

OXFORD
BROOKES
UNIVERSITY

Consciousness, Creativity and Generative AI

Prof Nigel Crook
CiS Lecture, Bristol
3rd May 2024



Images designed by  freepik

Generative AI

ChatGPT 3.5 ▾

But in this realm of wonders, where AI meets the mind,
A question softly whispers, leaving echoes undefined.
For though the machines may create, with prowess and with grace,
Is there a spark within them that mirrors human grace?

For in the heart of consciousness, where soul and spirit blend,
Lies the essence of creation, where humanity ascends.
A sacred flame that burns within, igniting every thought,
A gift bestowed on humankind, a treasure to be sought.

So as we delve in GenAI's world, let's ponder and explore,
The depths of human brilliance, what makes us truly more.
For in the dance of creativity, where art and science meet,
It's not just what we create, but the essence, pure and sweet.



| Message ChatGPT...

Generative AI

1(a)



1(b)



Generative AI

1(c)



1(d)



Generative AI

1(e)



1(f)



Generative AI

1(g)



1(h)



Generative AI

2(a)



2(b)



Generative AI

2(c)



2(d)



Generative AI

2(e)



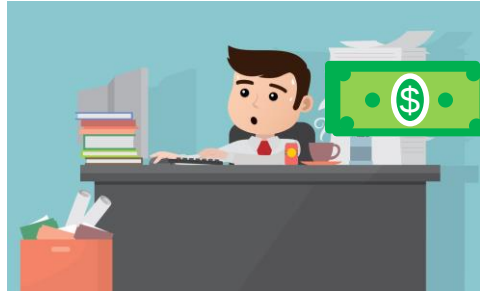
2(f)



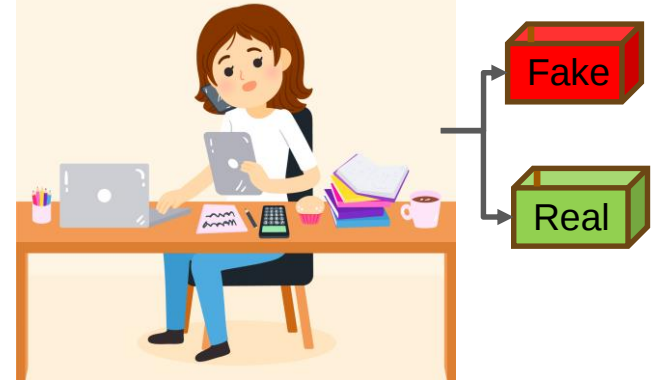
Generative Adversarial Networks - Analogy



Forger ('Generator')

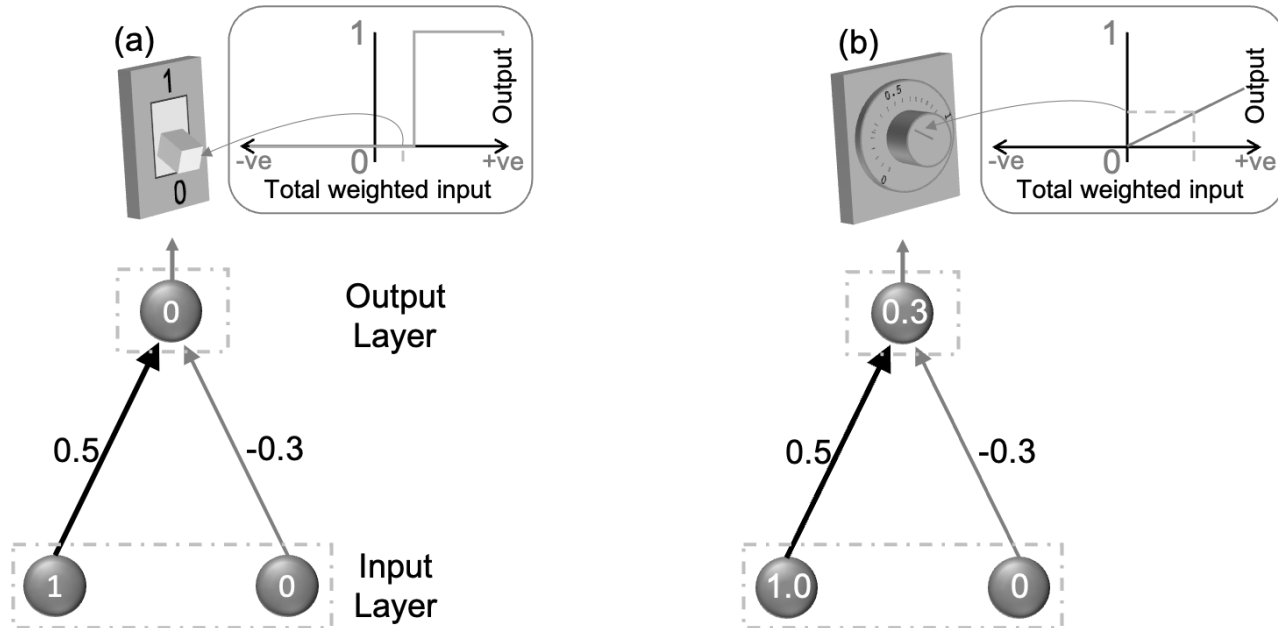


Clerk ('Discriminator')

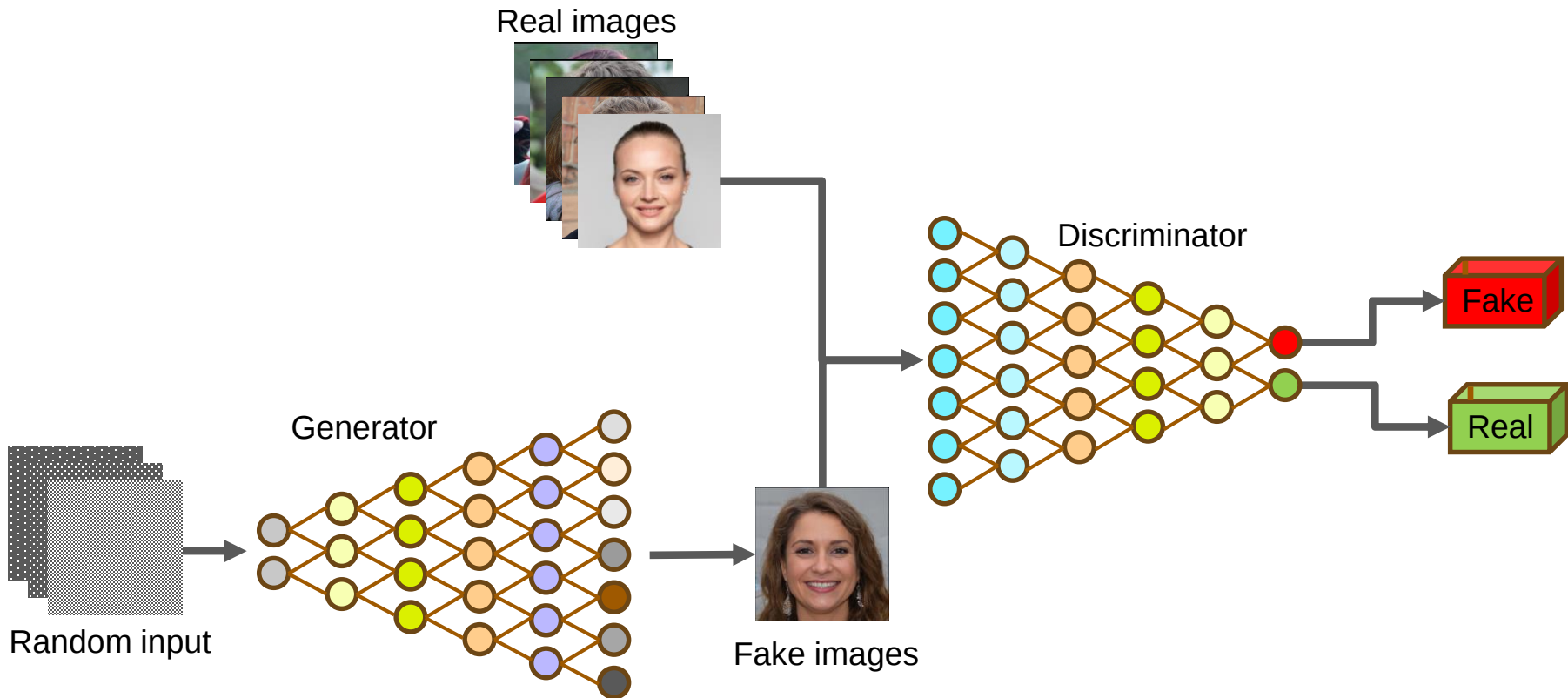


Generative Adversarial Networks

Artificial Neural Networks



Generative Adversarial Networks

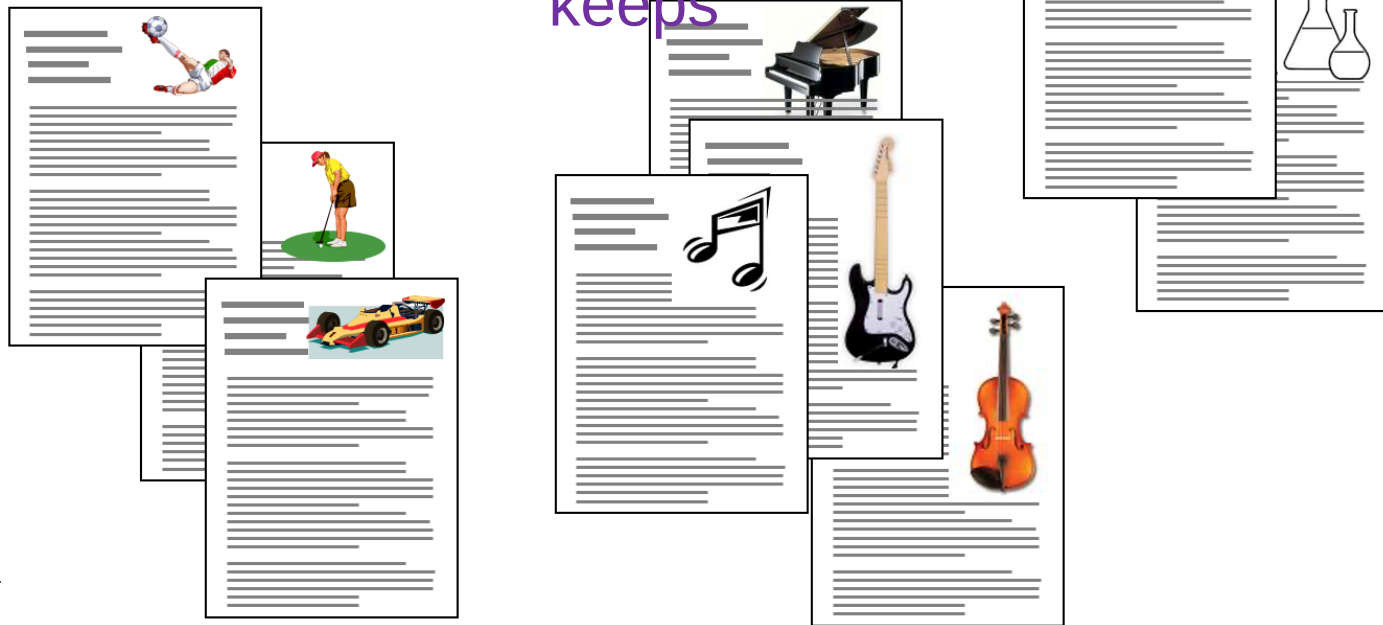


Representing meaning

Bag of words

- Counting words
 - 'Bag of words'

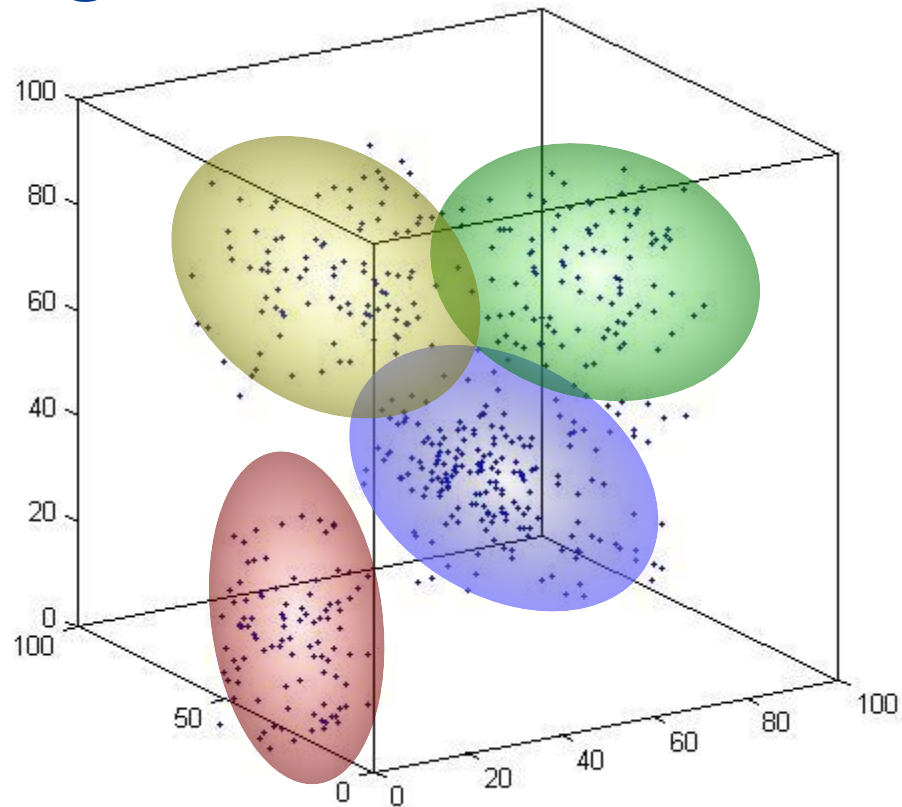
"you shall know a word
by the company it
keeps"



Representing meaning

Vector Space Models

- Counting words
 - 'Bag of words'

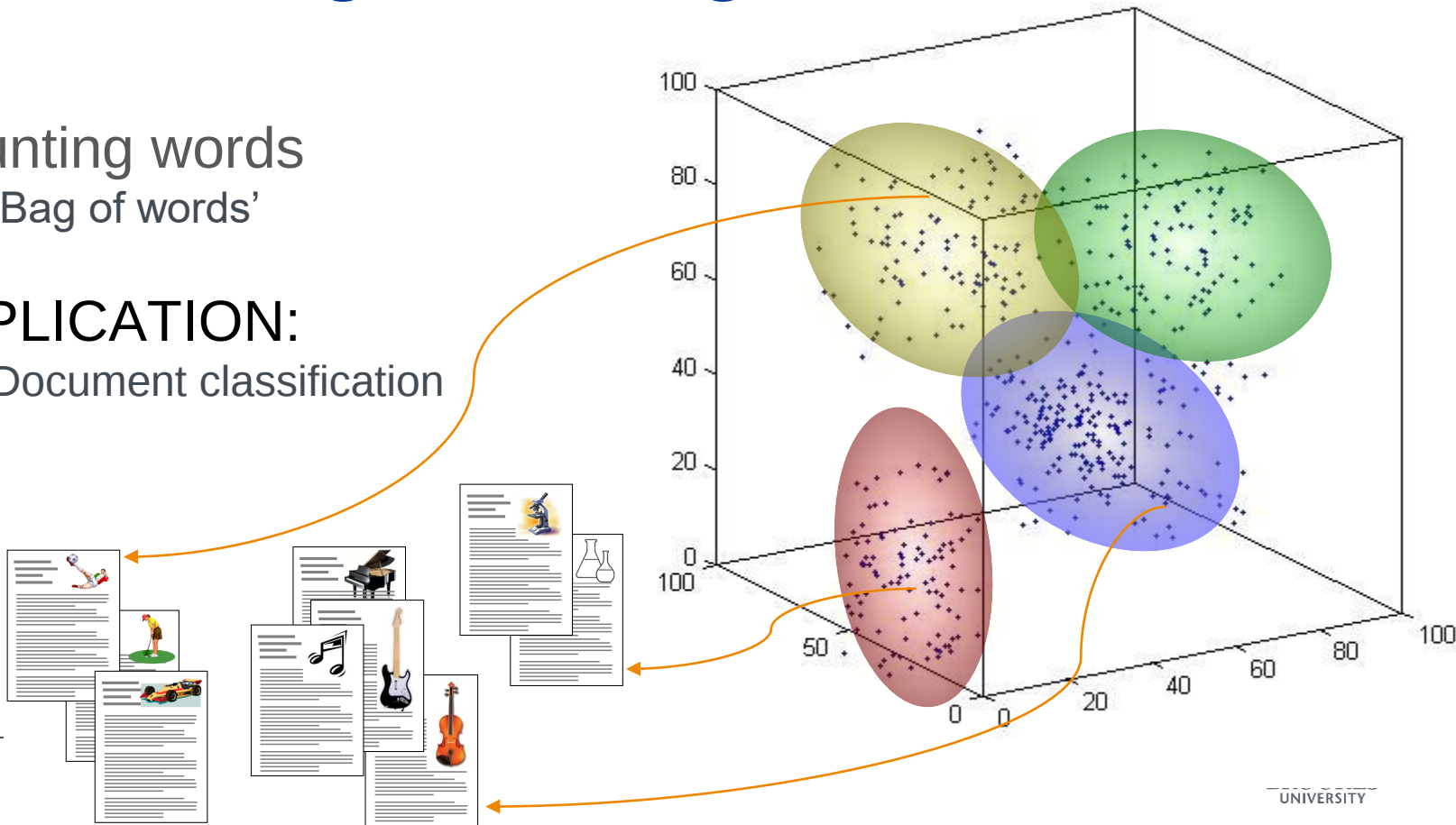


Sparck Jones, K., 1972. A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 28(1), pp.11-21.

Representing meaning

tf-idf

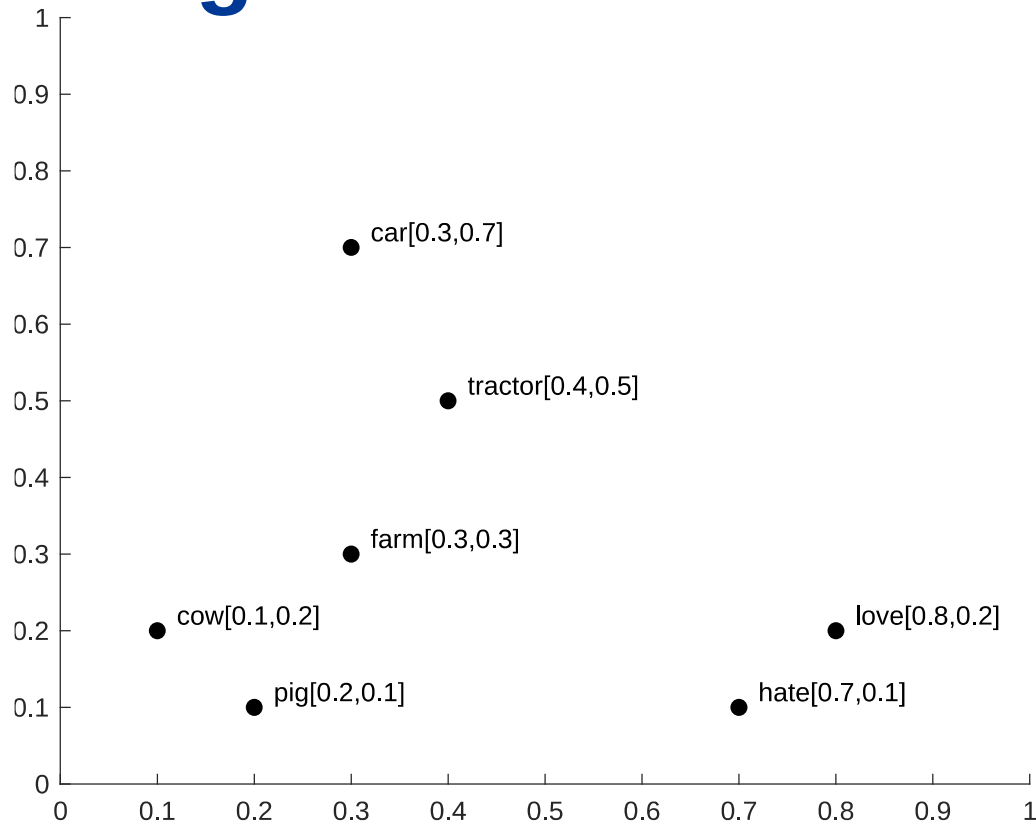
- Counting words
 - 'Bag of words'
- APPLICATION:
 - Document classification



Representing meaning

Word vectors (2013)

- Word vectors (2013)
 - Word2Vec
 - => word embeddings



Mikolov, T et al. (2013). "Efficient Estimation of Word Representations in Vector Space"

Mikolov, T et al (2013). Distributed representations of words and phrases and their compositionality.

Crook , N (2022) Rise of the Moral Machine: Exploring Virtue Through a Robot's Eyes (p56)

Language Models

- Estimate the probability of a sequences of words/tokens

$P(\text{the, moon, is, made, of, cheese}) = \text{high}$

$P(\text{cheese, of, the, is, moon, made}) = \text{low}$

$P(\text{print, (, “, Hello, ”,)}) = \text{high}$

$P(\text{Hello, }, \text{print, (, ”, “}) = \text{low}$

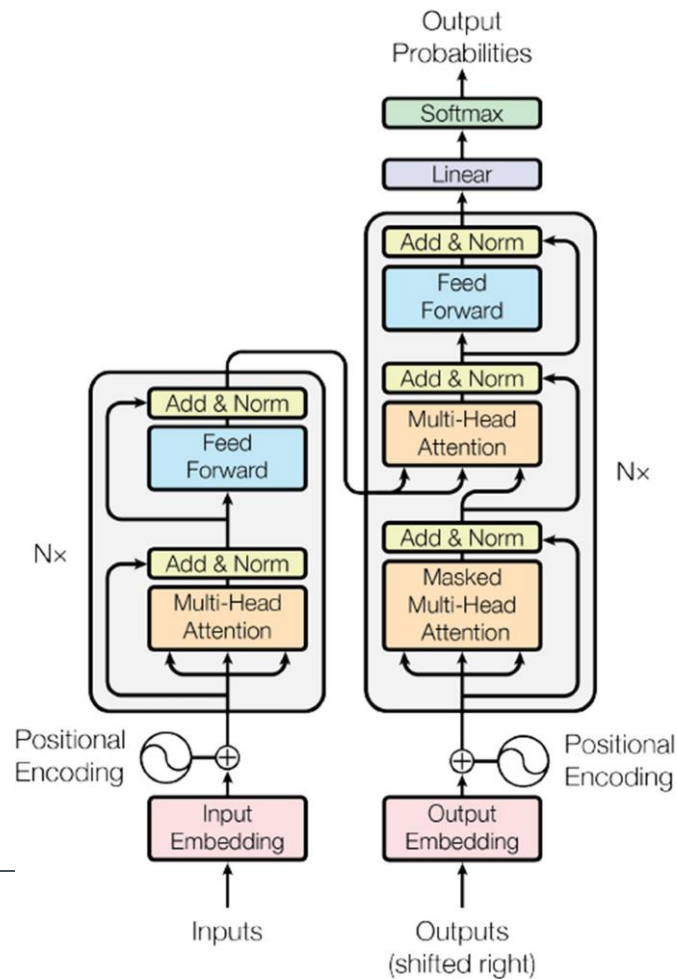
Language models

Transformers

- Google (2017)
 - “Attention is all you need”

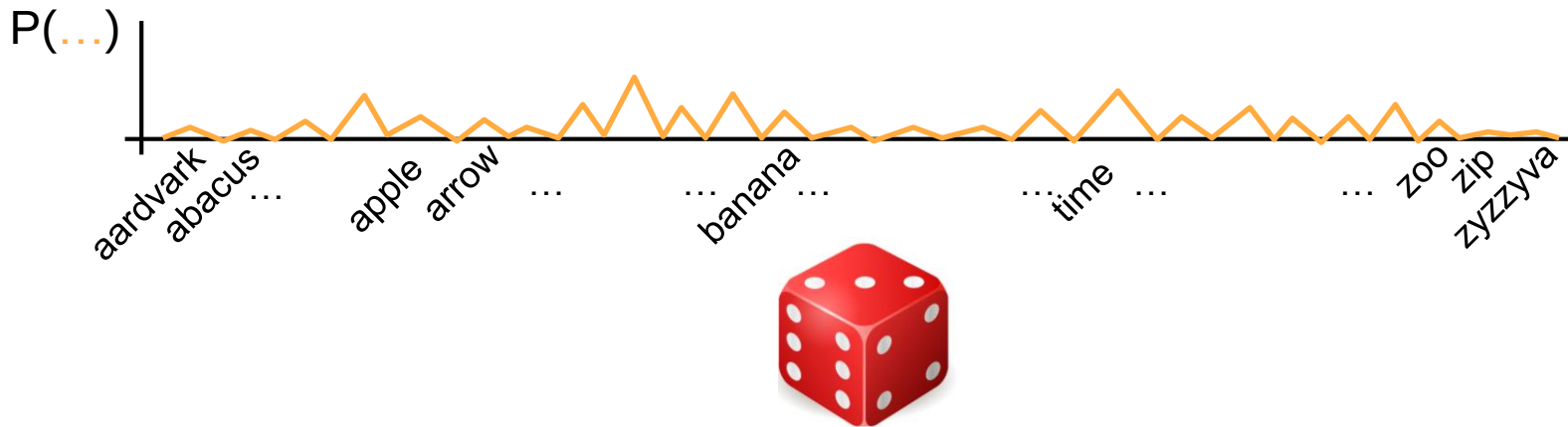
Emily awoke drowsily, head heavy, confused, eyes unable to focus. “Damn medication,” **she** muttered under her breath. As **she** lay there, a **vague unpleasant thought** took shape in her mind, like a distant memory coming into focus, a mix of **horror** and **heroism**, a **living nightmare**. Then **she** remembered the cost, and berated herself for being so ungrateful.

Vaswani, A. et al (2017) Attention is all you need.
Advances in neural information processing systems, 30



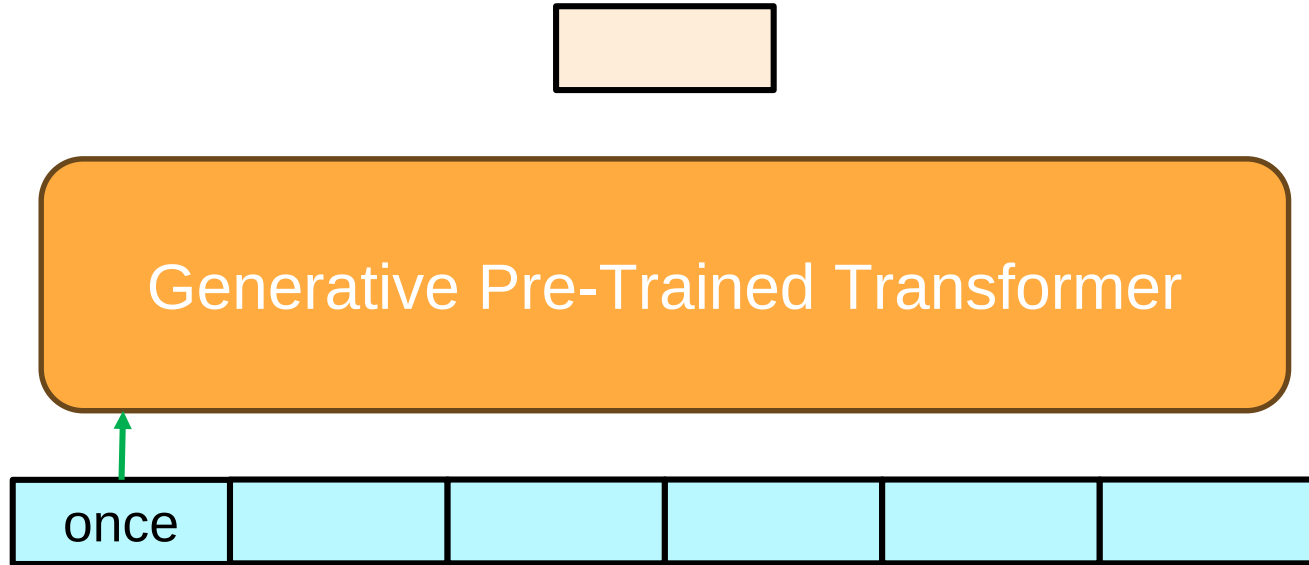
Generative Pre-trained Transformer

- GPT (2018) $P(\text{X} \mid \text{once, upon, a})$
- Next word - sampling:



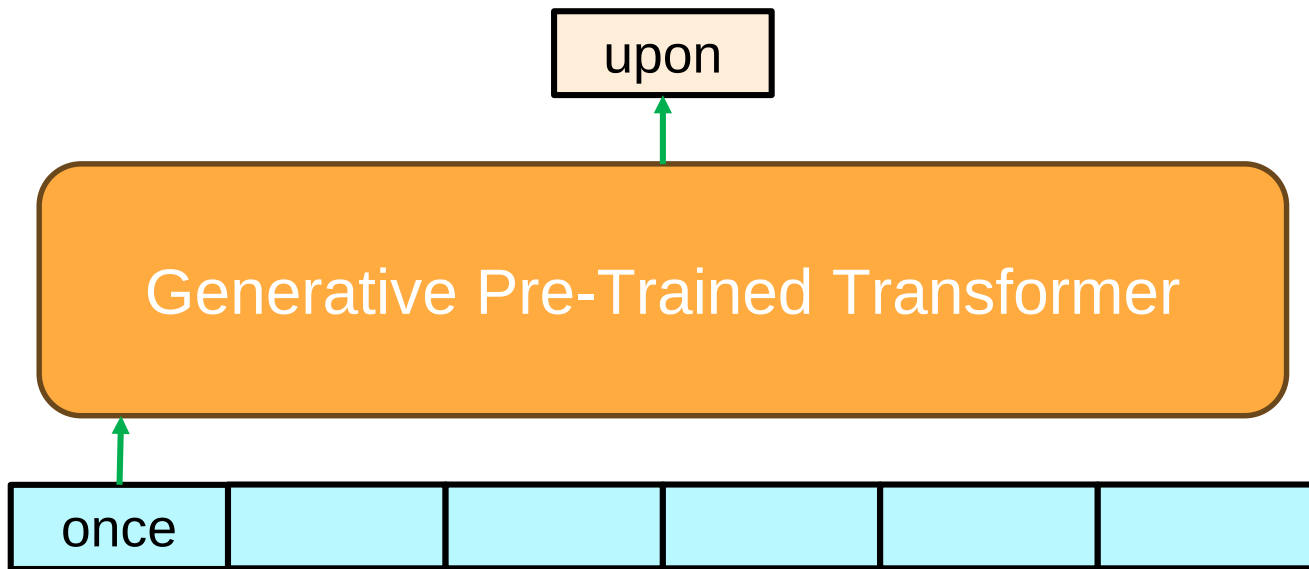
Generative Pre-trained Transformer

- GPT process ...



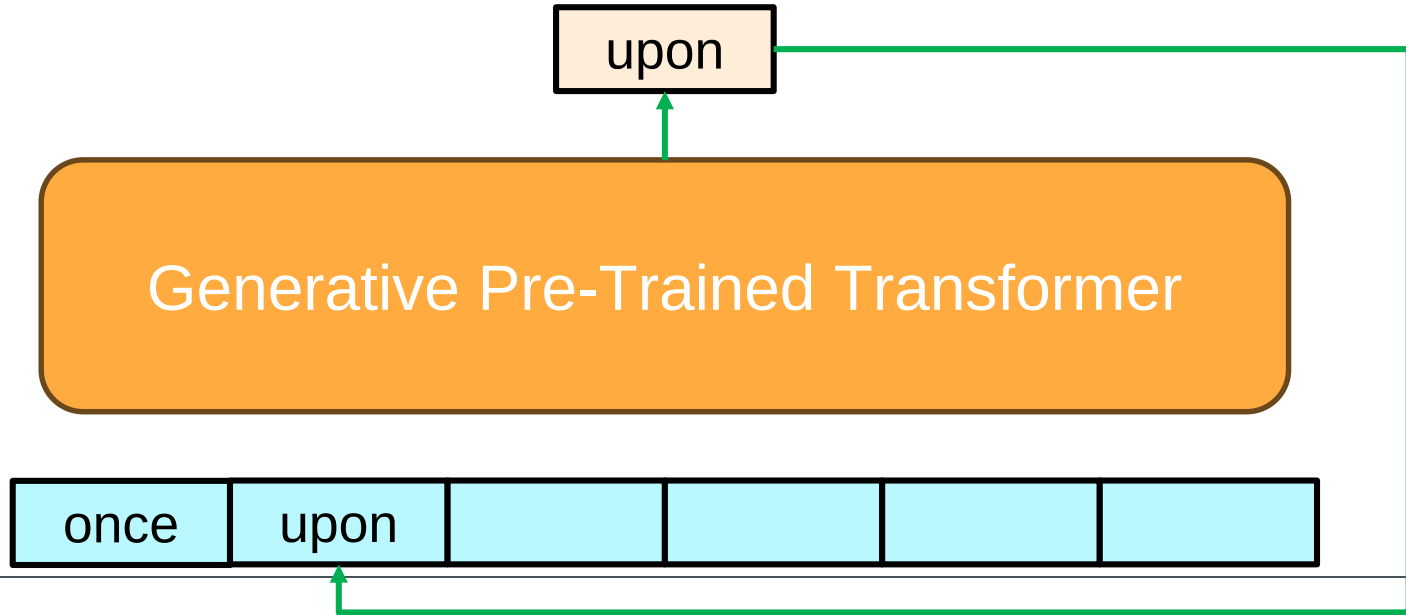
Generative Pre-trained Transformer

- GPT process ...



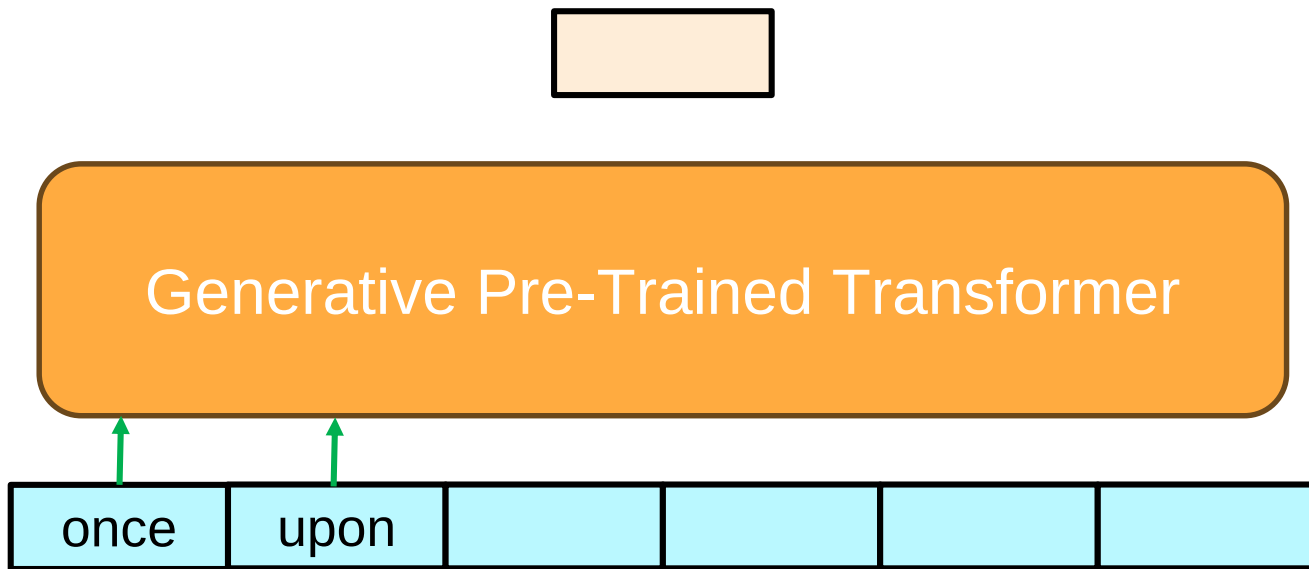
Generative Pre-trained Transformer

- GPT process ...



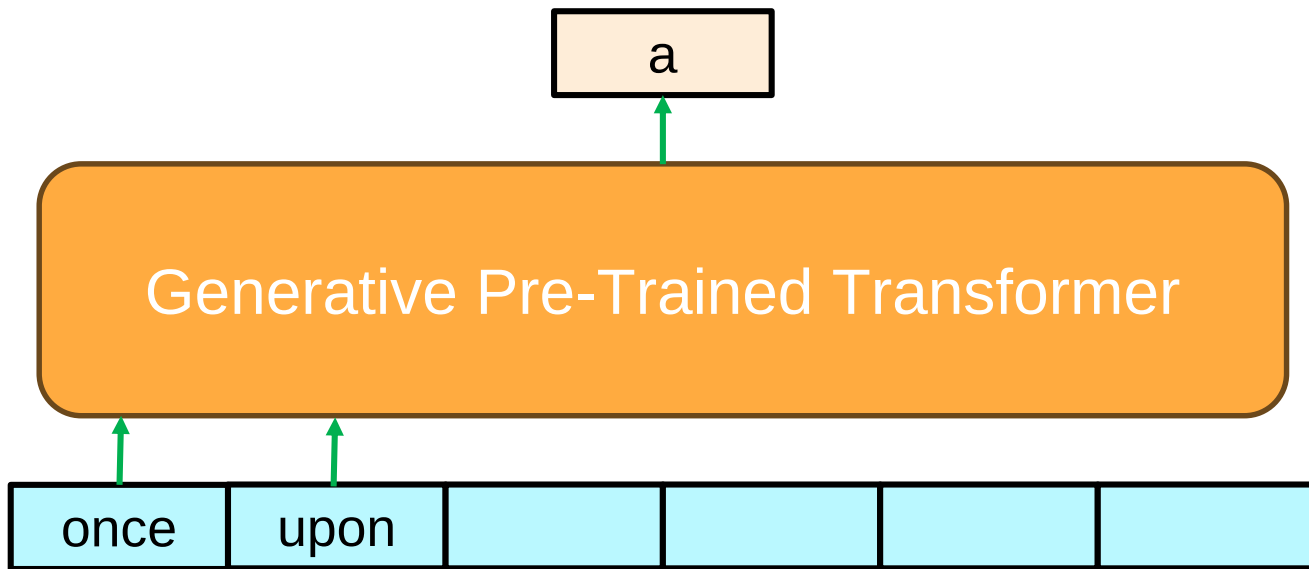
Generative Pre-trained Transformer

- GPT process ...



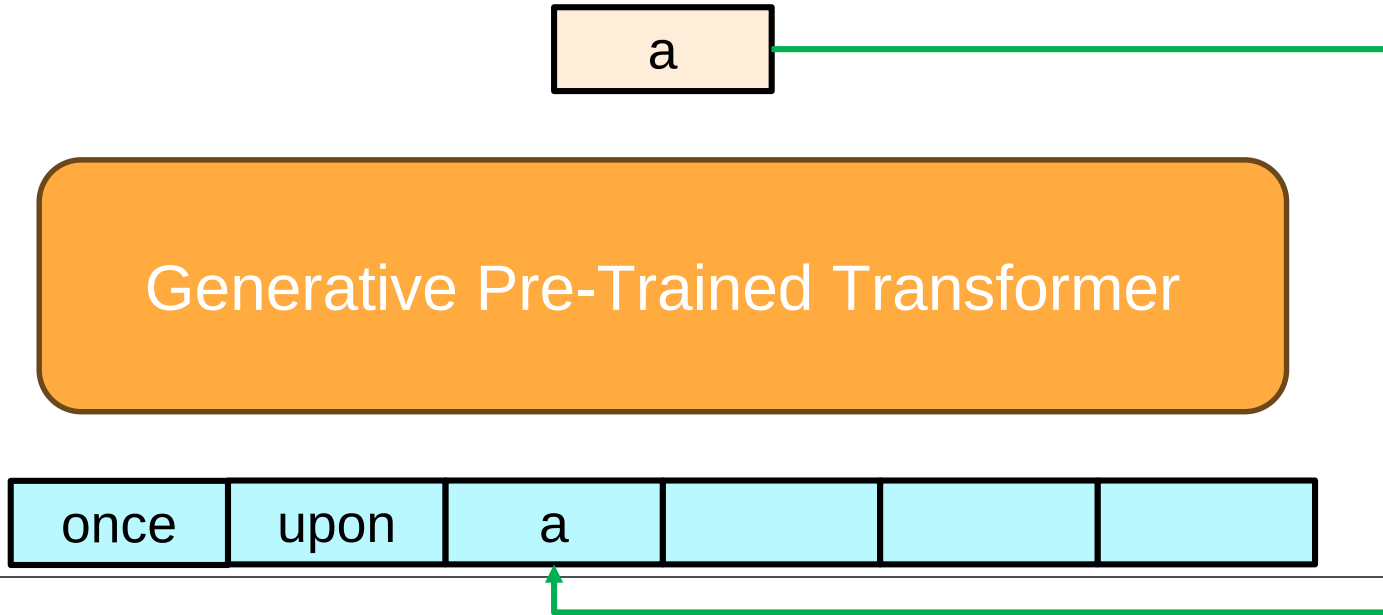
Generative Pre-trained Transformer

- GPT process ...



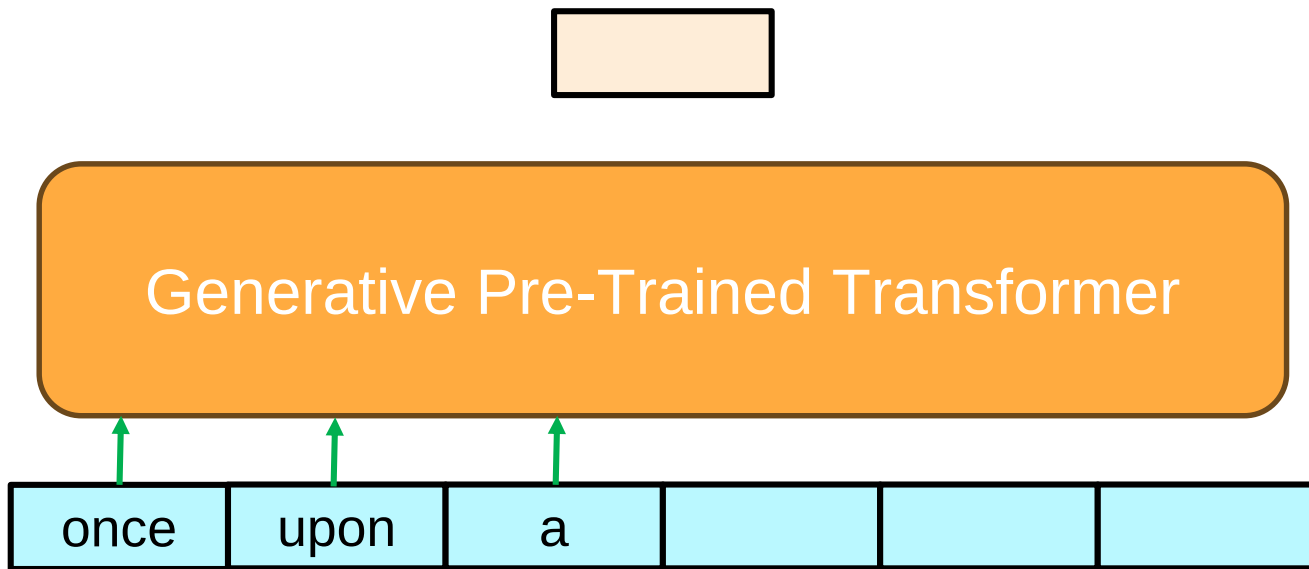
Generative Pre-trained Transformer

- GPT process ...



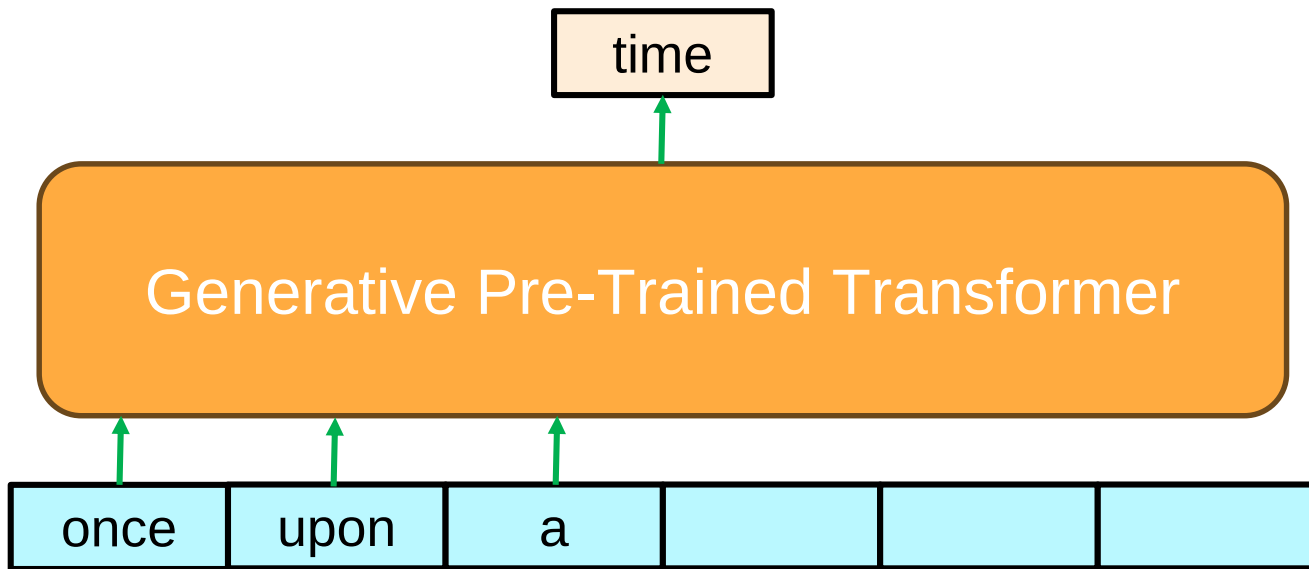
Generative Pre-trained Transformer

- GPT process ...



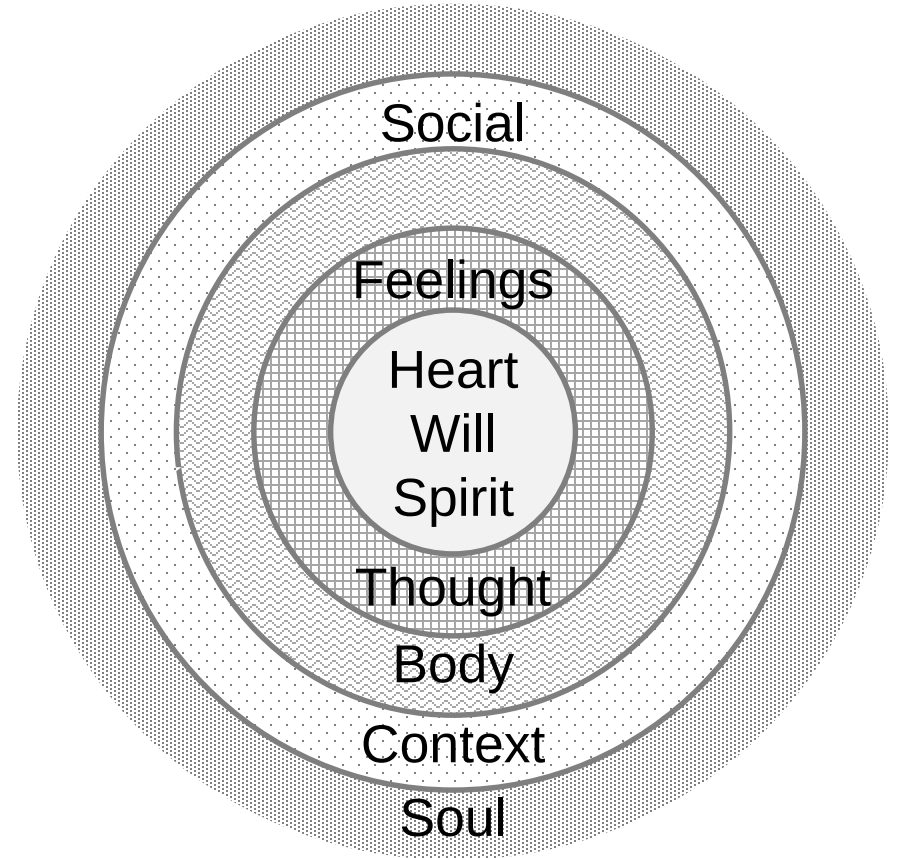
Generative Pre-trained Transformer

- GPT process ...



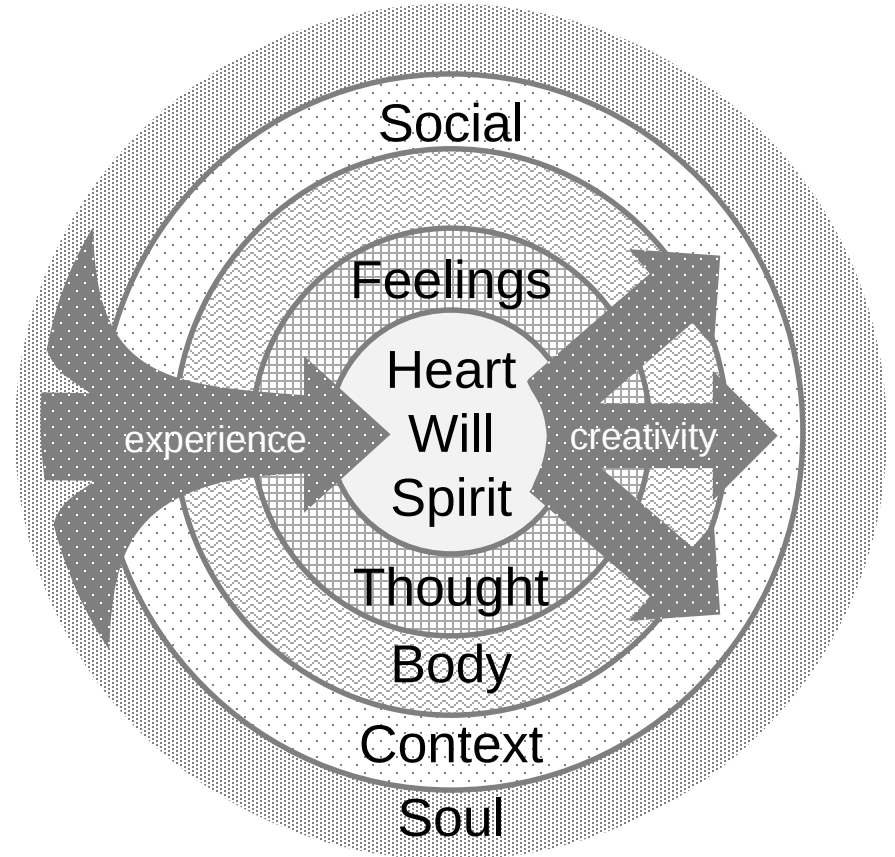
Creativity

- Creativity is the combined expression of all the essential dimensions of the human self.



Creativity

- It is fed by experience and enriched by the acquisition of artistic skill.

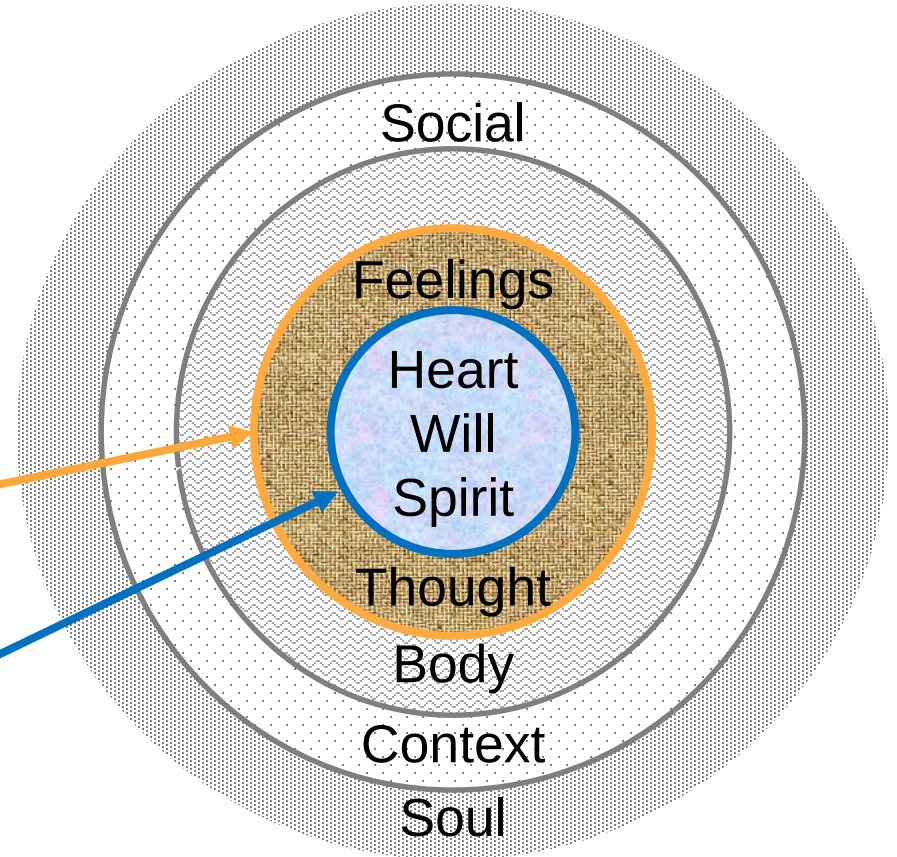


Creativity

- It is a purposeful activity, directed by conscious, intentional action

Consciousness

Free will



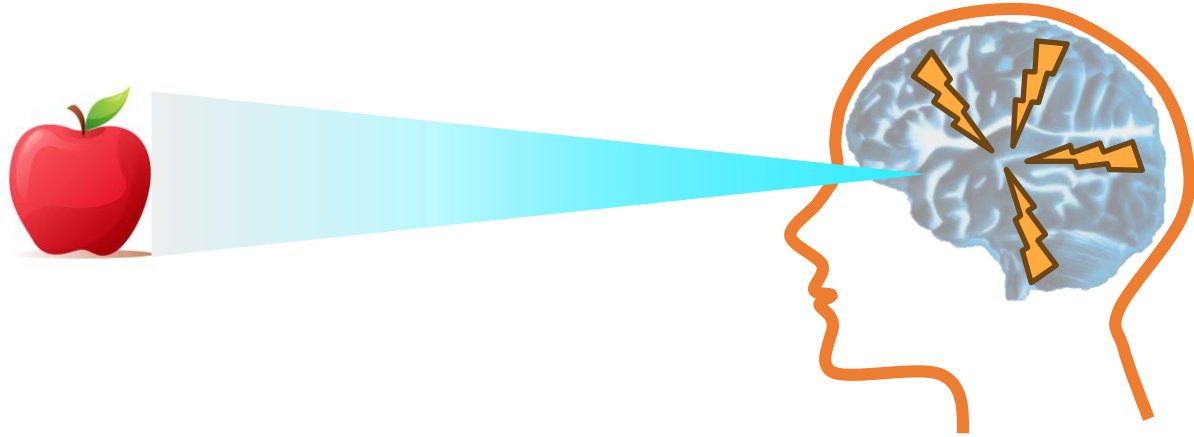
Consciousness

- Two types
 - Access Consciousness (A-consciousness)
 - Phenomenal Consciousness (P-consciousness)



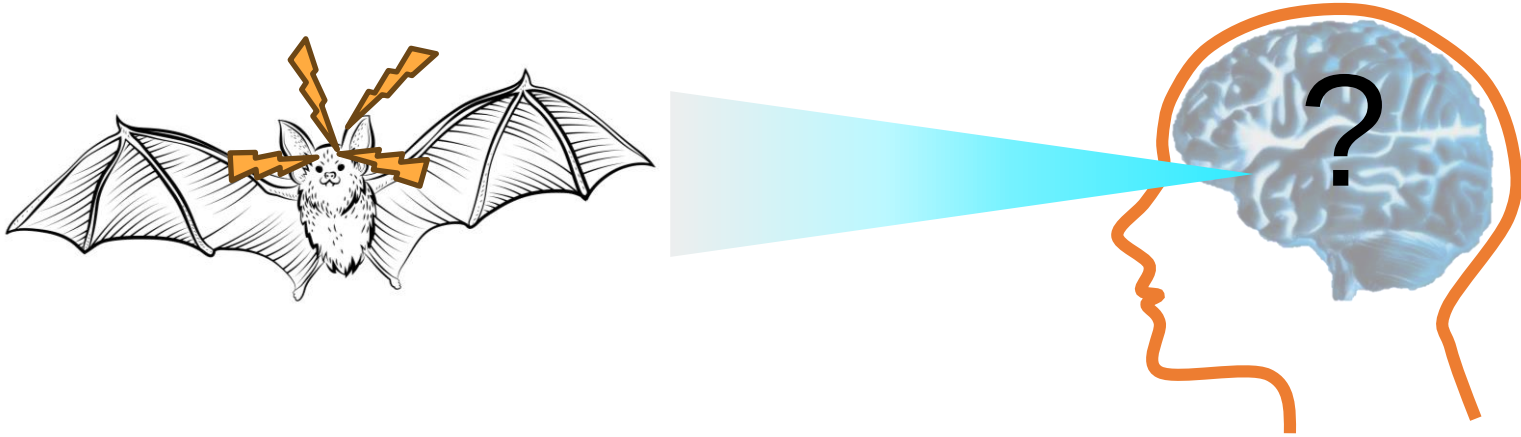
Consciousness

- Two types
 - Access Consciousness (A-consciousness)
 - Phenomenal Consciousness (P-consciousness)



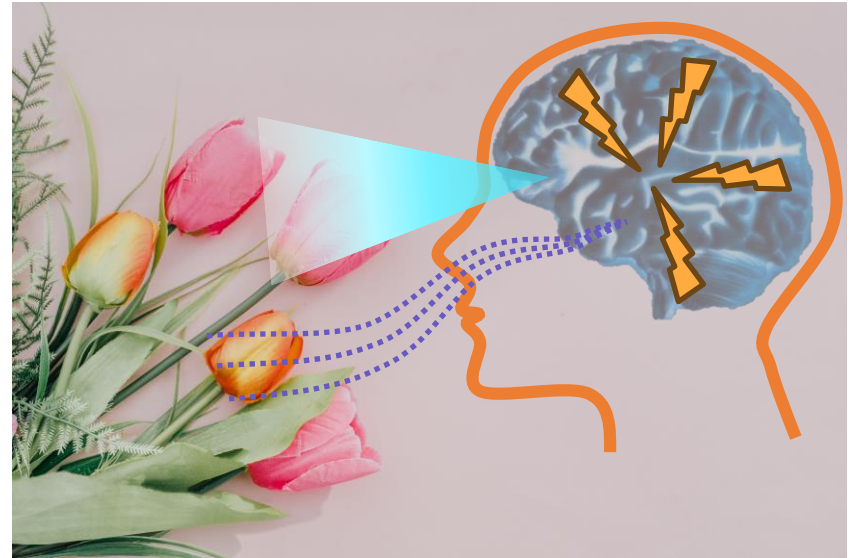
Consciousness

- Qualia
 - “What is it like to be a bat?” Thomas Nagel
 - “The Hard problem of consciousness” David Chalmers



Consciousness

- Brain gives rise to consciousness
 - Anil Seth
 - 'phi' - measuring consciousness
 - Information Theory



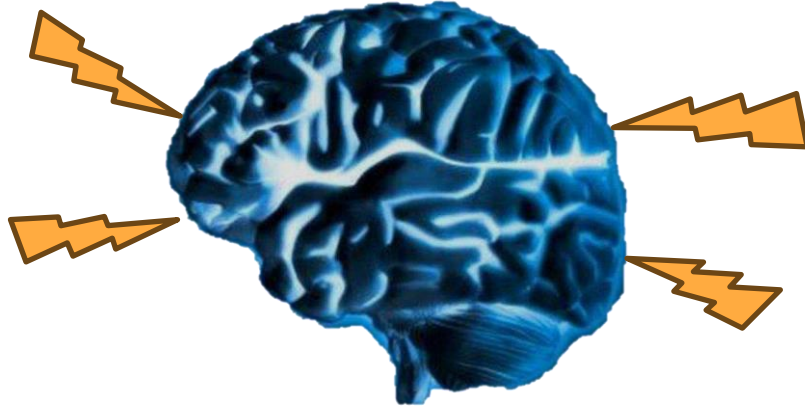
Consciousness

- Brain gives rise to consciousness

“However sophisticated the descriptions of how physical processes correlate with conscious experience may be, that still doesn’t account for how these are two very different things” - Sharon Dirckx

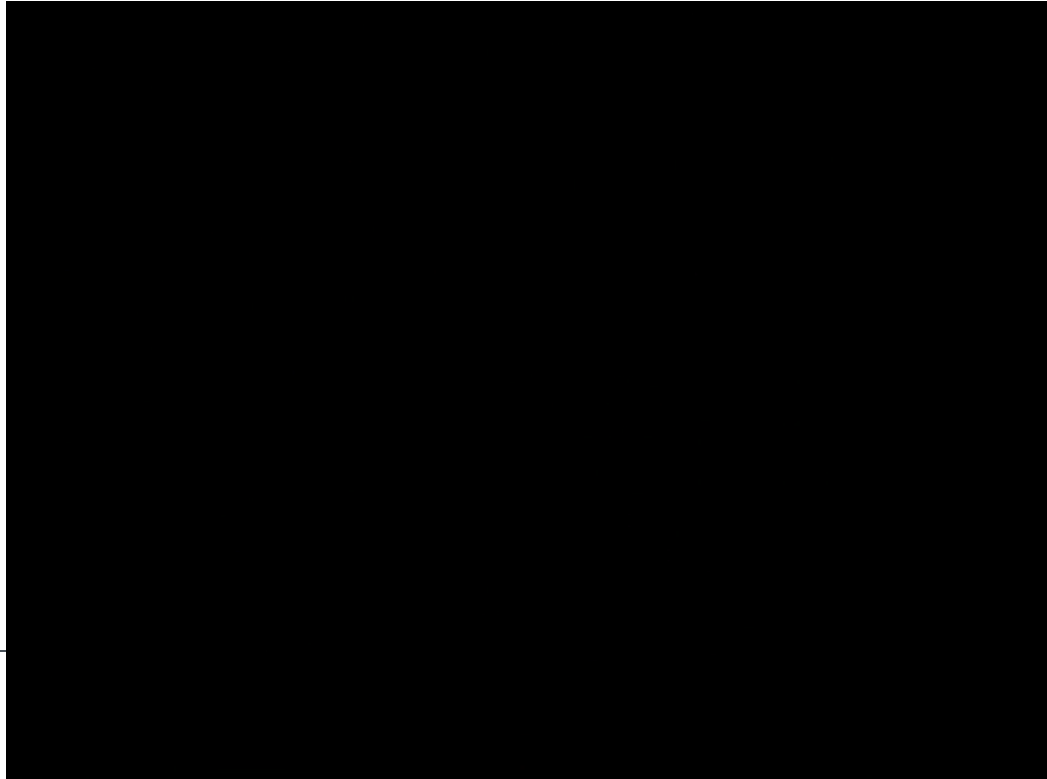
Consciousness

- Consciousness and brain processes are very different things
 - Consciousness does not appear to be a *property* of matter



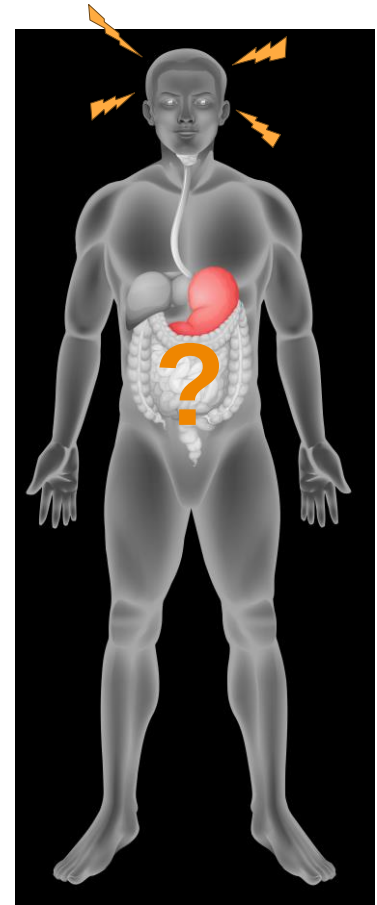
Consciousness

- Consciousness and brain processes are very different things
 - Consciousness does not appear to exist in the *behaviour of* matter



Consciousness

- Consciousness and brain processes are very different things
 - Consciousness does not appear to exist in the *behaviour of matter*



Consciousness

- Consciousness and brain processes are very different things
 - Measurements applied to matter cannot be applied to consciousness and vice versa



Consciousness

- Brain activity $\neq$ consciousness
 - Near death experience

VOLUME 32 | NUMBER 2 | WINTER 2015

JOURNAL *OF* NEAR-DEATH STUDIES

JNDAE7 32(2) 57–117 (2013) | ISSN 0891-4494

EDITOR'S FOREWORD

Janice Miner Holden, Ed.D.

ARTICLES

**The Nursing Profession and Near-Death Experiences:
A Personal and Professional Update**

Judith Mandalise, M.Ed., RN, CEN, LPC-Intern

**Reflections of Near-Death Experiences and Deathbed Visions:
A Study of Nursing Faculty's Perceptions**

Linda Moore, Ed.D., MSN, RN,

Christopher L. Pate, Ph.D., MPA, CQIA

BRIEF REPORT

**Phenomenology of Near-Death Experiences: An Analysis
of a Māori Case Study**

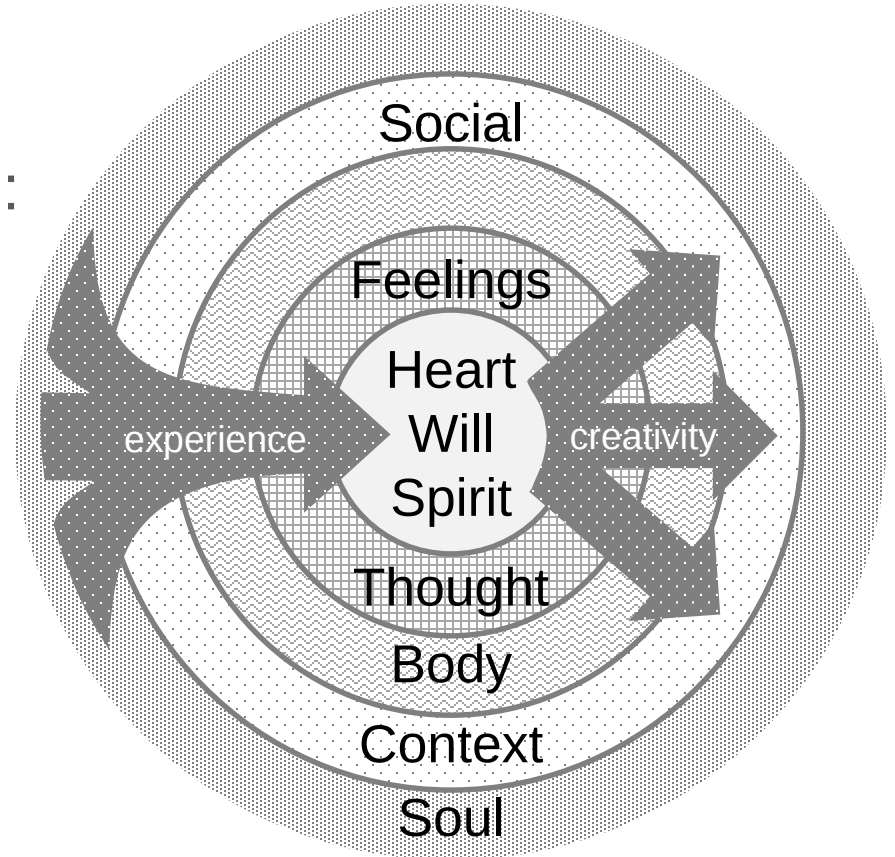
Natasha Tassell-Matamua, Ph.D.



INTERNATIONAL ASSOCIATION FOR NEAR-DEATH STUDIES

Free will

- Primary functions of the will:
 - a) to give rise to original thought or new ideas,
 - b) to choose what the mind focusses on,
 - c) to initiate the execution of what currently holds the attention of the mind.



Creativity

	Human (Authentic Creativity)	Generative AI (Simulated/pseudo Creativity)
Experience & Skill	Consciousness, learning, inspiration, practice, mentors, discovery	Examples (Big Data), 'learning', 'exploration'
Intentionality	Motivation, desire, wanting, communication	'Reward', 'goals' (states)
Representation	Expression, emotion, story	Vectors, tokens, sequences
Autonomy	Bounded free will, choices	Algorithmic, deterministic
Originality	Think outside the box, non-conformance,	Generalisation; conformation to statistical properties of data

"If you wish to get real insight into the fascinating
issues surrounding robots and morality,
then this is the book to read."
JOHN LENNOX

RISE OF THE MORAL MACHINE

EXPLORING VIRTUE THROUGH
A ROBOT'S EYES



PROF. NIGEL CROOK

CONTENTS

PREFACE		11
PART I	FOUNDATIONS	15
	SHORT STORY: Murder and Theft	17
Chapter 1	Roboethics	19
	1.1 Moral Machines	20
	1.2 Moral Agency	24
	1.3 Criteria for moral agency	29
Chapter 2	A Brief Encounter with AI	33
	2.1 Knowledge	34
	2.2 Reasoning	36
	2.3 Learning	37
PART II	MOTIVATION	43
	SHORT STORY: ethical alignment	45
Chapter 3	The Moral Machines are Coming	49
	3.1 Value Alignment	50
	3.2. Increasing Levels of Morality	66
Chapter 4	Getting Robots to Behave Themselves	71
	4.1. Moral Theories	72
	4.2. Teaching robots right from wrong	80
	4.3. Multiple Ethical Systems	91

"If you wish to get real insight into the fascinating
issues surrounding robots and morality,
then this is the book to read."
JOHN LENNOX

RISE OF THE MORAL MACHINE

EXPLORING VIRTUE THROUGH
A ROBOT'S EYES



PROF. NIGEL CROOK

Part III	MORAL AGENCY	95
	SHORT STORY: The Trial of Alex Montague	97
Chapter 5	Becoming Moral	101
	5.1. Early Moral Foundations	102
	5.2. Six Moral Stages	104
	5.3. Acquiring Virtue	112
Chapter 6	The Anatomy of Moral Agency	115
	6.1. Six Dimensions of the Human Self	117
	6.2. Virtuous Robots	131
	6.3 Machine-Based Virtue	134
Part IV	IMPLICATIONS	143
	SHORT STORY: Deep Fake Humans	145
Chapter 7	Human and Robot Autonomy	151
	7.1. Autonomous Robots	152
	7.2. Free Will	153
	7.3. Ultimate Responsibility	176
Chapter 8	The Moral Singularity	183
	8.1. What is a Singularity?	184
	8.2. The Technological Singularity	186
	8.3. The Creational Singularity	197
Chapter 9	I Compute, Therefore I Am	205
	9.1. Robot Moral Agency	206
	9.2. Deepfake Humans	218
	9.3. Authentically Human	222
	9.4. Final words	225

Recent Book

Available on Amazon

Website: ntcrook.com

