracy. High fidelity Text to Audio generators emerge.

2020: The first Conversational AI platforms emerge, synthetic audio across all domains improves, and commercialisation starts accelerating.

2015 — 2020 COMMERCIALISATION

2017: The first basic but passable synthetic AI voice generator emerges.

2019: A synthetic voice passes uncanny valley, more passable synthetic music and voice generators emerge, and an AI gets a commercial record deal. The future of "online audio" will be AI.

2021: Global COVID-19 pandemic accelerates the commercial adoption of Synthetic Audio and other complimentary technologies to fight disinformation. Development and investment get supercharged.

GAME CHANGERS TO BE AWARE OF:

EVERY FIELD we discuss in this Codex is moving FAST. It is important therefore to remember that in the main Creative Machines will one day be: Autonomous, Creative, Democratised, Free, Generalised, Imaginative, Integrated, Massively Parallel, Procedural, Scalable, Self-Evolving, and Self-Improving. For starters ...

2025: AI can perfectly generate most types of audio content including accents, emotions, and voices in most languages, and do it procedurally. Tech integration accelerates, regulators catch up, commoditisation begins.

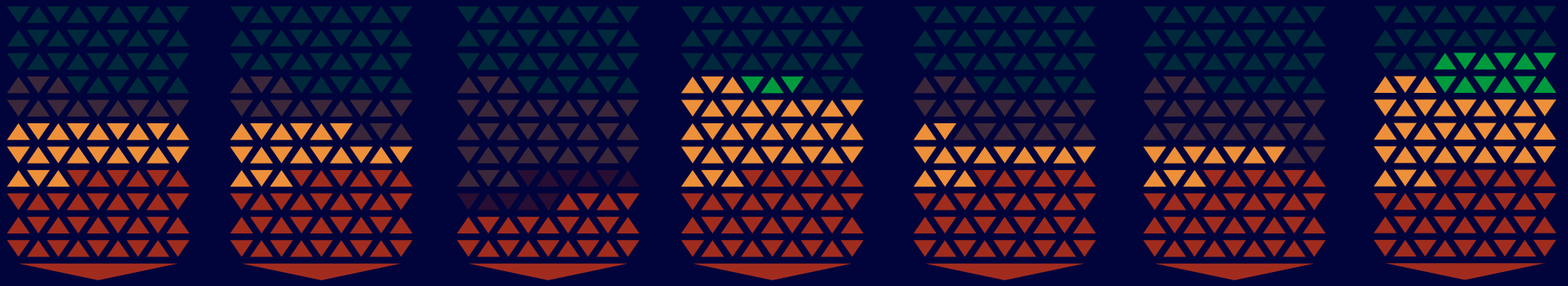
2025 COMMODITISATION & REGULATION

2030 SOCIETAL NORM

2030: AI has conquered almost all audio domains, including spatial audio, only niche domains remain. Automated generalised Synthetic Audio generators emerge, tech and workflow integrations deepen.

2035: Synthetic audio, in all its forms, is the accepted norm. Companies have large IP and "Virtual artist" holdings, as well as commercial propositions for all major content domains.

2035



CONTEXT

The development of machines capable of understanding and reproducing context.

DIALOGUE

The development of machines capable of understanding and generating dialogue.

CREATIVITY

The development of creative and imaginative machines.

DIALECTS

The development of machines that can understand and talk in multiple dialects.

EMOTIONAL RESONANCE

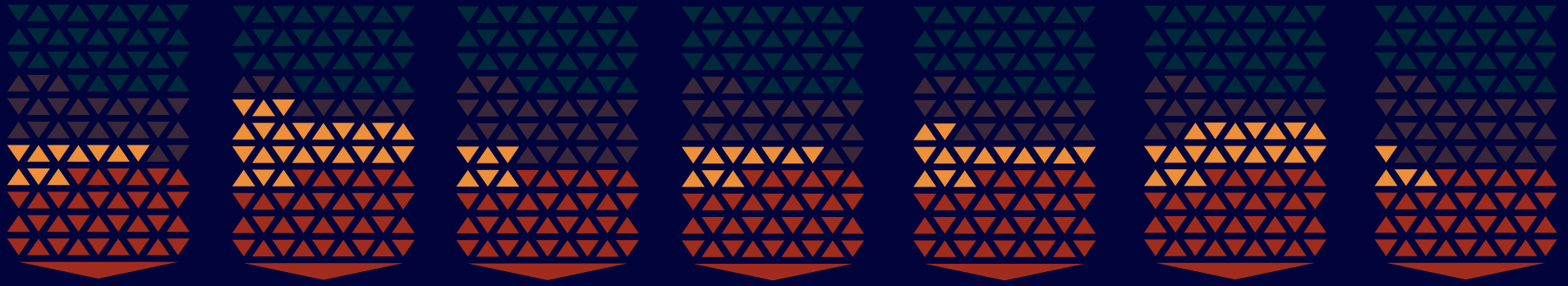
The development of machines that understand how to envoke specific emotional responses.

INTONATION

The development of machines capable of understanding and copying human intonation patterns.

NLP

The development of machines capable of understanding and generating natural language.



PROCEDURAL FLOW

The development of machines capable of generating content that flows.

SEMANTICS

The development of machines that understand implication, reference, and sense.

STORYTELLING

The development of machines capable of creating and telling great stories.

TEMPO

The development of machines capable of understanding and replicating the impact of tempo.

TONAL RANGE

The development of machines capable of understanding and replicating the impact of tonal changes.

TONE

The development of machines capable of understanding and replicating the impact of tones.

VALENCE

The development of machines capable of understanding and replicating specific emotional impacts.

Figure 1. The current state of Synthetic Audio development by AI discipline. Source: Multiple Sources

CREATING THE PERFECT GENERATOR

While many of us might think that creating a machine that's capable of spitting out decent audio is easy, whatever form that audio takes, you can clearly see from the disciplines highlighted on the previous page that it isn't.

Ultimately, reaching the point where developers are finally able to create a single universal and generalised synthetic audio generator that can be used across a wide variety of use cases relies on developments not in any one single area, but many. And that's a tough ask by any measure.

For example, while there are an increasing number of creative machines capable of generating synthetic music, with a couple of them even being signed by major record labels now, trying to create music that people are willing to pay for, music that strikes an emotional chord and pulls at the heart strings, relies on a wide range of development areas including, but not limited to, the machine's ability to understand context, emotion, and semantic meaning, and then, at the least, be able to translate those into compositions that have the appropriate flow, rhythm, and tonal qualities.

Even though there have been significant breakthroughs in the field over the past couple of years there is still clearly a lot to be done before we get to the point where we have the ultimate creative machine in this field - a machine, for example, where you can simply push a metaphorical button and have it churn out chart topping hit after chart topping hit. Which then brings me on to another point.

As we look at all of the synthetic audio platforms being developed today they are all being developed in what's called a "narrow" way. In other words, a debating machine will one day be good enough to hold its own against the world's top human debating teams, but ask that same machine to create a podcast or book a calendar appointment and it'll fall over faster than a one legged flamingo in a hurricane.

However, as time progresses all of the capabilities of these narrow AI systems will be combined beneath one umbrella that looks something akin to an Artificial General Intelligence (AGI) architecture that means that eventually one machine will be able to do it all, whether that's generating chart topping hits and catchy marketing content, or helping you converse in Mandarin during a meeting while at the same time generating the soundtrack to a new movie.

KEY: One triangle represents 1%. The more triangles the more mature the AI discipline, and when it reaches GREEN te technology is Good Enough for adoption.

TECHNOLOGY DEVELOPMENT AND DIRECTION

Over the past few years the number of companies focused on developing single use and multi-use synthetic audio generators has steadily increased. And, as predicted, that rate is accelerating as we see the first signs of commercialisation.

The early years of research in the field, which was mostly conducted on the fringes and in university labs, included alot of experimentation as the different teams tried to find the most commercially and technologically effective and efficient ways to generate high quality synthetic audio.

As the technology continues to break out into the wild and move from edgy lab experiment to serious venture it's clear to see that the companies at the forefront of the space have learned from their experiments and have developed generators that, for the very first time, have been able to carve out markets and stand up to real world customer and investor scrutiny. And while this doesn't mean that the technology has reached a tipping point in adoption we can now say with certainty that it's just simply a matter of time before we see that happen.

Furthermore, as the foundational technologies and tools these companies rely on continue to improve and mature, whether it's improvements in Artificial Intelligence modelling and training techniques, compute, user interfaces, or a myriad of other complimentary domains, we will continue to see the volume and variety of developments in the space accelerate.

Currently the primary areas of development and research in this field include:

- Conversational AI
- Emotive audio
- Music generation
- Multi-lingual audio
- Natural Language Processing
- Speech synthesis
- Text to Audio generation
- Voice cloning

AUDIO MIMICRY

In a world where we can increasingly clone everything from livestock and pets it's only natural that somewhere there are researchers that want to clone people's voices – with all of its obvious pros and cons. Over the past couple of years there has been a tremendous amount of focused development in this space, much of it focused on one goal – to use AI to create the most realistic artificial voices in the world, whether those voices are generated from scratch, as I'll discuss in the Spoken Word section below, or, in this case, to mimic the voices of real human beings. And the latter is hard, but for the enterprises that can pull it off they'll have a system that can get anyone to say anything.

Every individuals voice has its own patterns and nuances, and sampling someone's voice can now be done in as little as a minute from public video footage, after which a mixture of different deep learning algorithms disassemble the audio into its constituent components and analyse it based on accent, cadence, pitch, speed and more. And as these algorithms sort they learn, building up their own models.

Notes:

In many cases developing these AI's, most of which are built in PyTorch and TensorFlow frameworks and trained on petaflop scale computers in the cloud, and work by comparing every aspect of a recording to every other voice recording in their databases in order to build a digital profile with a few million data points required to build each voice – voices that are now starting to pass the uncanny valley test and even fool people's own families.

MUSIC GENERATION

Generative Adversarial Networks (GANs) are considered a state of the art method for generating everything from synthetic imagery to video. However, while they are a must have in those fields researchers focusing on generating other types of content, such as music, have struggled to apply them, and up until recently their only recourse was to try to adapt the many Autoregressive (AR) AI models such as WaveNets and Transformers that dominate the field because of their ability to predict single sample sounds – something that gives them a distinct edge when it comes to synthesising voices, but something's that makes them painfully serial and slow when generating music. The result of all this is, of course, that traditionally researchers in the field ended up having to augment their research with techniques including distillation or specialized kernels.

Now though researchers, fed up of this painfully slow approach are turning their attention to developing new types of GANs that are capable of generating entire music sequences in parallel, and synthesising audio significantly faster than real time - with some speeds exceeding 50,000 times faster than standard WaveNets.

Unlike traditional WaveNet autoencoders that rely on using a time-distributed latent code these new GANs generate the entire audio clip from a single latent vector, or trait, which makes it a lot easier for them to disentangle global features such as pitch and timbre from all of the other noise. Then, using datasets of musical instrument notes, such as the NSynth dataset, researchers have been able to independently control pitch and timbre using Progressive GAN architectures and upsample the tracks by using concolution generators and combining them with other single vector sounds in order to generate the full music track.

While the results that some of the researchers are pushing out now are

Notes:

impressive, with some music models even being signed by companies including Time Warner there is still a lot of work to do with some researchers thinking that their work might be further improved still by the development of multi-scale GAN conditioning models that can handle variable length outputs, and by replacing upsampling convolution generators with more flexible differentiable synthesisers.

SPOKEN WORD

Trying to create a synthetic speech platform that passes the uncanny valley test is complex, if for no other reason that every one second of human speech contains over 16,000 so called waveforms, and when it comes to developing state of the art platforms that can generate speech from scratch, complete with accents and emotions, one of the leading models in this field is the WaveNet model first developed by Google DeepMind whose second iteration, just a year later was over 1,000 times faster than the original and capable of generating far higher quality audio. What set this second model asides from the first was the development of a new Probability Density Distillation (PDD) technique that allowed it to work in a massively parallel computing environment.

The original WaveNet model used autoregressive connections to synthesise the waveform one sample at a time, with each new sample conditioned on the previous samples, and while this produced high-quality audio with up to 24,000 samples per second its sequential generation is too slow for production environments.

To get around this problem the researchers needed a solution that could generate long sequences of samples all at once with no loss of quality so they developed the PDD technique, where a fully trained WaveNet model taught a smaller “student” model that was better suited to modern parallel computational hardware. This student model is a smaller dilated Convolutional Neural Network (CNN) that was similar to the original WaveNet, but, crucially, the generation of each sample doesn’t depend on any of the previously generated samples meaning the researchers could generate the first and last word, and everything in between, at the same time.

During training, the student network starts off in a random state. It is fed random white noise as an input and is tasked with producing a continuous audio waveform as output. The generated waveform is then fed to the trained

Notes:

WaveNet model, which scores each sample, giving the student a signal to understand how far away it is from the teacher network's desired output, and over time, the student network can be tuned - via backpropagation - to learn what sounds it should produce. Put another way, both the teacher and the student output a probability distribution for the value of each audio sample, and the goal of the training is to minimise the so called Kullback Leibler Divergence between the teacher's distribution and the student's distribution.

Ironically, this new technique has parallels to the set up for Generative Adversarial Networks (GANs), that are used to create other forms of synthetic content, with the student playing the role of generator and the teacher as the discriminator. However, unlike GANs, the student's aim is not to "fool" the teacher but to cooperate and try to match the teacher's performance.

SYNTHETIC CHARACTERS

"Transmedia characters that can jump between metaverses will increase the future addressable market opportunity exponentially."

"[Digital Humans] will re-create human interaction at infinite scale and forever change our relationship with technology."

THE RULES of everything, from the rules of engagement and entertainment, to the rules of commerce itself, are about to change in ways that even today to many are unimaginable, and synthetic characters will be the driving force.

Ever since humans first began walking the savannah's of Africa we've been biological entities. But significant technological advances over the past few years are recasting humans in digital form as virtual beings - complete with their very own neural network brains and errant behaviours and personalities. And it's not just humans that will be digitised en masse.

Spurred by people's increasing willingness to engage with digital realms in all their forms and formats in time we'll see the same technologies used to create the first primal human-like virtual beings adapted and used to develop a whole host of synthetic characters whose forms and attributes will only limited by the imagination of the people and creative machines making them.

Then, as for the market opportunities, from the digital hangouts we all know and love, to through to the augmented, mixed, and virtual reality metaverses, it's difficult to see how synthetic characters won't come to dominate them all in much the same way that biological beings came to rule the Earth.

Furthermore though, unlike their biological "celebrity" equivalents the ability of synthetic characters to infinitely replicate and multiply themselves in the digital realm means that they present the companies who developed and own them with previously unimaginable marketing and revenue opportunities, especially when you consider the fact that, for example, not only could they appear in a million places all at once doing a million different things, but also the fact that thanks to technologies like Blockchain they could jump and proliferate between all of the distinctly different metaverses.

To a synthetic character the common place human phrase "You can't be everywhere at once" is just an arcane saying that has no meaning.

USE CASES AND TRENDS ...

KEY USE CASES AND TRENDS

As we continue to see the development of increasingly capable and sophisticated synthetic characters in time there won't be a digital domain or realm they won't dominate. Then, given the fact that, as we previously discussed, they'll be able to be literally everywhere at once performing all manner of tasks it's highly likely that in time the total market opportunity for synthetic characters will be measured in the trillions of dollars.

TECHNOLOGY USE CASES

Metaphorically speaking synthetic characters will ultimately bring life to all of the other synthetic content elements we've discussed, and there's no doubt that in time new use cases will be innovated on top of them. After all, when it comes to digital the only limits are our imagination.

Firstly we have the markets where synthetic characters would be the primary content type, and then secondly we have the markets where they would be a complimentary component of the overall finished product:

Exclusive Use Cases

- Coaching, Teaching, and Training
- Concierge services
- Computer-Human interfaces
- Corporate communications
- Corporate entities
- Customer engagement
- Customer service agents
- Digital managers
- Digital twins
- Healthcare agents
- Marketing
- Personal assistants
- Personal avatars
- Personalised services
- Robo advisors
- Synthetic brand ambassadors
- Synthetic celebrities
- Synthetic influencers
- Synthetic newscasters
- Teaching and training

Complimentary Use Cases

- Game entities
- Multi-character content
- Multiverse creation
- Product modelling
- Transmedia opportunities

TRENDS BEHIND THE TREND

As the cost-performance of the individual technologies that support the development of synthetic characters, such as Artificial Intelligence, compute infrastructure, Machine Vision, NLP, rendering technologies, and sensors, continue to improve it won't be too long before the vast majority of synthetic characters pass uncanny valley and are able to interact with people in a life-like way.

At this point not only will the market start to take off in earnest, but there will be serious ethical and moral questions to be answered especially, for example, when it comes to the role of digital humans and their impact on the human workforce and on hard fought corporate diversity and inclusion strategies.

In the short term most digital humans will be limited to upper bodies representations and only offer, by human standards at least, a limited range of

Notes:

interaction, motion, and speech. As we accelerate the timeline out, however, we developers will in time expand upon these foundations to create full body digital humans and synths that offer a much more life-like full body full human experience.

Furthermore, while they are mastering that art, we will see an uptick in the number of developers creating personalisation options for their creations which will see users able to switch a whole variety of different elements in and out including, but not limited to, changing the synths accents, behaviours, emotions, environment, ethnicity, language, physical characteristics, regionalisation, and much more. And then, in time, as the companies behind the synths mature, and as the technologies they use to create them matures, we will naturally see them expand into other market areas and use their talents to create an increasingly wide variety of synthetic characters both human and non-human.

FUTURE HISTORY ...

THE FUTURE HISTORY OF SYNTHETIC CHARACTERS

2016: The first Digital Human startups emerge from stealth,developers focus their attention on embodying their creations with neural network brains and improving speech synthesis.

2018: Digital Humans near uncanny valley in basic use cases but lots of flaws and tells remain, researchers focus on the quality of facial expressions and 3D movements. Overall pre-prod process is CG heavy.

2020: Digital Humans close to uncanny valley, but avatars still have some flaws and tells, fluid motion is still a challenge. Motion capture, Mo Cap, and CG still dominant in pre-prod, speech synthesis is improving but has far to go.

2022: Digital Character generation and personalisation becomes increasingly democratised as GUI's become increasingly accessible and feature rich. Development of interactive environments accelerates.

2015 COMMERCIALISATION 2020

2017: Developers focus their attention on improving the conversational accuracy of their characters, render quality improves but the modelling process is CG heavy and reliant on skilled experts.

2019: Researchers expand the scope of their work to include full body synthetic characters, but most motions are limited and basic. Using VR and game engines in the creative process accelerates.

2021: Digital Humans evolve to incorporate more upper body movement. Level of automation increases as developers, still reliant on Mo Cap and CG try to use AI to create life-like movements and reactions from 2D models.

GAME CHANGERS TO BE AWARE OF:

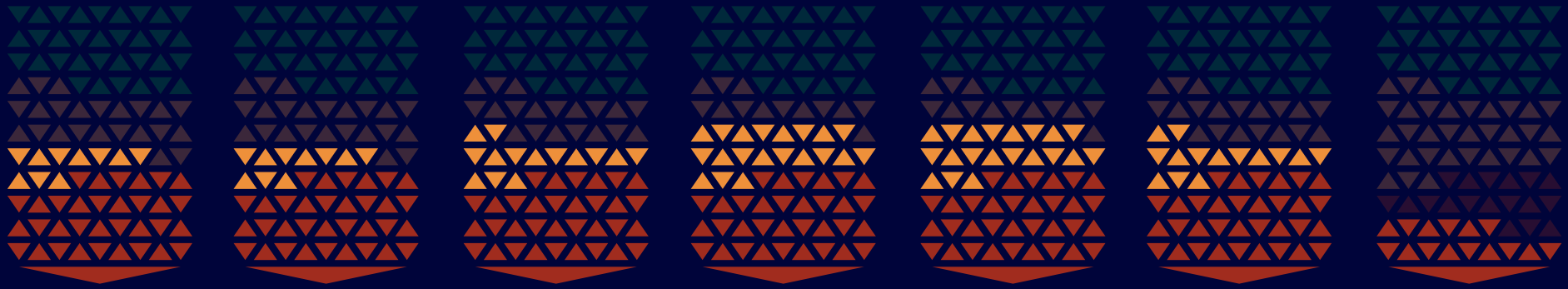
EVERY FIELD we discuss in this Codex is moving FAST. It is important therefore to remember that in the main Creative Machines will one day be: Autonomous, Creative, Democratised, Free, Generalised, Imaginative, Integrated, Massively Parallel, Procedural, Scalable, Self-Evolving, and Self-Improving. For starters ...

2025: AI eliminates Mo Cap and CG workloads by generating and animating 2D characters dynamically. Most part and whole form avatars pass uncanny valley. Clothing, environment, and personalisation increase.

2035: Interactive Synthetic Characters of all kinds are the norm, synthetic characters are generally impossible to distinguish without specialist tools. Regulators and governments are unable to cope with the downsides.

2025 COMMODITISATION & REGULATION 2030 SOCIETAL NORM 2035

2030: High quality Synths of almost all kinds are easy to create after many of the underpinning technologies matured and researchers focused on fine tuning user experience and interfaces.



AUTOMATION

The development of machines capable of automating the entire workflow and lifecycle of synths.

AWARENESS

The development of machines that are actively aware of themselves, their environments, and interactions.

BIOMECHANICS

The development of machines that understand biological biomechanics and principles.

CHARACTER OS

The development of character Operating Systems that give users full control over all aspects of the synths.

CONVERSATIONAL AI

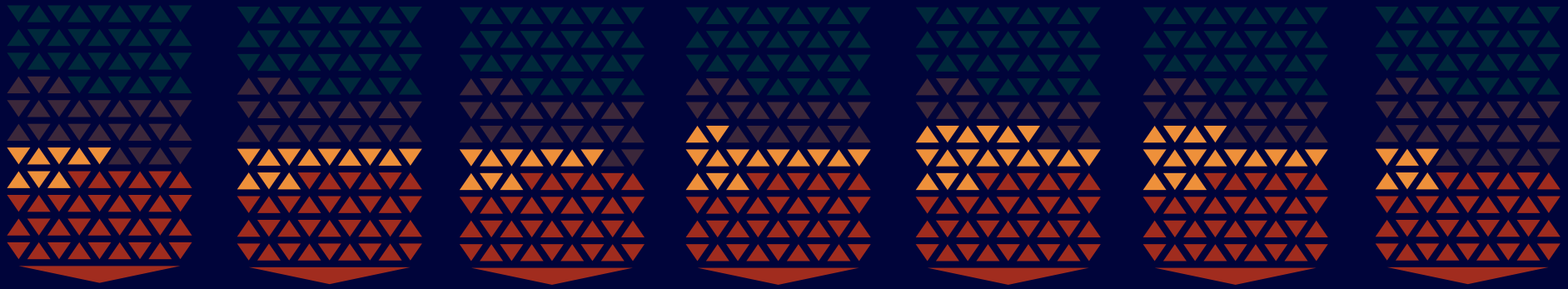
The development of machines capable of holding human-like conversations.

EMOTIONAL IQ

The development of machines capable of empathy and understanding, and displaying emotions.

ENVIRONMENT

The development of authentic and customisable digital or synthetic interactive environments that characters can interact with.



PERSONALISATION

The development of machines whose attributes, behaviours, environments, and looks can all be personalised.

PERSONALITY

The development of machines their own distinct and individual personalities.

REALISM

The development of machines that are in every way indistinguishable from living beings.

REGIONALISATION

The development of machines capable of being fully regionalised according to individual user requirements.

RENDER QUALITY

The development of machines capable of rendering scenes and synths indistinguishable from reality.

RESPONSIVENESS

The development of machines capable of responding to any and all kinds of inputs accurately and in real time.

WHOLE BODY GENERATION

The development of full body synths that act, behave, and move in life-like ways.

Figure 2. The current state of Synthetic Characters development by AI discipline. Source: Multiple Sources

CREATING THE PERFECT GENERATOR

TODAY THERE are arguably two types of companies interested in this market. On the one hand you have the high end creators who have a proud heritage of creating absorbing high quality CG characters for a range of content domains, and then on the other you have startups who have spied a market opportunity to bring a new breed of synthetic characters, predominantly in the form of digital humans, to a new market in new ways.

Needless to say, in time the two camps will end up clashing and competing with one another as the startups move up market and the high end creators move down market. And if high end creators aren't careful the startups could win.

After all, whereas high end creators often spend huge amounts of time, money, and resources developing their high end creations for their elite client base down in the weeds the startups are driven by different market forces, forces that mean they have to automate as much of their process as possible in order to reduce their costs and time to market as much as possible. And as they hone their creations in time creating a new character will be as simple as pushing a button, tweaking some digital dials, or telling a computer what you want it to create for you.

KEY: One triangle represents 1%. The more triangles the more mature the AI discipline, and when it reaches GREEN te technology is Good Enough for adoption.

TECHNOLOGY DEVELOPMENT AND DIRECTION

ONCE ENTREPRENEURS smell a market opportunity, especially one as large as the one offered by synthetic characters, the race to dominate the new market and create the most desirable and marketable products begins in earnest.

As a consequence, today, and just as we've seen elsewhere in the synthetic content space researchers in the synthetic character space are evolving a wide variety of different methods and tools to create their characters, all of which means that the race to dominate this space, to all intents and purposes, is just getting started.

Furthermore, aided by increasingly affordable, democratised, and powerful technologies that race, notably the time it takes to go from the starting line to the main event, is faster than it would at any other point in history.

Lest we forget though it's also worth bearing in mind that just a mere decade ago none of the companies in this space, no matter how well funded or ambitious they were, would have been able to create anything that comes close to what even the most basic synthetic characters are capable of today. That in itself is worth pondering for a moment as we dive deeper into the magic of how it's all done.

Currently the primary areas of development and research in this field include:

- AI dynamic character rendering
- API development
- Biomechanical modelling
- Character modelling
- Character Operating Systems
- Computer graphics tools
- Motion capture technology
- Rendering technology
- Static 2D image animation

COMPLEX CHARACTER MODELLING

In most reinforcement learning benchmarks simulated characters, such as stuntmen and dancers, are represented using simple bio-mechanical and reference models that create only a crude approximation of those characters real world dynamics and movements. As a result, most characters that are created by more traditional AI models are therefore prone to exploiting idiosyncrasies within the simulation and can quickly develop unnatural behaviours that are infeasible in the real world.

By incorporating more realistic bio-mechanical models the team behind DeepMimic were able to overcome this to create characters that had more natural behaviours and dynamics, but that still let the team with the problem of generating high fidelity synthetic character models which, again, meant that the resulting movements could be unnatural.

In order to overcome this the team took an alternative approach and took a data-driven approach, where reference motion capture of humans provided them with examples of the natural motions they were interested in. The characters were then be trained to produce more natural behaviours by imitating the reference motions, and while the results ended up appearing more natural they are still far away from being able to faithfully

Notes:

reproduce the wide range of motions needed in, for example, movies. But nonetheless they are much closer to their ultimate goal of creating characters with lifelike movements than before.

While there is still work to be done the new framework now means that the team have the necessary techniques to develop new models that will allow them to mimic a rich repertoire of challenging skills that range from locomotion to acrobatics, martial arts to dancing.

FEW SHOT ADVERSARIAL LEARNING

AI has already been used to generate realistic DeepFake videos of people, from famous actors to US presidents, thanks in part to the availability of an abundance of visual data on those individuals, but new AI techniques, such as Few-Shot Adversarial Learning, are letting researchers create similar photo-realistic video-like DeepFakes of people when source data is rare. To prove the point researchers used this technique to, literally, bring the Mona Lisa to life – in digital form at least.

The Few Shot Adversarial Learning method is trained on existing talking head datasets. The model extracts face landmarks from video sequences in these datasets, transforms these landmarks into a set of realistic photographs based on the target person, for example in this case the Mona Lisa, then combines the images together to generate a video or gif of them.

Furthermore, as the Mona Lisa showed the technique is capable of even one shot Adversarial Learning. Needless to say though the more training frames the higher quality the final result will be. This few shot learning superpower however does not come easy, as extensive pre-training, or meta-learning, on a large corpus of talking head videos, such as the VoxCeleb1 and VoxCeleb2 corpus, is required.

As illustrated above, the first steps in meta-learning involve translating the head images to embedding vectors with an embedder network. The corresponding results can then be used to predict the generator's adaptive parameters. Then, a generator with updated parameters maps the input of face landmarks into output frames through a set of convolutional layers.

Finally, the objective function of perceptual and adversarial losses, with the latter being implemented via a conditional projection discriminator, are chosen to compare the resulting image

Notes:

with the ground truth image.

NEURAL NETWORK CHARACTERS

Ultimately companies and consumers are going to want to be able to automate the generation of life-like synthetic characters using nothing more than their behavioural computing interfaces, or voices, and a single app that does it all for them without them ever having to lift a creative finger, and that means leaning more on Artificial Intelligence, eliminating the bulky and expensive motion capture and photogrammetry rigs, and automating all the pre and post production CG workflows and rendering tasks.

Once all these technologies and workflows are digitised and automated then Artificial Intelligence will finally be free to generate and dynamically render all the individual character elements, from language and facial expressions, to emotional responses and language, that it needs to create an autonomous life-like digital human, synthetic character, or even virtual being.

DigiDoug from the team at Digital Domain is one step that direction as they hone the technologies needed to create an autonomous digital human that can dynamically interact with everyone it interacts with. To create DigiDoug the team took thousands of images of his donor's face and body, the source, using motion capture and photogrammetry rigs to capture as much data as possible, and then embedded machine learning into their workflows.

Once they had done that, and here's the interesting part, they constructed a deep neural network that would take all those images and learn every detail about the source face – from the mix of expressions to physical details – so it could compute all the information and render it all in real time that, in turn, would allow DigiDoug to behave and react like a real person in real time. And while DigiDoug isn't fully autonomous yet he's getting closer.

Notes:

SYNTHETIC IMAGERY

"In time static 2D Synthetic Image generators will become the foundation of more advanced 3D dynamic Synthetic Video generators."

WE ARE bombarded by imagery every day, whether it's in adverts and online, in comic books and publications, and a myriad of other places, but when it comes to our ability to detect and sort the synthetic imagery from real imagery some categories of synthetic images are easier for us to detect than others.

The reason for this, of course, is obvious - the more real imagery we are bombarded with throughout our daily lives the easier it is for us to spot the fakes, with prime examples being the human face and human form.

As a result the enterprises that are working to develop creative machines capable of generating imagery that passes the uncanny valley test face a barrage of critics - the public. But as the technology, techniques, and tools improve rapidly it's clear for everyone to see that the platforms and the outputs they generate have improved dramatically in just the past five years alone, and we are only going to accelerate from here.

As a consequence it won't be too long now before the vast majority of imagery that these systems generate, whether it is simple 2D cartoon like imagery or more complex and dynamic 3D imagery, will surpass anything that human creatives can produce.

USE CASES AND TRENDS ...

KEY USE CASES AND TRENDS

IMAGERY IS all around us, whether it's the family photos that adorn our homes, the livery and logos that adorn buildings and vehicles, or it's in the articles and marketing that we consume every hour of every day.

There's no doubt that imagery is powerful, for a whole variety of reasons. After all, one incendiary image can cause a riot and topple a government, while at the same time another image can bring about a world wide outpouring of compassion.

As Creative Machines get better at not just generating synthetic imagery ad verbatim but also generating their own creations using their new found codified "imaginations" it's no exaggeration to say that we are entering an extraordinary era that will see machines take on more human-like creative traits and which, fundamentally, will challenge our own views of both ourselves and humanity.

TECHNOLOGY USE CASES

From a developer and market perspective though, let's face it, we're more interested in the use cases with the biggest market opportunities and returns. That said though we can split them into two main categories.

Firstly we have the markets where imagery is the exclusive content type. And secondly we have the markets where imagery is a complimentary component of the overall finished product:

Exclusive Use Cases

- Animations
- Anonymisation
- Avatars
- Artwork
- Fashion modelling
- Image re-touching
- Marketing
- Stock imagery
- Synthetic training datasets

Complimentary Use Cases

- Synthetic video generation
- Synthetic worlds generation

TRENDS BEHIND THE TREND

The impact of synthetic imagery will at first be felt by creatives working in the art, artistic, and photographic sectors whose livelihoods today mainly rely on their ability to create and sell a diverse range of image based content.

However, as the technology matures it will impact the careers and livelihoods of not only the people who appear in those images, such as models, but also the people who own and lease photographic studios. In short the ripple effects will be much greater than many people anticipate.

As jobs are lost though new jobs will emerge to support the new status quo, jobs that include everything from creating and improving these machines, until they do that themselves of course, to helping find new ways to protect and secure them, to defending society against their malicious use, and many more.

Notes:

The same technology that takes away will also democratise creativity in new ways and give non-creatives everywhere, irrespective of their ability, powerful new tools that help them generate their very own imagery and, literally, let their imaginations go like never before.

Moving back to the trajectory of the domain, as with all of the synthetic content generators we're discussing, researchers will continue to develop new machine models, iron out the kinks, and eventually those machines will leave the labs, be adopted by enterprises and integrated with other products and services, and be adopted by consumers - while the regulators play catch up.

FUTURE HISTORY ...

THE FUTURE HISTORY OF SYNTHETIC IMAGERY

2014: Research on using AI to generate Synthetic Imagery is a fringe experiment and produces narrow, low quality, and frankly weird results.

2016: Advances in the field slow down as researchers focus their attentions on refining the efficiency of their AI models.

2018: Further increases in GPU power and AI General Adversarial Network model breakthroughs push some Synthetic Imagery domains past uncanny valley.

2020: Researchers continue to develop new tools to control and influence the final image results. The first platforms are commercialised.

2015 — FIRST 2D-3D CROSS OVER — 2020 — COMMERCIALISATION

2015: Greater GPU power and better AI datasets and training models result in a giant leap forwards for the field.

2017: More researchers around the world start focusing on the issue of Synthetic Imagery generation and the outputs become far more refined and realistic.

2019: Researchers broaden the scope of images their systems can generate, develop new 2D to 3D Domain Transfer techniques, and start using AI to bring static images to life in Synthetic Video and World form.

2021: Developers focus on increasing the range of Synthetic Imagery that can be generated including full body and animated imagery, and teach their models about the fine art of aesthetics and other intrinsic "image" values.

GAME CHANGERS TO BE AWARE OF:

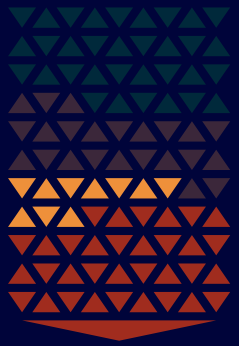
EVERY FIELD we discuss in this Codex is moving FAST. It is important therefore to remember that in the main Creative Machines will one day be: Autonomous, Creative, Democratised, Free, Generalised, Imaginative, Integrated, Massively Parallel, Procedural, Scalable, Self-Evolving, and Self-Improving. For starters ...

2025: The output will be fully controllable and researchers will have turned their attention to developing 3D and dynamic content generators, tech integration, and embodying their models with "imagination."

2030: The field will be fully mature, only areas left to conquer will be those relating to specialist content types. AI is now capable of "true" creativity and imagination, researchers move on to new tasks.

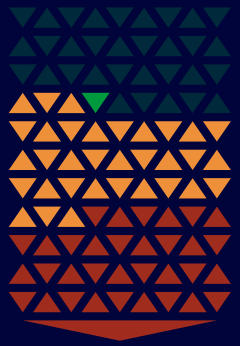
2035: The technology is now firmly the societal norm and innovators are now innovating on top of it and combining it with all manner of other technology formats including immersive ones.

2025 COMMODITISATION & REGULATION 2030 SOCIETAL NORM 2035



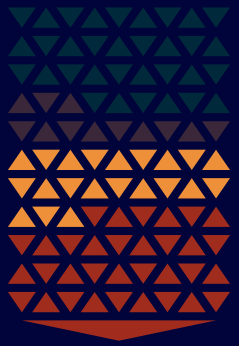
AESTHETICS

The development of machines capable of understanding and replicating beauty.



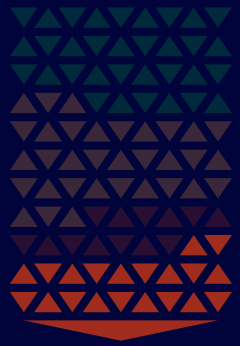
BIOMECHANICS

The development of machines that understand biological biomechanics and principles.



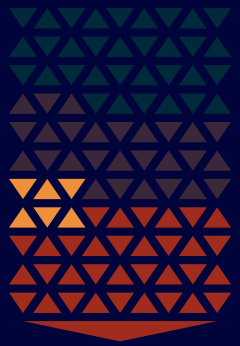
CONTEXT

The development of machines capable of understanding and reproducing context.



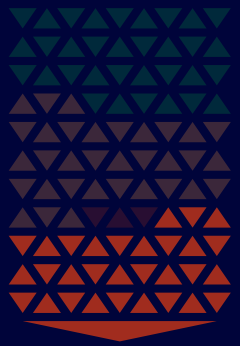
CREATIVITY

The development of creative and imaginative machines.



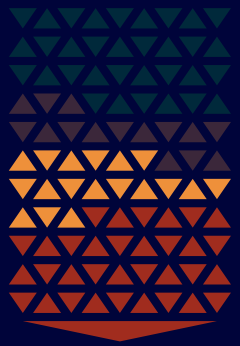
EMOTIONAL RESONANCE

The development of machines that understand how to envoke specific emotional responses.



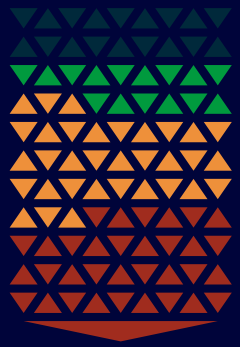
LAWS OF NATURE

The development of machines that understand the laws of nature, including materials and physics.



PERSPECTIVE

The development of machines capable of understanding and replicating the nuances of perspective.



RENDER QUALITY

The development of machines capable of rendering photo-realistic content.

Figure 3. The current state of Synthetic Imagery development by AI discipline. Source: Multiple Sources

CREATING THE PERFECT GENERATOR

OF ALL the different synthetic content types arguably the easiest type of synthetic content for creative machines to learn how to generate is synthetic imagery, because unlike some of the other synthetic content fields, such as synthetic video, generating good enough synthetic imagery requires a much smaller number of AI disciplines to combine and come together before researchers are able to finally produce a single universal and generalised synthetic imagery generator.

Furthermore, to make things easier there is already a huge corpus of training content for them to learn from, and an increasingly vast pool of computing power that they can draw on.

All that said though there is no getting away from the fact that images convey a large amount of visual information, and because sight is, arguably, our primary sense, if every little detail in the images isn't generated perfectly then people notice - and people can be harsh critics.

As a result there is, arguably, a lot more pressure on the researchers in this field to go past just "good enough" and create state of the art machines that fool even the harshest critics - and those machines are already emerging.

KEY: One triangle represents 1%. The more triangles the more mature the AI discipline, and when it reaches GREEN te technology is Good Enough for adoption.

TECHNOLOGY DEVELOPMENT AND DIRECTION

THE SYNTHETIC Imagery space has all the hallmarks of a sheep in wolfs clothing, and what I mean by that is that in time the Artificial Intelligence agents that are being developed to generate high quality 2D imagery will be iterated and tweaked to help developers turn those static 2D images into increasingly animated, dynamic, and sophisticated 3D synthetic content.

In short, whether it is directly or indirectly, the development of the systems in this space will more than likely help drive many of the other synthetic content domains forwards.

Today, the field is pretty much dominated by a type of Artificial Intelligence called Generative Adversarial Networks, or GANs for short.

However, while the synthetic imagery that early GANs generated was often impressive enough to get people to double take the outputs they generated were very difficult to control and tailor post production. As a result, and as researchers start creating the next generation of GANs adaptability and usability are now two of the key areas of focus.

Currently the primary areas of development and research in this field include:

- Image to Image Domain Transfer
- Image to Image generation
- Image generation and manipulation
- Speech to Image generation
- Text to Image generation

CARTOON GAN

Everybody loves animations and cartoons which is why, as you'd expect, some researchers are focused on developing creative machines capable of generating animation content.

Unlike generating traditional imagery though this form of content, such as Japanese Anime, has distinct aesthetics, and today relying on traditional manual transformation techniques to generate real world scenes require considerable expertise and expense as artists must painstakingly draw lines and shade colours by hand in order to create high quality scenes and reproductions.

In order to try to solve this problem in the past researchers in this space turned to Non Photo-realistic Rendering (NPR) as well as more traditional Convolutional Neural Networks (CNN) in order to try to develop working solutions, but both of these techniques proved either too time consuming, and, or, impractical because they didn't produce satisfactory cartoon-like results. After all, different cartoon styles have unique characteristics involving high-level simplification and abstraction, and cartoon images especially tend to have clear edges, smooth colour shading and relatively simple textures, which presented challenges for the Texture-Descriptor based loss functions used in existing methods.

After working though all of the available options most researchers , it appears, are now focusing their energies developing Generative Adversarial Network (GAN) frameworks composed of two CNNs that enable style translation between two unpaired datasets - a Generator for

Notes:

mapping input images to the cartoon manifold, and a Discriminator for judging whether the image is from the target manifold or synthetic, with residual blocks being introduced to simplify the training process.

And in order to avoid slow convergence and obtain the high quality stylization, dedicated semantic content loss and edge-promoting adversarial loss functions they need they've integrated an initialisation phase into this new, so called "cartoonisation" architecture.

An example of this is shown in a GAN named CartoonGAN that shows the importance of each component - the initialisation phase performs a fast convergence to reconstruct the target manifold; sparse regularisation copes with style differences between cartoon images and real-world photos while retaining original contents, and the adversarial loss function creates the clear edges.

Using this technique both real world photos and cartoon images are used for the initial model training phase while the test data contains only real world pictures.

When compared to recently proposed other CNN based image transformation frameworks, such as CycleGAN or Gatys et al's Neural Style Transfer (NST) method, the new technique more successfully reproduces clear edges and smooth shading while accurately retaining the input photo's original content.

However, while the developments so far have been impressive NST only uses a single stylisation reference image for model training which means it cannot deeply learn a particular anime style, especially when there are significant content differences between the stylization reference image and the input images. But, as ever, with more development and training the results will improve.

IMAGINATIVE MACHINES

Creating AI's that are imaginative is an incredibly interesting research area, full of both opportunities and mine fields, that has a wide range of applications – especially in the generation of synthetic imagery and video. And, as difficult and as onerous as it might sound to try to create a machine with its own imagination, fortunately the research teams involved in these types of projects have their own creativity and ingenuity to draw on.

Notes:

For example, if you're handed a note that asks you to draw a picture of a bird with a yellow body, black wings and a short beak, then chances are you'll start with a rough outline of a bird, then glance back at the note, see the yellow part and reach for a yellow pen to fill in the body, read the note again and reach for a black pen to draw the wings and, after a final check, shorten the beak and define it with a reflective glint. Then, just for completeness, you might sketch a tree branch for the bird to sit on. And now this ability, once just the domain of humans, can now be performed by an AI, and it's already probably better at drawing than you are.

The technology, which the researchers behind the most interesting project in the field, simply call "The drawing bot," can generate images of everything from ordinary pastoral scenes, such as grazing livestock, to the absurd, such as a floating double decker bus. But by far the most interesting aspect of what this particular AI is generating is the fact that each image contains details that are absent from the original text descriptions, indicating that this AI does indeed have its own imagination.

At the core of the model is a technology known as a Generative Adversarial Network (GAN). The network consists of two machine learning models, one that generates images from text the descriptions and another, known as a discriminator, that uses text descriptions to judge the authenticity of generated images. The generator then attempts to get fake pictures past the discriminator, and working together, the discriminator rejects the images that don't pass muster and continuously pushes the generator toward perfection.

During their work the researchers trained the drawing bot on datasets that contain paired images and captions, which allowed the model to learn how to match words to the visual representation of those words. The GAN, for example, learns to generate an image of a bird when a caption says bird and, likewise, learns what a picture of a bird should look like.

While GANs work well when generating imagery from simple text descriptions such as a blue bird or an evergreen tree, the researchers found that the model quality stagnated with more complex text descriptions such as a bird with a green crown, yellow wings and a red belly, and that's because the entire sentence serves as a single input to the generator, and the detailed information of the description is lost. As a result, they found that the original images the model

Notes:

created were often a blurry greenish-yellowish-reddish bird instead a close, sharp match with the original description, and this is where they, again, took inspiration from humans.

As humans draw, we repeatedly refer to the text and pay close attention to the words that describe the region of the image we are drawing, so to capture this human trait, the researchers created what they call an Attentional GAN (AttnGAN) that mathematically represents the human concept of attention. It does this by breaking up the input text into individual words and matching those words to specific regions of the image, and once they wrapped this development into their models they found that the synthetic imagery improved dramatically.

STEERABLE GAN

An open secret in traditional machine learning circles is that many AI models work beautifully when used to generate imagery based on standard benchmarks, but that they lack the flexibility and usability that users expect once they leave the lab and start being used in production to create tailored content. This is something that researchers have attributed to the fact that the models used to create certain imagery are often trained on biased, narrow datasets, for example, of cats and dogs that are all centred in the middle of images, that lock their models into generating imagery with very specific styles and viewpoints that users then find difficult to modify to their individual briefs and requirements when they decide to use them to generate content.

More modern generative models too are no exception, and while there have been significant advancements in using GANs to generate different types of dynamic imagery the fact remains that once generated those images are, again, difficult to modify. And this is where so called "Steerable GANs" now come into play that allow users to easily tailor and transform these generated images, such as by adjusting camera angles, colour, and lighting, in any way they like, and they work by letting users drag mouse icons around a 2D latent space, much in the same way we use colour pickers today, to change a wide variety of details in real time.

Researchers in this specific field go on to hypothesise that the degree of distributional shift that can then be applied to images post production is related to the breadth of the training data distribution, and as we use more diverse data sets to train our models

Notes:

the adaptability and flexibility of this technique will improve.

STYLE GAN

Since Generative Adversarial Networks (GANs)were first developed in 2014 by Google Researcher Ian Goodfellow the technology has been widely adopted by researchers interested in using the technology to create synthetic imagery and domain transfer.

After some early failures GANs have made huge breakthroughs recently and can now produce highly convincing fake images of animals, landscapes, human faces, and much more. But while researchers know what GANs can do a lack of transparency in their "Black Box" inner workings means that most GAN results are still achieved mainly through trial and error, which, needless to say offers only limited control over the synthesised imagery.

In order to overcome this problem more modern techniques, from companies such as DeepMind and Nvidia, incorporate so called Style Transfer Techniques (STT) to create new generator architectures that are capable of learning and separating out the different aspects of an image unsupervised; and that enable intuitive, scale-specific control of the resulting synthesis.

And here's an example of how they work: Given an input facial image, the style-based generator can learn its distribution and apply its characteristics on an entirely new and novel synthesised image. While previous GANs couldn't control what specific image features they, or their users rather, wanted to synthesise, using a technique called Style Mixing (SM) these new generators can control the effect of a particular style, for example, high level facial attributes such as identity, pose, and shape, without changing any of the other generated features. Not only does this enable better control over specific features, such as eyes and hair styles, and stylistic content, but it now gives users the ability to more accurately create the images they desire.

Stochastic Variation (SV) is another key property that researchers have introduced that allows GANs to realise the randomisation of certain image attributes, such as the placement of facial hair, freckles, pores, and stubble density, in imagery of human faces.

As a next step researchers are now planning on "quantifying interpolation quality and disentanglement" and propose a variety of new more

Notes:

automated techniques that include Perceptual Path Length (PPL) and Linear Separability (LS) that can be applied to future generator architectures, and as a result of all of these changes researchers have seen impressive results using these modified GANs to generate images of bedrooms, cars, and cats, as well as anime maestri and real world scenery.

SUMMARY

The ability to produce a diverse range of high quality imagery on demand will be very appealing to those enterprises and individuals that today have to pay large sums of money to access and publish copyrighted original works, however, the main impact of the field will be the sudden explosion in new content.

SYNTHETIC TEXT

Notes:

ENTERPRISES AROUND the world have been trying to use Artificial Intelligence (AI) in all its various forms to analyse text for arguably a couple of decades now, whether it's for national security purposes or academic ones. But the phenomenon of using AI to generate text, for a variety of reasons, is a relatively new phenomenon that has its roots in the world of intelligence gathering where national security agencies such as the NSA, faced by the huge exascale volumes of intelligence data, wanted to find new ways to condense and summarise it all into easily digestible narratives that were easier for their human analysts to digest.

The global national security community, however, aren't the only ones who face a data deluge, and over time many of the concepts developed and honed in their field were translated across to other sectors and, where necessary tailored accordingly. As a result the first commercial platforms to emerge were news generators, so called Robo-Journalists capable of ingesting large volumes of data, making sense of it all, and then condensing it down into easy to digest news articles.

USE CASES AND TRENDS ...

KEY USE CASES AND TRENDS

THE IMPORTANCE of text, or to be more specific the written word, is all too easy to underestimate. However, despite this it's easy to argue that text is one of humanities most important forms of content if not the most important form.

Over the past millennia it's been the foundation of every human society. From helping us create a historical record of the world, as well as forecast the future, to being the tool we use to educate the world and helping us open gateways into new worlds that help unleash our innate human creativity in new ways.

Today text has evolved, no longer simply encapsulated in ink on paper it's electrons and photons and runs our digital planet.

As a consequence, as Creative Machines become increasingly capable of generating all forms of synthetic text not only can they re-write human history and change our perceptions of the future, and open the door to new imaginary worlds, but they can also re-wire the digital world and have an impact in everything that relies on code.

TECHNOLOGY USE CASES

The use cases for synthetic text generators are as staggering as they are diverse

From a developer and market perspective though, let's face it, we're more interested in the use cases with the biggest market opportunities and returns. That said though we can split them into two main categories.

Firstly we have the markets where imagery is the exclusive content type. And secondly we have the markets where imagery is a complimentary component of the overall finished product:

Exclusive Use Cases

- Article writing
- Blogging
- Book writing
- Copy writing
- Conversational AI
- Lyrics
- Machine programmers
- Poetry
- Procedural text generation
- Report generation
- Research papers
- RPG game generation
- Script writing
- Summaries
- Synthetic training datasets

Complimentary Use Cases

- Synthetic audio generators
- Synthetic video generators
- Synthetic world generators

TRENDS BEHIND THE TREND

Text is the foundation of more things in our society than we might at first realise so it shouldn't come as any surprise that as synthetic text generators improve more people are becoming aware of the impact they'll have on everything from human jobs, such as journalists, programmers, and script writers, through to the rate of software innovation, and beyond, and the consequences thereof.

Over the past few years, fuelled by

Notes:

billions of dollars in grants and access to almost limitless compute resources, the rate of innovation in the domain has been nothing short of spectacular.

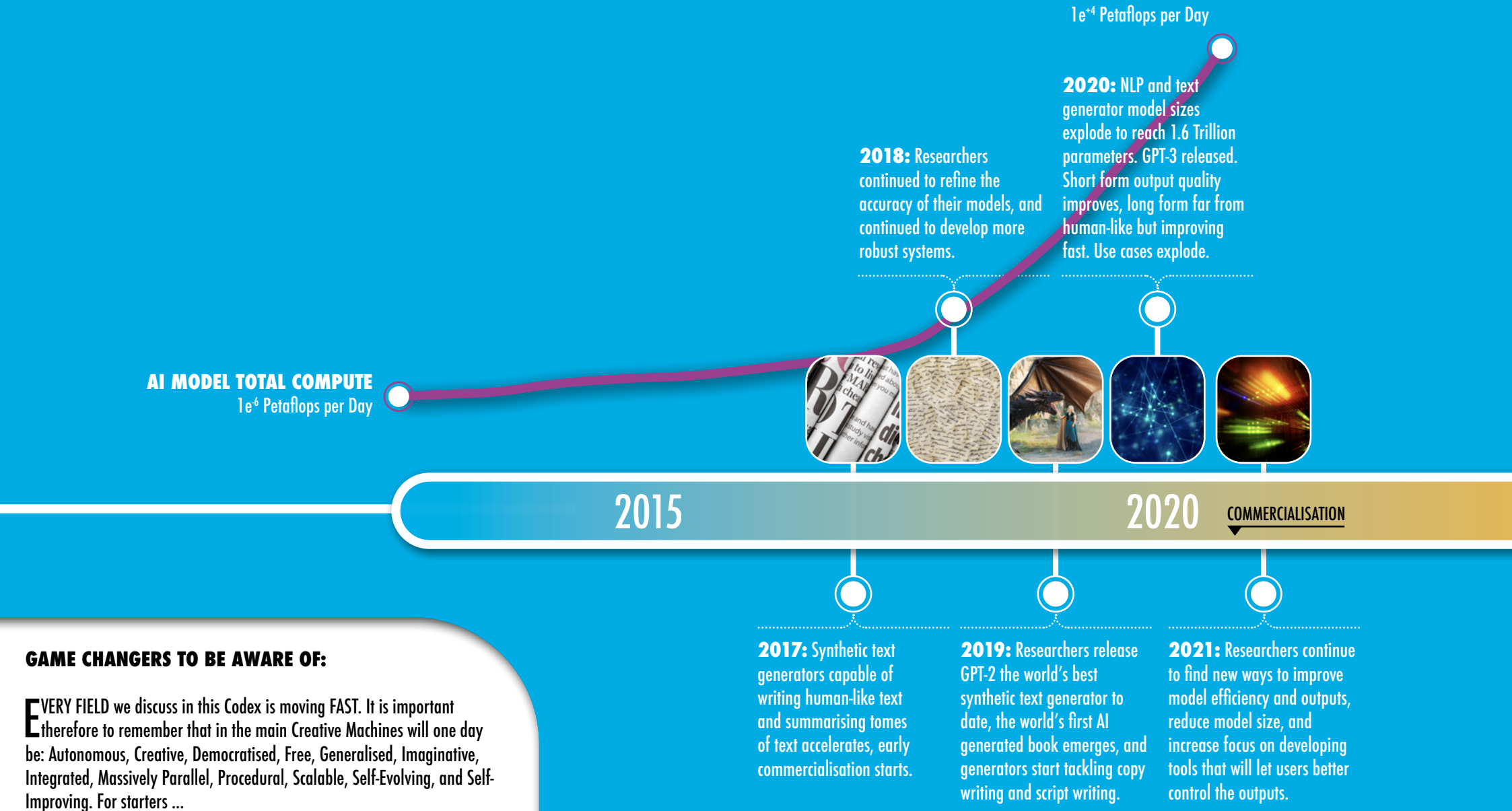
However, that progress has come at a significant cost and is now showing signs of slowing as the sectors dominant players grapple with the economic implications of developing models that at first were tens of billions of parameters in size, then hundreds of billions, and that are now measured in the trillions.

Asides from the commercial implications though there are also the brand and societal implications of continuing to develop AI's that consume the energy of a small country, all of which means that many think that an AI winter is coming. It also means that many of the smaller players in the sector are finding it increasingly difficult to compete and are increasingly choosing to tap into the incumbents models rather than keeping on funding their own.

In time though researchers will find ways to improve the energy efficiency of their massive models, and find new methods to improve model efficiency.

FUTURE HISTORY ...

THE FUTURE HISTORY OF SYNTHETIC TEXT



GAME CHANGERS TO BE AWARE OF:

EVERY FIELD we discuss in this Codex is moving FAST. It is important therefore to remember that in the main Creative Machines will one day be: Autonomous, Creative, Democratised, Free, Generalised, Imaginative, Integrated, Massively Parallel, Procedural, Scalable, Self-Evolving, and Self-Improving. For starters ...

2025: Tech is good enough or as close as for most commonplace applications with only specialist and multi-lingual use cases outstanding. Complimentary technology integrations accelerate.

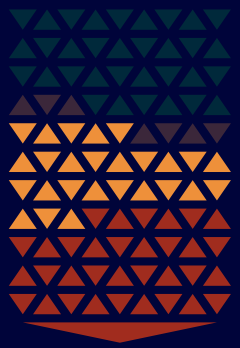
2025 COMMODITISATION & REGULATION

2030 SOCIETAL NORM

2030: Tech is fully commercialised and mature for all but the most specialist applications. AI reaches cultural tipping point, developers focus on innovating new use cases on top of the existing models.

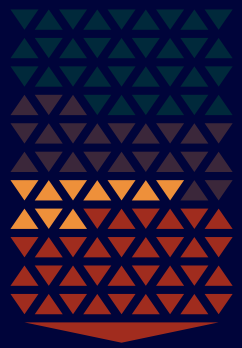
2035: Tech is fully mature and has become the societal norm.

2035



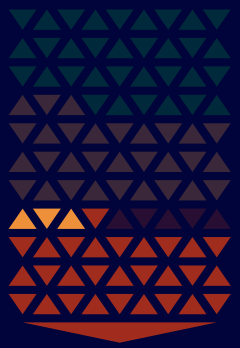
CONTEXT

The development of machines capable of understanding and replicating context.



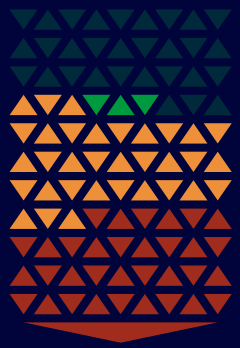
DIALOGUE

The development of machines capable of understanding and generating dialogue.



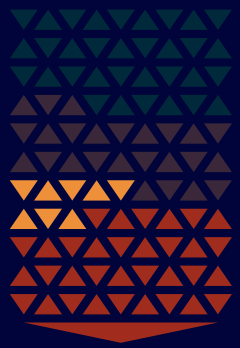
CREATIVITY

The development of creative and imaginative machines.



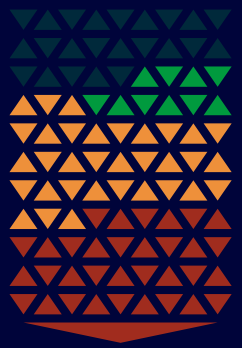
DIALECTS

The development of machines that can understand and talk in multiple dialects.



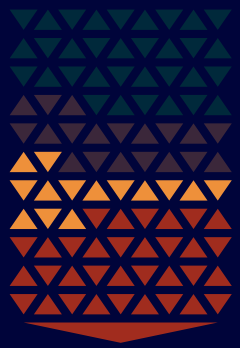
EMOTIONAL RESONANCE

The development of machines that understand how to evoke specific emotional responses.



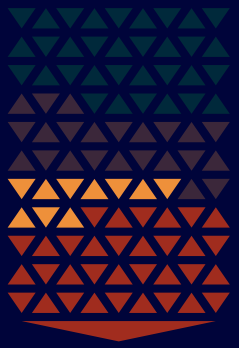
NLP

The development of machines capable of understanding and generating natural language.



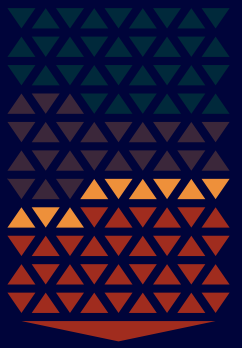
PROCEDURAL FLOW

The development of machines capable of generating content that flows.



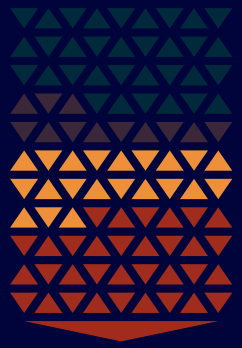
STORYTELLING

The development of machines capable of creating and telling great stories.



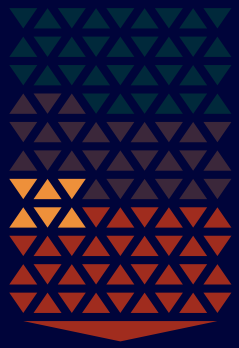
TEMPO

The development of machines capable of understanding and replicating the impact of tempo.



TONE

The development of machines capable of understanding and replicating the impact of tones.



VALENCE

The development of machines capable of understanding and replicating specific emotional impacts.

Figure 4. The current state of Synthetic Text development by AI discipline. Source: Multiple Sources

CREATING THE PERFECT GENERATOR

UNLIKE MANY of their other synthetic content cousins, there are already a number of good synthetic text generators in the wild and in commercial use in fields as diverse as copy writing and journalism, and the rate of development in this field is accelerating fast, so much so that we have already seen these platforms write the first robo-novels and even rudimentary movie scripts.

That said though, like many of the technologies in this megatrend, it's not a hill they have to conquer it's a mountain, and as a result there are a number of different AI discipline areas that need to combine together before we have a truly universal and generalised synthetic text generator.

As a result of their early successes many researchers in the field are now turning their attention to developing creative machines capable of generating more long form content - a feat that's much more difficult to achieve because, again, like the other synthetic content fields, these machines have to draw on a variety of other disciplines that include everything from learning that stories have a beginning, a middle, and an end, all the way through to being able to understand context, and generate text capable of soliciting the right emotional responses.

As these machines command of the natural language improve though in time researchers in the field will be able to devote more of their time training their models to generate the right rhythm and flow for pieces, find new ways to help them unleash their creativity and imagination, and create stories to be proud of.

KEY: One triangle represents 1%. The more triangles the more mature the AI discipline, and when it reaches GREEN te technology is Good Enough for adoption.

TECHNOLOGY DEVELOPMENT AND DIRECTION

RECENTLY A number of enterprises, such as OpenAI, a non-profit AI research firm in the USA, have released a range of new AI language models that are capable of generating convincing passages of prose. So convincing, in fact, that not only have some of these caught the media's attention and imagination, but in some cases the researchers have refrained from open sourcing the models in the hope of stalling their eventual and, unfortunately, inevitable weaponisation – especially when it comes to using them to mass produce fake news. Putting this into context it's not for no reason that the team of researchers behind one of the most advanced models called their creation "potentially most dangerous AI out there," because it was that good at creating fake content.

While these impressive results are a remarkable leap beyond what traditional language models have been able to achieve up until now the technique involved in creating them isn't all that new. One of the latest breakthroughs, for example, was driven by feeding the model an ever increasing amount of training data - a technique that's also behind many other recent advancements in the field which now include many of the models being used by everyone from the Wall Street Journal, to the researchers who are using them to create the first ever AI generated books – the first of which, a research book that summaries all of the latest Lithium Ion battery research, has already been released.

"The latest [OpenAI] model is surprising people in terms of what they can create with more data and bigger models," says Percy Liang, a leading computer science professor at Stanford University.

Once trained the passages of text many current models produce are good enough to masquerade as something written by humans, however, this ability to write should not be confused with a genuine understanding of language, which is the ultimate goal of the subfield of AI known as Natural Language Processing, or NLP for short.

In fact, getting machines to the point where they have a genuine understanding of natural language is a challenge that has so far largely eluded NLP researchers, and it's a goal that many experts think could take years, even decades, to achieve and involve techniques that haven't been developed yet.

That said though there have been some noteworthy developments in the field with the development of new models that are being training to generate new scripts for a wide array of popular series including Game of Thrones.

At the moment four different philosophies of language that currently drive the development of NLP techniques, so let's look into them all.

The primary areas of research in this field today include:

- Contextual meaning
- Copy writing
- Long form text generation
- Reasoning
- Short form text generation
- Script generation
- Semantic meaning

Notes:

- Summarising

DISRIBUTIONAL SEMANTICS

Linguistic philosophy. Words derive meaning from how they are used. For example, the words "cat" and "dog" are related in meaning because they are used more or less the same way. You can feed and pet a cat, and you feed and pet a dog. You can't, however, feed and pet an orange.

How this translates to NLP. Algorithms based on distributional semantics have been largely responsible for many of the recent breakthroughs in NLP, and, like many fields, it's a field that's accelerating. Researchers use machine learning to process text, finding patterns by essentially counting how often and how closely words are used in relation to one another. The resultant models can then use those patterns to construct complete sentences or paragraphs, and power things like auto-complete or other predictive text systems.

In recent years, some researchers have also begun experimenting with looking at the distributions of random character sequences rather than words, so models can more flexibly handle acronyms, punctuation, slang, and other things that don't appear in the dictionary, as well as languages that don't have clear delineations between words.
Pros and Cons. These algorithms are flexible and scalable because they can be applied within any context and learn from unlabelled data, but on the negative side the models they produce don't actually understand the sentences they construct - at the end of the day, they're writing prose using word mere associations.

FRAME SEMANTICS

Linguistic philosophy. Language is used to describe actions and events, so sentences can be subdivided into subjects, verbs, and modifiers, such as who, what, where, and when.

How this translates to NLP. Algorithms based on frame semantics use a set of rules or lots of labelled training data to learn to deconstruct sentences. This makes them particularly good at parsing simple commands, and thus very useful for creating chatbots or voice assistants. If, for example, you asked Alexa to "Find a restaurant with four stars for tomorrow," such an algorithm would have to figure out how to execute the sentence by breaking it down into the action "find", the what "restaurant with

Notes:

four stars," and the when "tomorrow."

Pros and Cons. Unlike distributional semantic algorithms though, which don't understand the text they learn from, frame semantic algorithms are able to distinguish between the different pieces of information in a sentence, and these can be used to answer questions like "When is this event taking place?" On the negative side however currently these algorithms can only handle very simple sentences and therefore fail to capture nuance because they require a lot of context specific training. Also as a consequence they're relatively brittle and not particularly flexible.

MODEL THEORETICAL SEMANTICS

Linguistic philosophy. Language is used to communicate human knowledge.

How that translates to NLP. Model theoretical semantics is based on an old idea in AI that all of human knowledge can be encoded, or modelled, in a series of logical rules. So, if you know that birds can fly, and hawks are birds, then you can deduce that hawks can fly. This approach is no longer in vogue, however, because researchers soon realised there were too many exceptions to each rule, for example, penguins are birds but they can't fly, which would result in an error.

But algorithms based on model theoretical semantics are still useful for extracting information from models of knowledge where the information is always consistent and logical, such as databases. Like frame semantics algorithms, they parse sentences by deconstructing them into parts, but whereas frame semantics defines those parts as the who, what, where, and when, model theoretical semantics defines them as the logical rules encoding knowledge.

For example, consider the question "What is the largest city in China by population?" A model theoretical algorithm would break it down into a series of self-contained queries such as "What are all the cities in the world?" "Which ones are in China?" "What are the cities' populations?" "Which population is the largest?" It would then be able to traverse the model of knowledge to get you your final answer.

Pros and Cons. These algorithms give machines the ability to answer complex and nuanced questions, but on the negative side they require a model of knowledge, which is time consuming to build, and are not flexible across

Notes:

different contexts.

GROUNDED SEMANTICS

Linguistic philosophy. Language derives meaning from lived experience. In other words, humans created language to achieve their goals, so language must be understood within the context of our goal oriented world.

How this translates to NLP. This is the most modern approach and the one that many experts in the field thinks holds the most promise. It tries to mimic how humans pick up language over the course of their lifetimes, so the machines start with a blank state and learn to associate words with the correct meanings through conversation and interaction.

In a simple example, if you wanted to teach a computer how to move objects around in a virtual world, you would give it a command like "Move the red block to the left" and then show it what you meant, and then over time the machine would learn to understand and execute the commands without the need for human interaction.

Pros and Cons. In theory, these algorithms should be very flexible and get the closest to a genuine understanding of language, but on the negative side teaching is very time intensive - and not all words and phrases are as easy to illustrate as "Move the red block."

In the short term, most experts think that the field of NLP will see much more progress from exploiting existing techniques, particularly those based on distributional semantics, but in the longer term they all believe they have limits.

"There's probably a qualitative gap between the way that humans understand language and perceive the world and our current models," says Liang.

Closing that gap would probably require a new way of thinking, he adds, as well as much more times.

SUMMARY

Like every creative machine field it's fair to say that generating synthetic text, for whatever purpose, is a field that on the one hand is coming under increasing scrutiny because of the impact that the ability to generate synthetic text will have on society and especially trust, and on the other hand over the past year research in the field has progressed rapidly with an increasingly large

Notes:

number of AI models being produced that many experts have deemed too "dangerous" to be released into the wild.

However, as the number of models produced, and the volume of research and papers grow, it is inevitable that the technology will escape into the wild and be democratised for better and worse.

Notes:

SYNTHETIC VIDEO

"In just 3 years DeepFakes went from expensive low grade lab project to being high grade and democratised for free in apps at global scale."

OF ALL the different content types synthetic video is by far the most difficult for creative machines to generate by a long shot, and that's before we try to combine so called flat synthetic video with other media types in order to create rich, fully featured output.

The reason for all of this is, of course, fairly obvious, after all video is a much more dynamic medium than many of the other media types and it's also packed with far more information, most of which, in one way or another, has to interplay with one other - whether it's the environment interplaying with the characters, the characters with the objects, or the soundtrack with the plot lines, and so on.

So while, for example, our ears might forgive the odd trip up in an audio stream now and again, or a slightly out of place squiggle in an image, any weaknesses in a synthetic video generator would be picked up instantly, and then, because of the dynamic nature of the media multiplied many fold which would, as a consequence, make any imperfections increasingly difficult to ignore.

In short, the bar to create high fidelity synthetic video that passes the uncanny valley test is much higher than it is with other media types, and that's one of this category's greatest hurdles.

USE CASES AND TRENDS ...

KEY USE CASES AND TRENDS

IF THE modern age belongs to one specific type of content then it would be easy to argue that its video. But, despite the ubiquity of smartphones that have democratised video creation for the billions of people that use them creating high quality video content, let alone high quality long form video content, is still for the most part beyond the vast majority of people.

However, as increasingly emerging technology helps developers create increasingly powerful tools in time all this too will be democratised at scale, and that's before we discuss the eventual and inevitable emergence of Creative Machines that give users to create any and every kind of content either autonomously or at the push of a proverbial button.

TECHNOLOGY USE CASES

If a picture says a thousand words then the bar for video has got to be extremely high, and when we think about the use cases for synthetic video generators, let alone the fact that the markets for those use cases will be multiplied hundreds fold as new space based and terrestrial communications technologies, such as 5G and 6G, help double the number of people who are connected to the internet we could, again, quite easily argue that in terms of the addressable market opportunity video has only just started realising its eventual potential.

And all of that is before we discuss how it will become a key component of new forms of content, such as Augmented, Mixed, and Virtual Reality, or Synthetic Worlds, and the Transmedia Metaverse.

From a developer and market perspective though, let's face it, we're more interested in the use cases with the biggest market opportunities and returns. That said though we can split them into two main categories.

Firstly we have the markets where video is the exclusive content type. And secondly we have the markets where video is a complimentary component of the overall finished product:

Exclusive Use Cases

- Animations
- Documentaries
- Educational content
- Marketing
- Movie generation
- Procedural content generation
- Stock videography
- Synthetic training datasets
- TV series generation
- Video shorts

Complimentary Use Cases

- AR/MR/VR augmentation
- Synthetic world generation
- VR world generation

TRENDS BEHIND THE TREND

Today we can think of video as, ironically, being 2D and flat, which is more of a function of the two dimensional devices and displays we use to watch it on rather than the medium itself.

In the future, however, as we see new

Notes:

devices continue to emerge, which will include everything from Brain Machine Interfaces and Smart Contact Lenses, to full immersion headsets and haptic clothing, the common-a-garden video we experience today will be increasingly 3D, in volumetric terms, and immersive.

Furthermore, the development of new consumption devices will also give content creators and video developers everywhere to let their imaginations go and bring consumers a whole new world of experiences, whether, for example, it's the introduction of haptic sensations and the ability to experience virtual smells and tastes, or the use of Virtual Reality floors and neuro-technologies to literally transport users minds and bodies to alternative realities.

Stretching the timeline further out in time Creative Machines will autonomously create their own procedural video content and they too will be able to enrich it with a wide variety of new components.

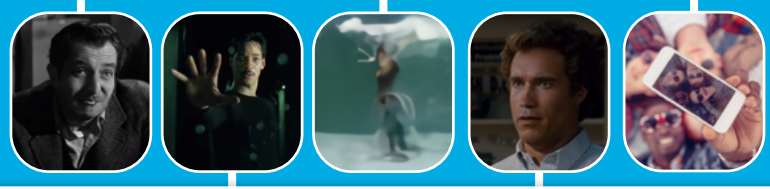
FUTURE HISTORY ...

THE FUTURE HISTORY OF SYNTHETIC VIDEO

2018: First basic Text to Video, Video to Video, and Video to Video Domain Transfer generators emerge, and AI produces the world's second very awful short movie.

2020: Researchers accelerate development of models that generate video from scratch, and predictive AI models that can reason what video strings should come next. Startups leverage manual and automated tools.

2022: Techniques to generate Synthetic Video, including AI only, CG-Hybrid, Synthetic Imagery and Worlds, and Text to X, start converging, hardware is digitised, generation is increasingly automated.



2015 — 2020 COMMERCIALISATION

2019: First free DeepFake and Video Domain Transfer apps capable of transferring users into movies, full body DeepFakes and AI's able to generate Text to Video and short synthetic videos emerge.

2021: Researchers develop AI's that automate more of the video creation workflow and outside of Deepfake improvements most pass worthy synthetic video output still needs heavy human involvement.

GAME CHANGERS TO BE AWARE OF:

EVERY FIELD we discuss in this Codex is moving FAST. It is important therefore to remember that in the main Creative Machines will one day be: Autonomous, Creative, Democratised, Free, Generalised, Imaginative, Integrated, Massively Parallel, Procedural, Scalable, Self-Evolving, and Self-Improving. For starters ...

2025: Some Synthetic Video generators pass uncanny valley and are integrated with other technologies. Research turns to creating immersive content, tools to control the final output, and autonomous procedural content generation.

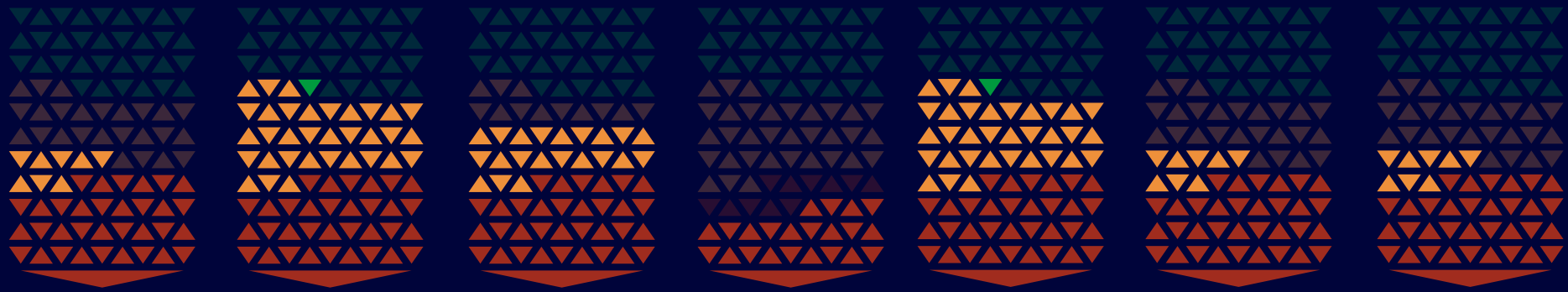
2025 COMMODITISATION & REGULATION

2030 SOCIETAL NORM

2030: Tech is nearing maturity for some major market categories, but is still error prone. Using Big Data to generate truly interactive, procedural experiences faces regulatory questions.

2035: All underlying technologies are near maturity and the technology has been democratised. High quality procedural synthetic movies are possible but likely need refinement.

2035



AESTHETICS

The development of machines capable of understanding and replicating beauty.

BIOMECHANICS

The development of machines that understand biological biomechanics and principles.

CONTEXT

The development of machines capable of understanding and reproducing context.

CREATIVITY

The development of creative and imaginative machines.

DIALECTS

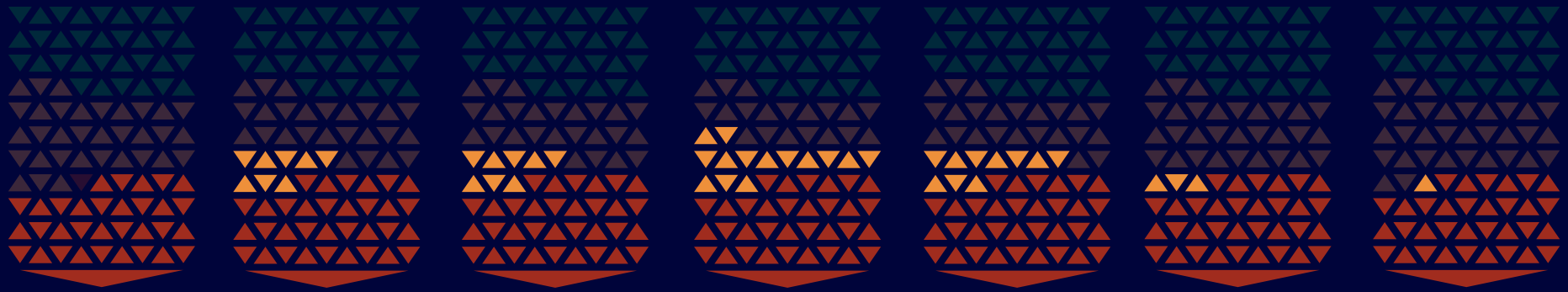
The development of machines that can understand and talk in multiple dialects.

EMOTIONAL RESONANCE

The development of machines that understand how to envoke specific emotional responses.

INTONATION

The development of machines capable of understanding and copying human intonation patterns.



LAWS OF NATURE

The development of machines that understand the laws of nature, including materials and physics.

PERSPECTIVE

The development of machines capable of understanding and replicating the nuances of perspective.

PROCEDURAL FLOW

The development of machines capable of generating content that flows.

RENDER QUALITY

The development of machines capable of rendering photo-realistic content.

SEMANTICS

The development of machines that understand implication, reference, and sense.

STORYTELLING

The development of machines capable of creating and telling great stories.

TEMPO

The development of machines capable of understanding and replicating the impact of tempo.

Figure 6. The current state of Synthetic VFX development by AI discipline. Source: Multiple Sources

CREATING THE PERFECT GENERATOR

WHEN WE look at all the different types of creative machines that are emerging, arguably, the ones that everyone are looking forwards to seeing the most are those that will help democratise video content creation.

After all, from a marketing perspective at least video is the new E-Mail. And while there have been significant breakthroughs in the space, as I've discussed, a lot of those developments are being driven by the interest in so called DeepFake technologies that first rose to fame a few years ago, but before researchers are able to create a single universal and generalised synthetic video generator obviously a lot of separate AI disciplines have to mature and combine.

Unlike the other synthetic content fields synthetic video generators are one of the harder technologies to develop because unlike synthetic imagery, for example, creating a viable system is reliant on a whole host of complimentary disciplines that, as you can see from the diagram, include everything from being able to understand the laws of nature, so that, for example, the system knows that if someone drops a ball then gravity acts on it, all the way through to being able to combine a vast amount of different content types together to create something that gets anywhere near resembling a modern day movie - complete with emotional twists, plot lines, soundtracks, and special effects.

KEY: One triangle represents 1%. The more triangles the more mature the AI discipline, and when it reaches GREEN te technology is Good Enough for adoption.

TECHNOLOGY DEVELOPMENT AND DIRECTION

TRYING TO create convincing synthetic video is one thing, but ultimately that's not most people's end goal, unless that is they're looking for stock videography.

For the majority of consumers in this category their end goal is to produce rich video content that at a minimum is complete with characters, soundtracks, and storylines - irrespective of how long they want that content to be, or its use case - that can also be, in an ideal world, easily edited and modified.

The primary areas of research in this field today include:

- Imaginative AI
- Predictive AI
- Text to Video generation
- Video to Video Domain Transfer
- Video to Video generation

IMAGINATIVE MACHINES

Creating AI's that are imaginative is an incredibly interesting research area, full of both opportunities and mine fields, that has a wide range of applications – especially in the generation of synthetic imagery and video. And, as difficult and as onerous as it might sound to try to create a machine with its own imagination, fortunately the research teams involved in these types of projects have their own creativity and ingenuity to draw on.

For example, if you're handed a note that asks you to draw a picture of a bird with a yellow body, black wings and a short beak, then chances are you'll start with a rough outline of a bird, then glance back at the note, see the yellow part and reach for a yellow pen to fill in the body, read the note again and reach for a black pen to draw the wings and, after a final check, shorten the beak and define it with a reflective glint. Then, just for completeness, you might sketch a tree branch for the bird to sit on. And now this ability, once just the domain of humans, can now be performed by an AI, and it's already probably better at drawing than you are.

The technology, which the researchers behind the most interesting project in the field, simply call "The drawing bot," can generate images of everything from ordinary pastoral scenes, such as grazing livestock, to the absurd, such as a floating double decker bus. But by far the most interesting aspect of what this particular AI is generating is the fact that each image contains details that are absent from the original text descriptions, indicating that this AI does indeed have its own imagination.

At the core of the model is a technology known as a Generative Adversarial Network (GAN). The network consists of two machine learning models, one that generates images from text the descriptions and another, known as a discriminator, that uses text descriptions to judge the authenticity of generated images. The generator then attempts to get fake pictures past the discriminator, and working together, the discriminator rejects the images that don't pass muster and continuously pushes the generator toward perfection.

During their work the researchers trained the drawing bot on datasets that contain paired images and captions, which

Notes:

allowed the model to learn how to match words to the visual representation of those words. The GAN, for example, learns to generate an image of a bird when a caption says bird and, likewise, learns what a picture of a bird should look like.

While GANs work well when generating imagery from simple text descriptions such as a blue bird or an evergreen tree, the researchers found that the model quality stagnated with more complex text descriptions such as a bird with a green crown, yellow wings and a red belly, and that's because the entire sentence serves as a single input to the generator, and the detailed information of the description is lost. As a result, they found that the original images the model created were often a blurry greenish-yellowish-reddish bird instead a close, sharp match with the original description, and this is where they, again, took inspiration from humans.

As humans draw, we repeatedly refer to the text and pay close attention to the words that describe the region of the image we are drawing, so to capture this human trait, the researchers created what they call an Attentional GAN (AttnGAN) that mathematically represents the human concept of attention. It does this by breaking up the input text into individual words and matching those words to specific regions of the image, and once they wrapped this development into their models they found that the synthetic imagery improved dramatically.

TEXT TO VIDEO EDITORS

Once we've generated our video content it's highly likely that the production version is at some point, for whatever reason, going to need to be edited. And this creates yet another challenge in a field full of challenges, and consequently needs researchers to develop yet more tools to solve the problem.

Fortunately however over the past couple of years, and as the volume of synthetic video in the wild increases, in all its forms, there have been a number of research teams around the world who've been focusing in on this very problem, and their research is needless to say still at the early stages.

One of the first fields that's been getting researchers attention is the ability to modify and manipulate talking heads and the spoken word, with recent developments in the field being likened to "Photoshop for voice."

In order to create their video fakes,

Notes:

which are used to replace the original synthetic video content, researchers have had to combine a number of different techniques together, with their first challenge being to accurately scan the target video to isolate phenomes spoken by the original subject.

Once this was complete they were then able to match their phonemes with corresponding visemes, which are the facial expressions that accompany each sound. And finally, they created a new 3D parametric model of the lower half of the subject's face using the target video as a reference.

When someone edits a text transcript of the video, the software combines all this collected data - the phonemes, visemes, and 3D parametric face model - to construct new footage that matches the text input. This is then automatically pasted onto the source video at the right timestamp to create the final result.

Despite impressive results so far there are, obviously limitations to these techniques such as the algorithms only work on talking head style videos and require 40 minutes of input data. Also, the edited speech can't differ too much from the source material, and in their best quality fakes, the researchers actually had to ask subjects to record new audio to match the changes and then used AI to generate the corresponding modified video frames.

And then, furthermore, the research isn't yet at the point where we can change the mood or tone of the speaker's voice – but all of these problems will be solved in time.

TEXT TO VIDEO GENERATORS

Generating images from text is a well studied field, but generating video, whether it's just clips, shorts, or even one day entire movies, from text is still a very nacent field of research that is yet to be explored as extensively, but increasingly it's an area of interest for many research teams around the world. And as a consequence there still isn't a clear consensus on which models and techniques are the best, but this, I think is one that shows the greatest promise – even if it isn't the finished article yet.

Previous work on the generative relationship between text and short video clips has focused on producing text captioning from video, but unsurprisingly the inverse problem of producing videos from text has alot more degrees of freedom and is a challenging problem for traditional modelling methods. For

Notes:

example, a key consideration when using text as the foundation to generate new video content from is that both the scenes and object motion must be directly determined by the text input, and what many of the researchers in the field have found so far is that simply adapting traditional Text to Image generation methods results in videos where the progression of the video quickly descends into chaos.

In the latest research from the US though researchers have been considering how to tackle motion and background synthesis from text, which, they feel, is related to video prediction where, ultimately the goal for new models is to learn how to navigate non-linear transformation functions between specific frames so that they can predict what the next frames in the sequence should be. However, simply being able to predict future frames is not enough to generate a complete video clip.

In order to tackle some of the inherent problems researchers have found that decomposing the individual video elements into a static background, a mask and moving objects, has helped significantly and had the additional benefit of simplifying their model's architectures, and then once the first stage is complete next comes the process of generating the video, the Generative Adversarial Network (GAN) stage.

This stage, which is where the magic happens, and where all the corresponding layers and content, complete with all of their associations, are generated, is broken down into two components. First, a conditional Variational Autoencoder (VAE) model is used to generate the "gist" of the video from the input text, where the gist is an image that gives the background colour and object layout of the desired video. The content and motion of the video is then generated by conditioning on both the gist and text input using a not too dissimilar process that mimics how humans create art. Specifically - artists often draw a broad draft and then fill in the detailed information afterwards.

In other words, the gist generation step extracts static "universal" features from the text, while the video generator extracts the dynamic "detailed" information from the text to generate the final video output.

SUMMARY

The ability to generate high quality synthetic video content on demand will be on the creative machine's crowning

Notes:

moments, and despite the impressive achievements in other synthetic content fields it has to be said that nothing sparks the human imagination in the same way that, for example, movies can. Furthermore, videos are the ultimate education and story telling medium, but the road to creating machines that master this medium, as well as all of the other nuances that go along with it, will be a long one.

Notes:

SYNTHETIC VFX

EXPLORE MORE. Click or scan me to learn more about this emerging tech.

To explore more about the topics discussed in this section click the corresponding metatags:

ROTOSCOPING

SYNTHETIC DUBBING

VIRTUAL STUNTMEN

BEING ABLE to generate high quality synthetic video is one thing, however, let's face it most movies probably wouldn't be the blockbusters they are today if it wasn't for special effects, and in some respects that is one of the next fields to be tackled and conquered by creative machines in a long line of things to be conquered.

As the ability of creative machines to generate synthetic content improves though the researchers behind them are increasingly getting noticed and hired by some of the world's largest studios who over the past couple of years have ramping up their interest in the field and the people behind it before experimenting with these new tools and debating how best to incorporate the technology into their work flows when it's mature enough.

However, given the scale of re-tooling that would be needed to happen in order to convert their studios to include these deep learning driven work flows it's likely that it's going to be some time before CGI, the technology that was first ushered in by Jurassic Park, has its own Jurassic Park moment, and finally gets consigned to the history books in the same way that it consigned the era of mechanical props and photochemical VFX to the history books that preceded it.

Before that happens though one of the main issues that the studios have with today's developments is that the outputs aren't completely controllable or predictable. That said though where there's a problem there's a researcher eagerly trying to solve it, and as research in the field accelerates it will be only a matter of time before we see these technologies enter the studios special effects pipelines.

FUTURE HISTORY ..

THE FUTURE HISTORY OF SYNTHETIC VFX

2018: Researchers continue develop AI's that automate some of the most common VFX tasks such as move matching and rotoscoping.

2020: Building on earlier models developers AI's conquer the art of Colour Grading and SFX mastering.

2015 — 2020 COMMERCIALISATION

2017: Researchers develop new AI powered character models to automate and digitise stunt people. Overall variety of AI driven cross domain VFX use cases increases as AI training methods improve.

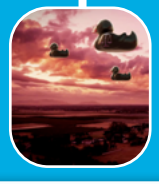
2019: Researchers continue to create more intricate facial maps and meshes and train AI's to model human attributes and behaviours.

2021: Researchers focus on using AI to automate domain and style transfer between different content types.

GAME CHANGERS TO BE AWARE OF:

EVERY FIELD we discuss in this Codex is moving FAST. It is important therefore to remember that in the main Creative Machines will one day be: Autonomous, Creative, Democratised, Free, Generalised, Imaginative, Integrated, Massively Parallel, Procedural, Scalable, Self-Evolving, and Self-Improving. For starters ...

2025: Developers have automated many of the most common VFX use cases, they start focusing attention on automating niche VFX workflows and tasks. Integration with other tech and tools accelerates.



2025 COMMODITISATION & REGULATION

2030 SOCIETAL NORM

2030: Vast majority of VFX effects have been automated by AI and are photo-realistic. AI at the edge supercharges democratisation, machines and people can auto-generate excellent VFX content.

2035: Synthetic VFX of all forms and in all formats, including immersive use cases, have become the societal norm.

2035

TECHNOLOGY DEVELOPMENT AND DIRECTION

WITH SO many different tasks in the content workflow to choose from researchers in the space who are trying to automate them are spoilt for choice. As a consequence many of them have naturally chosen to pick off and automate some of the most labour intensive and inefficient tasks to automate first and, bit by bit, they are chipping away and starting to make significant inroads across the board.

DUBBING

We've all been subjected to horrible dodgy dubbing, but thanks to Artificial Intelligence (AI) all that could be a thing of the past thanks to new several breakthroughs in the field - some that include traditional facial mapping, such as Synthesia's solution which I outline below, and others, such as Adobe and Princeton University's Text to Video technology that can generate the corresponding lip and facial movements for a subject using nothing more than a text transcript.

David Beckham, for example, does not speak Arabic, Hindi or Mandarin. But when the soccer legend starred in an advert for malaria awareness, he effortlessly switched among these and six other languages thanks to a new type of cutting edge technology that could soon change how Hollywood localises its movies and TV shows.

The ad in question was produced with technology from Synthesia, a UK startup that uses AI for dubbing. In Beckham's case, the company recorded video footage of the soccer star as well as native speakers in each of the languages it wanted to use. Then it fed all that raw footage to an algorithm that "learned" the facial expressions for each word in languages like Spanish and Yoruba, and tweaked video of Beckham's face accordingly to create a puppeteer video that actually lip-synced.

Unsurprisingly, this type of AI based video editing though has also gotten a lot of attention for its darker side, best known as the DeepFake phenomenon – porn videos that have been altered by so called deep learning based algorithms to convincingly feature the faces of celebrities, for example. That said though Synthesia's work shows that the technology also has a lot of potential for Hollywood, with dubbing being a key area of interest for the glitterati.

Traditionally, dubbing has been done without altering the source video with local script writers aiming to match translated dialogue to the action on screen - and it takes a lot to prepare those scripts. Furthermore, voice actors have to time their delivery perfectly to make sure the dub doesn't feel off, a process that can take weeks for blockbuster movies. And in one fell swoop, almost anyway, these new developments change all that. Now an AI can tweaks an actor's mouth movements to fit the local language, whatever that is, and that could significantly shorten those timelines and make dubs even more accurate.

It's no surprise therefore that Disney is also dabbling with the technology, citing that if they could change lip movement to post production then that would be a huge deal, as well as have the upside that it points a way to make lip-sync

Notes:

indistinguishable from what the actor would do if he or she actually spoke that language.

Dubbing is becoming ever more important for Hollywood as media companies target consumers worldwide with their own streaming services. Netflix, for instance, is dubbing in 31 languages, and began to target English language audiences with dubbed versions of foreign originals last year so naturally the streaming giant has also been exploring methods for automating some aspects of its dubbing work.

Despite some major advances in the field though there are still technical issues. Synthesia's technology, for example, works best when actors look directly into the camera and is less effective with profile shots or action scenes. Teaching the algorithms the facial idiosyncrasies of each individual actor is also challenging because while an algorithm can do a good job of making it look like someone's mouth forms foreign language words, it doesn't know how an actor such as Leonardo DiCaprio would deliver a line in Mandarin.

Employing AI for dubbing will be less demanding in animation though, where the technology can also be used to better match the mouth movements of a character to the original language, and as the developments keep coming many think that within the next couple of years we could see the first real applications of this exciting technology.

MOTION CAPTURE AND STUNTS

Trying to create high quality character motion in videos is difficult, very difficult – even after decades of trying with CGI. But recently a team of researchers in the US from Berkeley managed to develop a new AI reinforcement learning model, called DeepMimic, that overcame the challenges inherent in existing techniques, and that was able to realistically simulate the simple and more complex movements of characters, that ranges from people, doing things such as back flips and stunts, to dinosaurs and dragons, doing other things, so that all their movements looked realistic.

Furthermore, they also managed to eliminate some of the erroneous behaviours such as creating characters with asymmetric gaits, excessive limb movement, jitters, and unnatural behaviours that plagued previous models.

The team accomplished this amazing feat by first drawing on the wealth of physics-

Notes:

based character animation information in the computer graphics world where the physics-based simulation of natural movements and the creation of so called "controllers," that direct the movements, have been a subject of intense development for decades - especially in the gaming, movie and VFX industries where a huge amount of emphasis has been placed on trying to get motions right – and then by combining this corpus of data with their RL models.

That said though, when it came to the controllers the team quickly discovered that many of them were only capable of generating specific movements for specific tasks or characters, and that consequently they couldn't handle more complex or dynamic movements, such as stunt work.

In order to solve this problem the team drew inspiration from both of these fields, the old CGI world, and the new AI world, and took advantage of the generality afforded by deep learning models that let their models generate natural behaviours that rivalled the state-of-the-art in full body motion simulation found in the computer graphics industry.

The result was that they created a conceptually simple RL framework that enabled simulated characters to learn highly dynamic and acrobatic skills from reference motion clips, and given a single demonstration of a skill, such as a spin-kick or a backflip their characters, which included everything from people to dragons, were able to learn a robust policy to imitate the skill in simulation to produce RL policies that produced motions that were nearly indistinguishable from real life motion capture.

ROTOSCOPING

Back in 2018, inside an old auto body shop in Silicon Valley a gentleman called Stefan Avalos pushed a movie camera down a dolly track. He and a small crew were making a short film about self-driving cars. They were shooting a rusted blue 1963 Austin Mini that, via special effects, would be transformed into an autonomous vehicle that looked more like the DeLorean from the movie "Back to the Future."

Stepping back from the camera at the time Avalos referred wryly to the movie he was filming as "Project Unemployment" because the film he and his crew were making was a way of testing new technology from a start-up called Arraiy, a company that at the time was trying to automate the creation

Notes:

of digital effects for movies, television and games – all tasks that today are as tedious as they are laborious – using a new Artificial Intelligence (AI) system.

The system in question was also being developed by software giants Adobe and AutoDesk, as well as in other technology industry research labs, including Nvidia, and ultimately, the hope at the time, which has been born out since, was that once the technology was good enough and embedded into studio workflows it would spell the end to all the drudgery associated with creating and refining special effects – as well as many of the people involved in the process. As is the way with automation.

Since the advent of Jurassic Park back in the 80's, Computer Generated Imagery (CGI) has transformed how movies and TV series are made. But for all CGI's promise creating digital effects is still a painstaking and enormously tedious process. For every second of movie time armies of designers can spend hours isolating people and objects in raw camera footage, digitally building new images from scratch and combining the two as seamlessly as possible to create the final result that you see on the big screen.

Arraiy, the company behind the system Alvaros was testing, was one of the first to use AI to automate as many of these laborious tasks as possible, and thanks to improvements in deep learning we're now at the point where these systems can edit noise and mistakes out of images as well as, as you've seen elsewhere in this codex, create a wide range of synthetic content.

Inside Arraiy's offices teams of data scientists were hard at work building computer algorithms that could learn design tasks by analysing years worth of work by movie effects houses. That includes systems that learn to rotoscope raw camera footage - carefully separating people and objects from their backgrounds so that they can be dropped onto new backgrounds.

In order to train their algorithms the team collected a decade of rotoscoping and other visual effects work from various design houses, as well as adding their own work to the collection. After filming people, mannequins and other objects in front of a classic green screen company engineers were quickly rotoscope thousands of images relatively quickly that could then be added to the training corpus, and once the algorithm was trained the team unleashed it and let it rotoscope images without help from a green screen.

Notes:

This is the work that Avalos tested while all the while thinking the work could ultimately supplant work done by his own effects house. But as he saw it one of the big upsides in his case was the fact he would no longer have to outsource the tedious work the system was helping automate to workers in other countries.

Meanwhile, elsewhere, Kognat, a company started by Rising Sun Pictures debuted its Rotobot deep learning rotoscope and compositing tool that it made available with NUKE that, similarly, thanks to its intense training is able to isolate all of the pixels that belong to a certain class into a single mask, called segmentation. The effect is the isolation of portions of the image, just like rotoscoping that can isolate all of the pixels of a single instance of a class, for example where a class could be a "person" or a "cat," and so on, into its own layer.

As hardware improves these techniques and tools will soon be able to help Special effects artists generate preliminary effects right on the movie set in real time, and help give filmmakers a better idea of what the finished product will look like – a challenge that was notably taken up by The Mill, an effects house based in London, that soon after Alvaros' test took the same system and used it to create one that could instantly apply effects that changed the look of cars in their critically acclaimed film "The Human Race."

STYLE TRANSFER

Style transfer is the process of combining the content of one image with the style of another piece of content, whether it's a static image, a game, a video, or even a virtual world, to create something new. Up until recently though if you've wanted to change the look and feel of a game, for example, you first had to access its source code, but this is no longer the case and new styles can be transfered on the fly, and the teams managed it by creating a set of parameterisable Vulkan postprocessing shaders which mean that new effects can be applied to different types of content instantly and without having to crack open any code.

In the latest examples teams have used the weights from trained TensorFlow models and exported them to a games so called "shader system" which allows the styles to be changed at runtime. The final model weights are quite small, around 512KB, and the shaders run in real time on top of the game stream.

Notes:

Translating from an illustrated, 2D piece of concept art into a fully realised game environment would ordinarily require custom texture painting, modeling, material crafting, lighting, and tuning, and now real-time artistic style transfer potentially allows developers to go straight from looking at a concept to testing it in a live, interactive game environment immediately - something that will inevitably enable rapid iteration of a video game's art style.

The real-time execution of artistic styles in this way will open up new forms of video game interaction, including the shifting of visual styles during gameplay, and allow the creation of individually customised artistic styles and in game personalisation by game players - and much more.

SYNTHETIC WORLDS

EXPLORE MORE. Click or scan me to learn more about this emerging tech.

Notes:

WHILE WE could be forgiven for just thinking about synthetic content in terms of creative machines ability to generate and copy more “traditional” media types, such as imagery and video, it would be remiss of us to not include the new types of media that are emerging, and their both obvious and less obvious uses.

As everything within our world becomes increasingly driven and tied to the digital world the fact is that the overwhelming majority of the content we consume today only exists on in what I call surface media, for example, on the web, on screens and smartphones, but over time more of the content we consume will be more immersive, whether it’s consumed during sessions held in Virtual Reality environments, or, in time, via neural interfaces where content is streamed straight to and from our minds - something that, as science fiction as that might sound, has already been demonstrated with human volunteers multiple times.

However, when we start looking at alternative use cases for simulated and immersive content there are a wide number of interesting and promising growth areas. On the one hand researchers are now turning their attention to finding new ways to easily and quickly generating immersive procedural content for gaming applications and VR environments, but increasingly one of the hottest alternative uses cases in town right now is the ability for organisations to virtually test and train new AI models and products within simulated environments, as well as innovate and iterate them. And while there are now many examples of where these so called simulation engines have been used some of the most intriguing involve using them to cram hundreds of man years of learning into minutes in order to rapidly train neural network models and robots, and using them, in combination with other technologies, tools and techniques, to accelerate product innovation - from hardware through to software - by multiples of tens of thousands.

USE CASES AND TRENDS ...

KEY USE CASES AND TRENDS

TWO WORLDS COLLIDING

Having looked into all the different synthetic content fields in depth it's this field, the Synthetic Worlds field, that's perhaps the most interesting and understated especially when we consider that eventually the technologies, techniques, and tools that we use to create these increasingly life-like environments will be the same technologies, techniques, and tools that we use to create more of tomorrow's whole bodied synthetic content, especially tomorrow's dynamic, interactive, and procedural video content.

As our ability to create, and then importantly control and tailor, these synthetic worlds improves it's easy to see how many of the same technologies, techniques, and tools will be able to be adapted to allow Artificial Intelligence to auto generate everything from future AAA computer games to ultra-realistic movies, and so on.

A great example of this is Disney's Lion King 2019 re-creation where, in a world first, the company behind the film built and filmed the entire movie, everything from the characters to the environment, in Virtual Reality using the Unreal engine - an engine that up until that point had mainly been used to develop games.

As a result of this it is inevitable that enterprises in this space, for example the gaming companies, will eventually break out of their existing industry niches and challenge not just the big studios but also every other content creator out there.

In short the digital world of gaming and simulations will collide with the world of mainstream mass market digital creators and studios.

Furthermore, when we consider the implications of such a coming together on the user experience this means that user experiences can be more intense because more information, such as haptic information, can be encoded into these environments easier than they can using other more "traditional" synthetic video technologies, techniques, and tools.

Additionally, not only will this give creators more fine grained control over each individual experience as users flip between devices but it will also allow enterprises to adjust their business models in new ways and offer tame, moderate, and intense experience packages. And much more.

TECHNOLOGY USE CASES

Virtual Worlds gives us the opportunity to both digitise the real world, and create completely new worlds with new laws and behaviours. And in both of these examples these new worlds can be used for a variety of purposes.

Firstly we have the markets where video is the exclusive content type. And secondly we have the markets where video is a complimentary component of the overall finished product:

Exclusive Use Cases

- Digital Twins
- Entertainment
- Modelling

Notes:

- Product development
- Simulations
- Synthetic Innovation
- Synthetic training
- Synthetic training datasets
- Therapeutic healthcare treatments

Complimentary Use Cases

- Augmented Reality experiences
- Mixed Reality experiences
- Education and Training
- Movie production

TRENDS BEHIND THE TREND

The impact of synthetic simulations will be nothing short of revolutionary and will impact every aspect of global culture, industry and society, and they will, arguably, be one of the biggest drivers in accelerating the global rate of change.

One of the most exciting opportunities enabled by this field, for example, is to accelerate the rate of product development and innovation by multiples of thousands by, for example, using simulations to accelerate the rate of Artificial Intelligence (AI) model development, and by training drones, robots and other products millions of times faster than we can using today's traditional techniques.

Asides from this, however, as the quality of synthetic simulations improves we could also see a day when movies are constructed and filmed in immersive Virtual Reality (VR) worlds, and that is before we discuss the convergence of the digital and physical worlds and the opportunities that entails.

Meanwhile from a trajectory perspective as with all of the synthetic content generators we're discussing researchers will continue to develop new machine models, iron out the bugs and the kinks, and eventually those machines will leave the labs, be adopted by enterprises and integrated with other products and services, and be adopted by consumers - while the regulators play catch up.

FUTURE HISTORY ...

THE FUTURE HISTORY OF SYNTHETIC WORLDS

2018: Major advances in multiple AI disciplines allow researchers to create real-enough virtual environments that go on to be used to innovate and train new products billions times faster than traditional techniques.

2020: Researchers accelerate development of biologically inspired self-evolving AI's that create, modify, and solve their own environments and problems giving creators and innovators a very powerful new tool.

2015 — 2020 COMMERCIALISATION

2017: Researchers develop new Video to Video Domain Transfer models capable of digitising and altering streaming video in real time.

2019: Researchers develop the first AI capable of synthesising realistic game environments using voice commands, and development of photo realistic Video to "X" and VR generators accelerates.

2021: There is a continued focus to build generalised AI models capable of reasoning and having a genuine understanding of how the real world works. Work on creating multi-application life-like environments accelerates.

GAME CHANGERS TO BE AWARE OF:

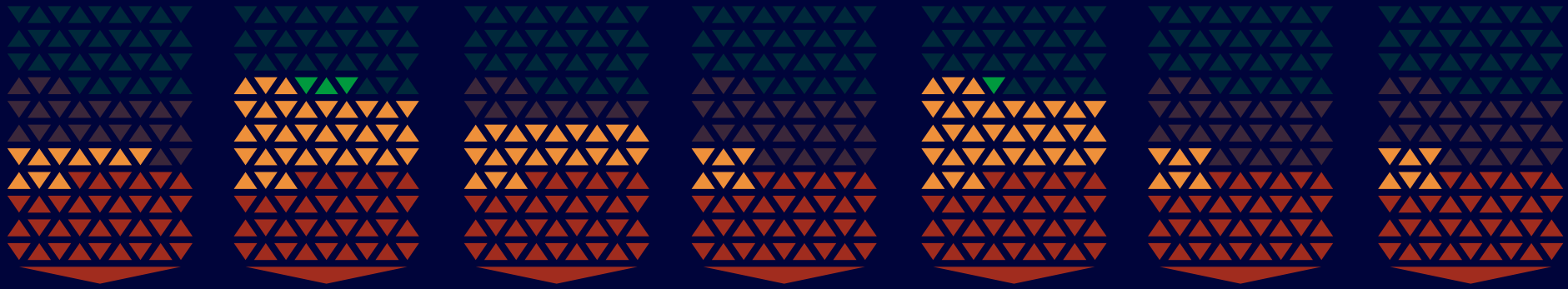
EVERY FIELD we discuss in this Codex is moving FAST. It is important therefore to remember that in the main Creative Machines will one day be: Autonomous, Creative, Democratised, Free, Generalised, Imaginative, Integrated, Massively Parallel, Procedural, Scalable, Self-Evolving, and Self-Improving. For starters ...

2025: Increasingly difficult to distinguish simulated worlds from the real thing, and the outputs are more controllable. Computer and network infrastructure performance still bottleneck commercial adoption.

2030: Simulated worlds are fast and easy to create, and are fully modifiable, albeit that they will likely still need a level of expertise to tweak. Researchers focus on expanding into new use cases.

2035: Technology is fully matured, commercialised, and ubiquitous across multiple platforms and domains. It also becomes yet another way to create procedural content of all forms, from metaverse game experiences to movies.

2025 COMMODITISATION & REGULATION 2030 SOCIETAL NORM 2035



AESTHETICS

The development of machines capable of understanding and replicating beauty.

BIOMECHANICS

The development of machines that understand biological biomechanics and principles.

CONTEXT

The development of machines capable of understanding and reproducing context.

CREATIVITY

The development of creative and imaginative machines.

DIALECT

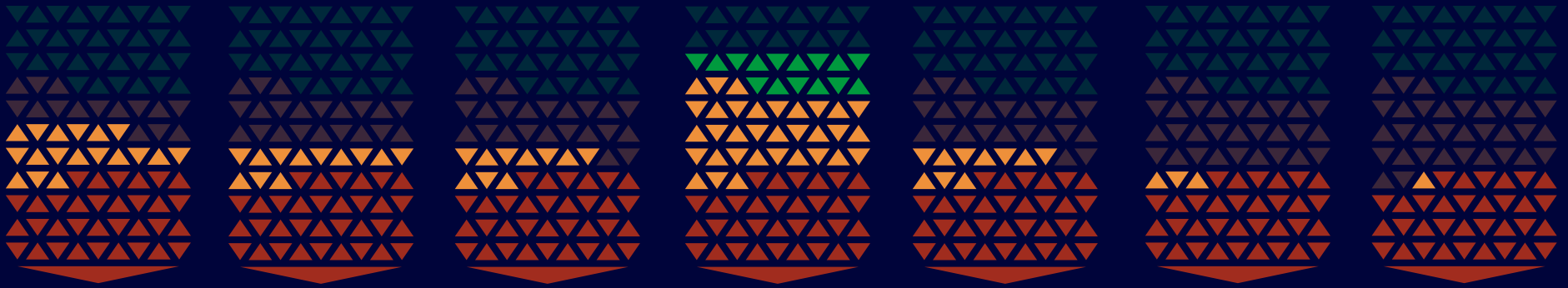
The development of machines that can understand and talk in multiple dialects.

EMOTIONAL RESONANCE

The development of machines that understand how to envoke specific emotional responses.

INTONATION

The development of machines capable of understanding and copying human intonation patterns.



LAWS OF NATURE

The development of machines that understand the laws of nature, including materials and physics.

PERSPECTIVE

The development of machines capable of understanding and replicating the nuances of perspective.

PROCEDURAL FLOW

The development of machines capable of generating content that flows.

RENDER QUALITY

The development of machines capable of rendering photo-realistic content.

SEMANTICS

The development of machines that understand implication, reference, and sense.

STORYTELLING

The development of machines capable of creating and telling great stories.

TEMPO

The development of machines capable of understanding and replicating the impact of tempo.

Figure 7. The current state of Synthetic Worlds development by AI discipline. Source: Multiple Sources

CREATING THE PERFECT GENERATOR

SYNTHETIC SIMULATIONS are, in my mind at least, one of the most fascinating areas of creative machine development today, if not only for the reason that on the one hand these are the platforms that will help us create truly expansive and immersive virtual reality environments, as well as realistic simulation environments that can be used to accelerate the development and training of future generations of AI's and robots, accelerate innovation and product development by thousands fold, and be used as the medium in which future movies, such as the recent Lion King, and other types of content are "filmed" and produced.

This is truly a revolutionary field - and for all the right reasons, but because of that it is also highly reliant on not one but many different AI disciplines coming together before researchers are finally able to create a universal and generalised synthetic simulation generator.

Furthermore the level of difficulty is increased several fold because, in short, the end goal for many researchers in this field is to create truly expansive and immersive digital environments and worlds that mirror the real world - real life, but in digital form. And better.

As a consequence, for this field to be considered fully mature there are a huge number of different disciplinary areas that need to mature and then be combined, and this is a formidable challenge. In time though, as each individual challenge is overcome the creative machines in this space will become some of the most game changing platforms on the planet.

KEY: One triangle represents 1%. The more triangles the more mature the AI discipline, and when it reaches GREEN te technology is Good Enough for adoption.

TECHNOLOGY DEVELOPMENT AND DIRECTION

WHILE THIS field is relatively new it's one that is accelerating rapidly as the amount of computing power enterprises and researchers have access to scales almost exponentially, so, as you'd expect as different teams of researchers experiment with new techniques it's also a field that is accelerating rapidly.

The primary areas of research in this field today include:

- Artificial General Intelligence
- CGI
- Simulation engines
- Text to Video generation
- Video to Synthetic Video generation
- Video to Video Domain Transfer
- Video to VR generation

So with that in mind now let's take a look at some of those technologies, tools, and techniques.

GAMING ENGINES

One of the most promising techniques at the moment involves combining streaming video input with a Video to Video synthesis AI architecture, and the Unreal Engine to create so called realistic "Semantic maps" of environments and scenes which is aided by the fact that the AI assigns every pixel on the screen with an identifier label, which then helps with post processing, and, if required gives consumers a simple way to edit and change different aspects and attributes of the environment. So, for example, in an urban environment some pixels will be labelled as "Car" and others as "Building," and so on. These clumps of pixels are then also given clearly defined edges which helps the Unreal Engine to produce a type of sketch of the environment that then gets fed into a neural network that renders the scene.

Despite the fact the field is advancing quickly though it's still likely that is will be a couple of years before these digitally rendered worlds will be ready to be incorporated into Virtual Reality worlds, and there are still many details and kinks to be ironed out before the renders become truly photo-realistic.

Notes:

CONCLUSION

PEOPLE SAY change is a constant, but in today's technology fuelled world this simple phrase is a deceiving, and often comforting, misnomer because change isn't constant, it's exponential, and the only boundaries to what we can achieve as individuals and as a global society are the ones that we invent for ourselves.

As researchers and scientists increasingly prove that nothing is impossible, that yesterdays science fiction is simply the future generations status quo, and as we all continue to bear witness to an increasingly rapid rate of change that's affecting and transforming every corner of global culture, industry, and society the future belongs to all of us equally, and we should never loose sight of that.

As you race into your own future I wish you well, and never forget you have all the friends and support you need around you as we all voyage through time and space together on this fragile living spacecraft we call Earth.

Explore More

MATTHEW GRIFFIN
Founder

Notes:

THIS IS NOT THE END.
EXPLORE MORE.

UK311-190321-DOC01